Influence of cobalamin scarcity on diatom molecular physiology and identification of a cobalamin acquisition protein

Erin M. Bertrand[a,b,1], Andrew E. Allen[c], Christopher L. Dupont[c], Trina M. Norden-Krichmar[c], Jing Bai[c], Ruben E. Valas[c], and Mak A. Saito[a,2]

[a]Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543; [b]Massachusetts Institute of Technology/Woods Hole Oceanographic Institution Joint Program in Chemical Oceanography, Woods Hole, MA 02543; and [c]Microbial and Environmental Genomics, J. Craig Venter Institute, San Diego, CA 92121

Edited* by François M. M. Morel, Princeton University, Princeton, NJ, and approved April 20, 2012 (received for review January 31, 2012)

Diatoms are responsible for ~40% of marine primary production and are key players in global carbon cycling. There is mounting evidence that diatom growth is influenced by cobalamin (vitamin B_{12}) availability. This cobalt-containing micronutrient is only produced by some bacteria and archaea but is required by many diatoms and other eukaryotic phytoplankton. Despite its potential importance, little is known about mechanisms of cobalamin acquisition in diatoms or the impact of cobalamin scarcity on diatom molecular physiology. Proteomic profiling and RNA-sequencing transcriptomic analysis of the cultured diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* revealed three distinct strategies used by diatoms to cope with low cobalamin: increased cobalamin acquisition machinery, decreased cobalamin demand, and management of reduced methionine synthase activity through changes in folate and S-adenosyl methionine metabolism. One previously uncharacterized protein, cobalamin acquisition protein 1 (CBA1), was up to 160-fold more abundant under low cobalamin availability in both diatoms. Autologous overexpression of CBA1 revealed association with the outside of the cell and likely endoplasmic reticulum localization. Cobalamin uptake rates were elevated in strains overexpressing CBA1, directly linking this protein to cobalamin acquisition. CBA1 is unlike characterized cobalamin acquisition proteins and is the only currently identified algal protein known to be implicated in cobalamin uptake. The abundance and widespread distribution of transcripts encoding CBA1 in environmental samples suggests that cobalamin is an important nutritional factor for phytoplankton. Future study of CBA1 and other molecular signatures of cobalamin scarcity identified here will yield insight into the evolution of cobalamin utilization and facilitate monitoring of cobalamin starvation in oceanic diatom communities.

micronutrient acquisition | proteomics | transcriptomics

Diatoms are responsible for an estimated 40% of marine primary production, and are therefore important players in global carbon cycling (1, 2). Although diatom growth in the oceans is thought to be controlled primarily by nitrogen and iron availability (3, 4), recent studies (5–8) support long-standing hypotheses that cobalamin availability can have an impact on marine phytoplankton growth and community composition (9–12). In the open ocean, cobalamin is present in exceedingly low concentrations and is depleted in irradiated surface waters, largely as a result of biological utilization (11). Because no eukaryotic organism is known to produce cobalamin (13), marine bacteria and archaea must supply auxotrophic (vitamin-requiring) phytoplankton with the vitamin, either through direct interaction (14) or through production and release into the water column upon death and cell lysis (15, 16). This chemical dependency is one of many that underlie interactions among marine microbial groups; assessing the role of these dependencies in oceanic processes is of considerable interest (17). Cobalamin availability may play a significant role in the climatically important Southern Ocean, where it appears to colimit the growth of diatom-dominated phytoplankton communities periodically (6) and is likely in short supply relative to other marine environments (18, 19).

The available genome sequences of marine diatoms (*Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, and *Fragilariopsis cylindrus*) lack proteins homologous to known metazoan and bacterial cobalamin acquisition proteins (20). As a result, the mechanisms by which these phytoplankton acquire the vitamin from their environment remain unclear. Cobalamin requirements in eukaryotic algae like diatoms arise primarily from its use in the enzyme methionine synthase (14, 21). Methionine synthase is responsible for generating methionine and tetrahydrofolate (THF) from homocysteine and 5-methyltetrahydrofolate, thus playing an essential role in cellular one-carbon metabolism (22). Some eukaryotic algal genomes encode only one version of this enzyme, cobalamin-dependent methionine synthase (MetH), which uses methylcobalamin as an intermediate methyl group carrier (23). These algae thus have an absolute cobalamin requirement. In contrast, other algal strains encode MetH as well as cobalamin-independent methionine synthase (MetE), an enzyme that accomplishes the same reaction as MetH but without cobalamin and with much lower efficiency (24). Organisms with MetE and MetH thus have a flexible cobalamin demand and use cobalamin when available but do not absolutely require it. Maintenance of the much lower efficiency MetE enzyme in phytoplankton genomes presumably allows for ecological flexibility in environments with scarce or variable cobalamin availability (21).

Because of its influence on methionine synthase activity, low cobalamin availability to phytoplankton has the potential to have an impact on a wide range of cellular and ecological functions. Methionine serves not only as a protein-building amino acid but as the precursor to S-adenosyl methionine (AdoMet or SAM), an important methylating agent, propylamine donor, and radical

Author contributions: E.M.B., A.E.A., and M.A.S. designed research; E.M.B., C.L.D., and J.B. performed research; E.M.B., A.E.A., C.L.D., T.M.N.-K., and R.E.V. analyzed data; and E.M.B., A.E.A., and M.A.S. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

Data deposition: The RNA sequencing data reported in this paper have been deposited in the National Center for Biotechnology Information's Sequence Read Archive (SRA) (accession nos. SRA051681-7, 051693-4, 052788, 052867, 052898, 052921-2, and 052955-7). Sequences used to prepare Fig. 5 were deposited in GenBank (accession nos. JX042584– JX042668).

[1]Present address: Microbial and Environmental Genomics, J. Craig Venter Institute, San Diego, CA 92121.

[2]To whom correspondence should be addressed. E-mail: msaito@whoi.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1201731109/-/DCSupplemental.

source. Methionine is also required for dimethylsulfonium propionate production, which is used by a subset of diatoms possibly as a cryoprotectant, osmolyte (25), or antioxidant (26), and is the precursor to the climatically important gas dimethylsulfide (27). In addition, impaired methionine synthase activity prevents efficient folate recycling, which has important implications for nucleic acid biosynthesis (14, 28).

Here, we used MS-based proteomics and RNA-sequencing (RNA-seq) transcriptomics to identify diatom gene products involved in cobalamin metabolism and to assess the overall impact of cobalamin scarcity on diatom molecular physiology. This study has yielded the identification of a unique protein involved in cobalamin acquisition and characterized multiple molecular level responses to cobalamin scarcity in diatoms.

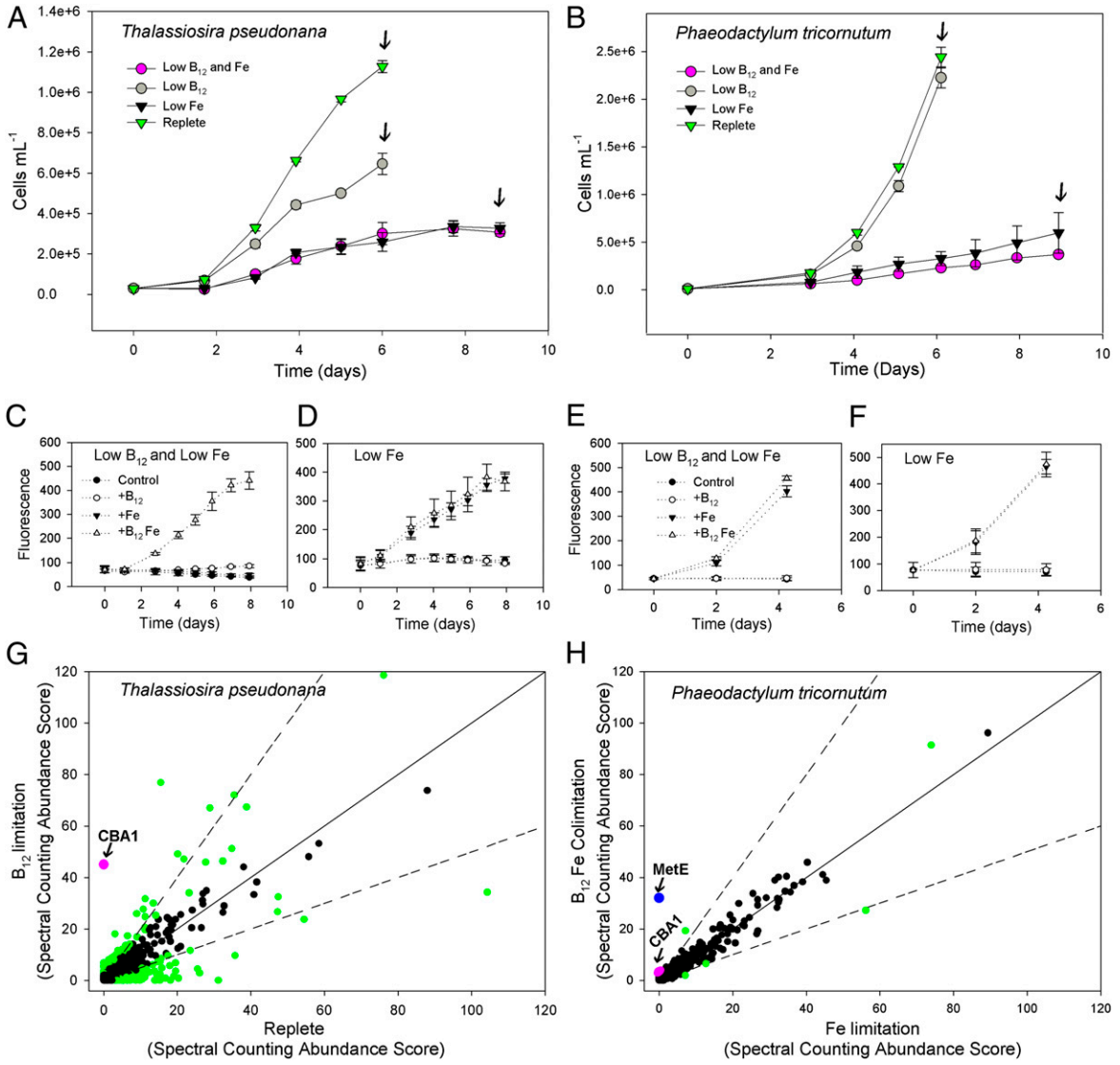

Fig. 1. Effect of low cobalamin and iron availability on growth and protein expression in two diatoms. Cell density over time for *T. pseudonana* (*A*) and *P. tricornutum* (*B*) grown under four different nutrient regimes: low vitamin B_{12}, low Fe, low vitamin B_{12} with low Fe, and replete. Values shown are means of triplicate cultures, and error bars are 1 SD. Arrows indicate where samples for proteomic and transcriptomic analyses were taken. Low cobalamin availability had a much larger impact on *T. pseudonana* growth than on the growth of *P. tricornutum,* likely attributable to *P. tricornutum*'s use of MetE as an alternative to the vitamin B_{12}-requiring MetH. Because of the low iron concentrations used, iron limitation had a more severe impact on growth than low vitamin B_{12} in this the experiment. (*C*–*F*) Limitation scenarios were verified by resupplying cultures with cobalamin and iron. Fluorescence over time is shown as means of single measurements of triplicate cultures, with error bars representing 1 SD. Each culture in *A* and *B* (at arrows) was split in four and resupplied with nothing (control), vitamin B_{12}, iron, or vitamin B_{12} and iron together. (*C*) Growth in low vitamin B_{12} with low iron cultures of *T. pseudonana* was only rescued by the addition of both vitamin B_{12} and iron together; fluorescence in +Fe cultures and +vitamin B_{12} and Fe cultures was significantly different (Student's paired *t* test, $P = 0.0004$; time = 8 d). (*E*) Iron addition alone rescued growth in low vitamin B_{12} and iron *P. tricornutum*, and coaddition of vitamin B_{12} and Fe together further enhanced growth; fluorescence in +Fe cultures and + vitamin B_{12} and Fe cultures was significantly different (Student's paired *t* test, $P = 0.0007$; time = 4 d). Shotgun MS analyses of *T. pseudonana* (*G*) in the low vitamin B_{12} vs. replete treatment and *P. tricornutum* (*H*) in the low vitamin B_{12} with low iron treatment vs. low iron treatment reveal vitamin B_{12}-responsive proteins. Each point is an identified protein, with the mean of its technical triplicate abundance scores in one treatment plotted against the mean of abundance scores in another treatment. The solid line is 1:1 abundance, and the dashed lines denote 2:1 and 1:2 abundances. Protein CBA1, Tp11697 and Pt 48322, is highlighted in pink, and MetE is highlighted in blue. Proteins in green are considered differentially abundant (Fisher exact test, $P < 0.01$).

Results and Discussion

We used quantitative proteomic and transcriptomic methodologies to investigate the response of two distantly related diatoms, *T. pseudonana* and *P. tricornutum*, to low cobalamin availability. *T. pseudonana* has an absolute cobalamin requirement because its genome encodes only the vitamin B_{12}-dependent methionine synthase MetH. *P. tricornutum* has a flexible cobalamin demand because its genome encodes both MetE and MetH. To facilitate the identification of gene products that specifically respond to low cobalamin supply in these diatoms, we used a multifactorial experimental design to evaluate their response to low cobalamin under conditions of both high and low iron availability (Fig. 1). Gene products displaying consistent responses under low cobalamin, regardless of iron nutritional status, were identified as specifically involved in cellular responses to cobalamin scarcity. The low iron treatment provided a control against which to judge whether gene product abundance changes were more general to stress or low growth rate, or whether they displayed a specific response to low cobalamin. Iron was chosen over other nutrients for this factorial design because simultaneous iron and cobalamin limitation has been documented in field communities of diatoms (6, 29). Therefore, any efforts to develop molecular markers for low cobalamin availability that can potentially be implemented in environmental samples must consider the impact of iron starvation on those markers.

The growth of diatoms under these limitation scenarios is described in Fig. 1 *A* and *B*. Proteome and transcriptome analyses were conducted on single (protein) or duplicate (transcript) cultures from each condition, which were harvested at the time points denoted with (Fig. 1 *A* and *B*, arrows). Later time points for the iron-limited cultures were chosen to facilitate the harvesting of additional biomass. Several major patterns emerged from these analyses. First, we measured a large difference in the extent of proteome and transcriptome rearrangement under cobalamin starvation that was dependent on the cobalamin requirements of each diatom (Fig. 1 *G* and *H* and Table 1). Second, we identified a previously uncharacterized protein [*P. tricornutum* 48322 and its ortholog in *T. pseudonana* 11697, hereafter referred to as cobalamin acquisition protein 1 (CBA1)] that was more abundant under low cobalamin availability in both diatoms (Fig. 1). Third, we identified changes in gene products involved in folate and methionine cycling, supporting the hypothesis that cobalamin starvation has implications for cellular one-carbon metabolism.

Verification of Nutrient Limitation Scenarios. Cobalamin and iron resupply experiments confirmed that these diatom cultures were starved for nutrients as intended, with iron rescuing growth of both low iron cultures and cobalamin rescuing growth only in the low cobalamin cultures of the cobalamin-requiring diatom, *T. pseudonana* (Fig. 1 *C–F* and *SI Appendix*, Fig. S1). Growth in the combined low cobalamin and low iron *T. pseudonana* culture was only restored upon the addition of both cobalamin and iron together, demonstrating that this culture was simultaneously limited by the availability of both nutrients (Fig. 1*C*). In contrast, growth in low cobalamin and low iron *P. tricornutum* cultures was rescued by iron addition alone and was further enhanced by the coaddition of cobalamin and iron (Fig. 1*E*). These differing responses were expected and are likely attributable to the different requirements for cobalamin in these diatom species.

Cobalamin Scarcity-Induced Proteome and Transcriptome Rearrangement. To detect changes in abundant proteins under environmentally relevant conditions, we used shotgun proteomic methods that allow for a high degree of reproducibility (30) (*SI Appendix*, Fig. S2) as opposed to those that are optimized for deep proteome coverage. In sum, 764 *T. pseudonana* proteins were detected from a total of 4,955 unique peptides with a 0.19% peptide false discovery rate. These proteins and peptides are listed with their putative function in Dataset S1. A total of 859 *P. tricornutum* proteins were detected from 5,172 unique peptides with a 0.22% peptide false discovery rate (Dataset S1). Of the total mass spectra acquired for *T. pseudonana*, 46% were assigned to peptides found in the genomic databases, whereas 52% of *P. tricornutum* spectra were assigned, corresponding to the identification of 6% and 8% of the genome-predicted protein models, respectively.

In the cobalamin-requiring diatom *T. pseudonana,* 19% of detected proteins were significantly differentially abundant under low cobalamin availability compared with the replete control (Fig. 1*C* and Table 1). This suggests that the diatom conducts a significant rearrangement of cellular function when grown under cobalamin limitation. Although some of these changing

Table 1. Pairwise comparisons of growth rate, cell yield, protein abundance changes, and transcript abundance changes between low cobalamin vs. replete growth, low cobalamin with low iron vs. low iron growth, and low iron vs. replete growth in two diatoms

	Low vitamin B_{12} vs. replete	Low vitamin B_{12} and Fe vs. low Fe	Low Fe vs. replete
% proteins differentially abundant			
T. pseudonana	19	18	30
P. tricornutum	5	1	20
% transcripts differentially abundant			
T. pseudonana	26	5	25
P. tricornutum	6	2	16
Fold cell yield decrease			
T. pseudonana	1.8 ± 0.1	1.0 ± 0.1	3.4 ± 0.1
P. tricornutum	1.1 ± 0.1	1.6 ± 0.2	4.1 ± 0.4
Fold growth rate decrease			
T. pseudonana	1.2 ± 0.1	1.2 ± 0.1	2.0 ± 0.1
P. tricornutum	1.0 ± 0.1	1.3 ± 0.1	2.8 ± 0.3

The percentage of proteins changing in abundance was calculated from the total number of identified proteins and those that had significantly different abundance between the two treatments compared (Fisher exact test, $P < 0.01$). The percentage of differentially abundant transcripts was calculated from the number of transcripts mapped to genomic locations that had $\log_2$-fold change RPKM values greater than 1 or less than −1 between the two treatments. Fold cell yield and growth rate decreases were calculated by determining the fold change between the maximum cell density or cell-specific growth rate in each treatment and are given as means of biological triplicates ± 1 SD. Growth rates are cell-specific and were calculated from the following time periods: *T. pseudonana* high iron, days 2–4; *T. pseudonana* low iron, days 3–5; *P. tricornutum* high iron, days 3–6; and *P. tricornutum* low iron, days 5–7.

Table 2. Proteins in higher concentration and significantly differentially abundant ($P < 0.01$) in both low vitamin B_{12} compared with replete and low vitamin B_{12} with low Fe compared with low Fe alone, shown with a putative functional description and average spectral counting scores for each treatment

JGI protein ID no.	Description	Low vitamin B_{12} and Fe	Low Fe	Low vitamin B_{12}	Replete	JGI Protein ID no.	Low vitamin B_{12} and Fe	Low Fe	Low vitamin B_{12}	Replete
T. pseudonana							Homolog in *P. tricornutum*			
270138	Possible glutamine synthetase	4.0	0.0	118.5	76.1	22357	91.3	74.0	73.5	122.5
269942	SHMT2, mitochondrial	29.5	16.1	49.1	20.2	54015	8.3	7.6	19.7	32.4
22483	Unknown, conserved protein	31.8	15.2	25.9	9.1	54686	31.0	22.0	0.7	2.0
11697	**CBA1**	42.4	0.0	45.1	0.0	48322	1.9	0.0	8.5	0.0
24346	Unknown protein	22.5	11.4	25.2	14.2	N/A				
26031	**SHMT1, cytosolic**	19.0	1.9	27.6	10.8	18665	19.2	7.2	21.0	0.0
42612	PLP synthase	18.9	5.0	18.0	3.1	29885	0.3	0.0	2.5	0.5
23556	Unknown protein	12.7	5.8	14.0	7.1	N/A				
23657	Globin-like protein	6.6	2.2	7.2	1.1	46237	0.0	0.0	0.0	0.0
24639	Unknown protein, conserved domains	5.4	1.3	8.0	1.4	42442	1.6	1.6	1.1	0.9
22096	Unknown protein with heme binding domain	3.2	0.0	8.0	2.8	bd1699	0.0	0.0	0.0	0.0
1896	Unknown protein	5.5	1.3	6.0	1.4	N/A				
41733	Thiamine biosynthesis protein ThiC	3.3	0.0	5.2	0.0	38085	0.0	0.0	5.5	0.4
1738	Clp-like protease	2.2	0.0	2.4	0.0	44382	1.6	1.3	0.0	0.0
P. tricornutum							Homolog in *T. pseudonana*			
18665	**SHMT1, cytosolic**	19.2	7.2	21.0	0.0	26031	19.0	1.9	27.6	10.8
28056	MetE, methionine synthase, cobalamin-independent	32.0	0.0	9.6	0.0	N/A				
48322	**CBA1**	1.9	0.0	8.5	0.0	11697	42.4	0.0	45.1	0.0

The average spectral counting scores for the homologous protein in the other diatom are also given. The two proteins highlighted in bold have protein abundances that appear to be driven by vitamin B_{12} availability in both diatoms. JGI, Joint Genome Institute; N/A, absence of a homologous protein encoded in the genome.

proteins are likely responding to the accompanying growth rate depression, there are many that display different behavior under cobalamin vs. iron limitation (Dataset S1) and have putative functions suggesting they are directly related to vitamin B_{12} metabolism (Table 2 and *SI Appendix*, Table S1). Even though iron limitation induced in this study had a much more severe impact on growth rate than low cobalamin did, changes induced in the *T. pseudonana* proteome by cobalamin starvation were nearly as large as those induced by iron limitation (Table 1 and *SI Appendix*, Fig. S3). In contrast, *P. tricornutum,* which can accomplish methionine synthesis without the use of the vitamin, displayed a relatively minor proteome change in response to cobalamin scarcity (Table 1 and *SI Appendix*, Fig. S3). Protein abundance changes under the combined low vitamin B_{12} and low iron treatment vs. treatment with low iron alone showed a similar pattern. *T. pseudonana*, even under severe iron limitation, rearranged its protein complement significantly to manage cobalamin scarcity, whereas *P. tricornutum* changed the abundance of less than 1% of the proteins in its detected proteome (Fig. 1*H* and Table 1). The small change detected between these two treatments in *P. tricornutum* reflects both the minimal metabolic rearrangement induced in these cells and the efficacy of the proteomic analyses applied here.

RNA-seq transcriptomic analyses revealed trends in diatom molecular physiology that were broadly coherent with those observed via proteomics; a similar percentage of the measured transcriptome and proteome changed as a result of each starvation scenario (Table 1). Given the deep coverage of the diatom genomes obtained via these RNA-seq analyses (10,404 genes with mapped transcripts in *P. tricornutum* and 11,778 genes with mapped transcripts in *T. pseudonana*; Dataset S2) and the coherence in the proteome and transcriptome datasets, these data suggest that the cobalamin-requiring diatom *T. pseudonana* conducts a significant rearrangement of its molecular physiology under cobalamin starvation. The diatom with a flexible cobalamin demand, *P. tricornutum*, changed a much smaller proportion of transcript abundances in response to cobalamin scarcity compared with the cobalamin-requiring *T. pseudonana*, also consistent with changes observed in the proteome.

Despite similar global trends in percentage of gene products changing under these conditions, both diatoms showed little coherence in the abundance patterns of specific proteins and transcripts observed in this study (Fig. 2 *A* and *B*). This is consistent with the limited available previous results, suggesting that there is substantial decoupling between transcript and protein abundance patterns in eukaryotic algae (31). Notable exceptions were certain nutrient-specific gene products that showed coordinated increased protein and transcript abundance in response to phosphorus deprivation previously (31, 32) and to cobalamin scarcity in this study, as discussed in subsequent sections (Fig. 2).

Identification of a Previously Undescribed Cobalamin Acquisition Protein. In this shotgun proteomic analysis, the protein that showed the largest response to cobalamin starvation in *T. pseudonana* was a previously uncharacterized hypothetical protein. Here, we identify this protein as a cobalamin acquisition protein, as described below, and refer to it as CBA1 (Figs. 1*G* and 3). A protein homologous to the *T. pseudonana* CBA1 was detected in the *P. tricornutum* global proteome, also only under low cobalamin availability, suggesting that this protein may be playing a similar role in both diatoms and that it is likely involved in cobalamin metabolism (Figs. 1*H* and 3*A*). These shotgun proteomic results were confirmed through absolute protein quantification via a more sensitive and quantitative technique, selected reaction monitoring (SRM) MS, revealing that the concentration of CBA1 protein was between 10- and 160-fold higher under low vitamin B_{12} availability in *P. tricornutum* (Fig. 2*C*). These SRM assays were developed by choosing two tryptic peptides diagnostic of CBA1 and designing specific MS detection assays for

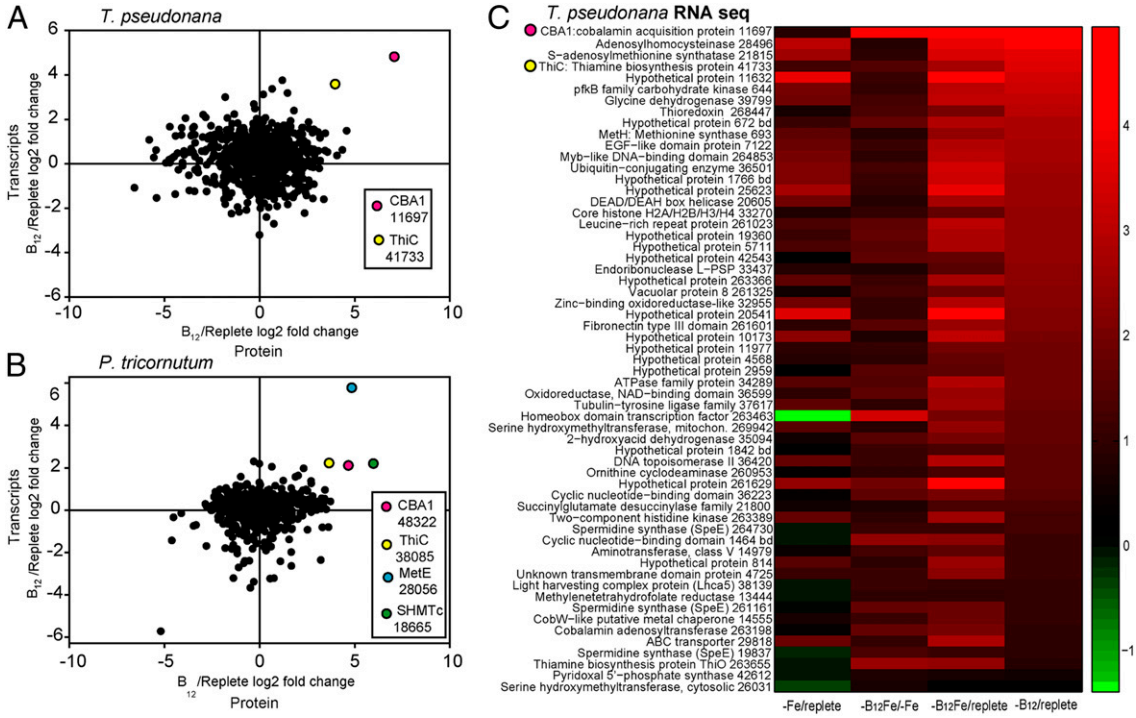

Fig. 2. Transcriptomic analyses reveal additional patterns in cobalamin-responsive gene products. (*A* and *B*) Comparative proteome and transcriptome responses to cobalamin deprivation are shown. All gene products for which there was both protein- and transcript-based quantitative information are displayed. The fold change (log_2) between the transcript abundance (RPKM value) in the cobalamin-starved and replete treatments is shown on the *y* axis, and the fold change (log_2) between the protein abundance (spectral counting score) in the cobalamin-starved and replete treatments is shown on the *x* axis. For the protein data, any null values were replaced with a spectral counting score of 0.33, the lowest measurable value in our experiments, to facilitate the computation. Generally, coherence between the proteome and transcriptome responses is limited to specific proteins that display enhanced abundance under cobalamin starvation in both the transcript and protein pools. These include CBA1, MetE, ThiC, and cytosolic SHMT, as noted by color and identified in the key. (*C*) Heat map displays select *T. pseudonana* transcript responses to cobalamin and iron starvation. Fold change RPKM values are shown for the low iron vs. replete, low vitamin B_{12} with low Fe vs. low Fe, low vitamin B_{12} with low Fe vs. replete, and low vitamin B_{12} vs. replete treatments, with up-regulation denoted in red and down-regulation denoted in green. The genes were selected by high-to-low ordering of the log_2-transformed fold change RPKM values and sorted by the comparison between low vitamin B_{12} vs. replete treatments. Gene products highlighted in *A* (ThiC, CBA1) are also highlighted in *C*.

each, as previously described (30) (*SI Appendix*, Table S2). This method involves the use of stable, isotopically labeled versions of two diagnostic tryptic peptides (Pt48322_1 and Pt48322_2) that were used as internal standards. Each behaved linearly over four orders of magnitude (*SI Appendix*, Fig. S4) and allowed for absolute quantification of each of these peptides, both generated from the CBA1 amino acid sequence, in *P. tricornutum* peptide samples. However, one peptide, Pt48322_2, was measured at consistently higher abundance in *P. tricornutum* than the other diagnostic peptide, Pt48322_1 (Fig. 3*C*). Careful examination

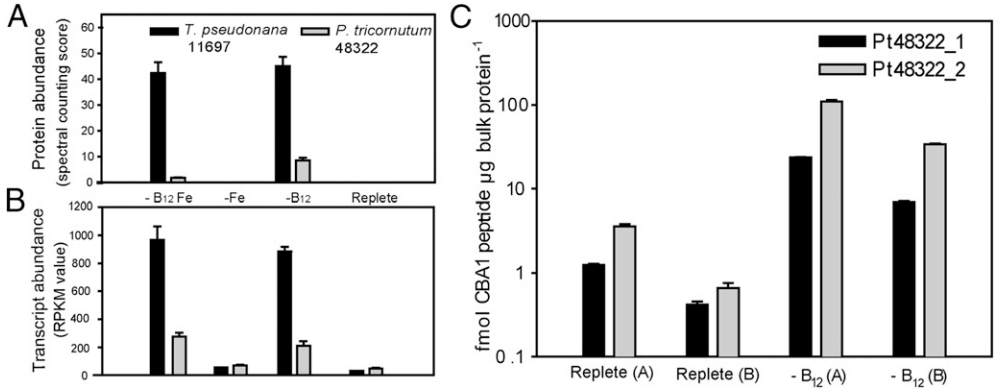

Fig. 3. CBA1 is much more abundant under low vitamin B_{12} availability via three independent quantitative analyses. (*A*) Bars are means of spectral counting abundance scores for protein CBA1 in four treatments in both diatoms as measured via shotgun ion trap MS. Error bars represent 1 SD about the mean of technical triplicate measurements. (*B*) Bars are means of transcript RPKM abundance scores for CBA1 sequences in four treatments in both diatoms in RNA-seq transcriptomic analyses. Error bars represent 1 SD about the mean of biological duplicate measurements. (*C*) Absolute abundance of two peptides from CBA1 in *P. tricornutum* measured via the highly sensitive and quantitative technique SRM MS in two low vitamin B_{12} and two replete cultures. Error bars are 1 SD about the mean of technical triplicate measurements.

of nucleic acid sequences amplified from cobalamin-limited *P. tricornutum* RNA extracts revealed that this variability was attributable to SNPs within allelic copies of the CBA1 coding sequences of this diploid diatom genome. Diatoms possess two copies of each chromosome, and sequence analysis revealed that there are minutely different versions of the gene encoding CBA1 on each of these copies. These slightly different genes produce CBA1 protein with sequences that differ by three amino acid residues. The diagnostic tryptic peptide target Pt48322_1 contains one of these variable amino acid residues, and hence is encoded by one of the two chromosomal copies, whereas the other peptide target, Pt48322_2, does not contain a variable site and is encoded by both copies (details are provided in *SI Appendix*, Fig. S5). The abundance of these peptides was linearly correlated across all samples (r^2 = 0.999; *SI Appendix*, Fig. S6), and Pt48322_2, the peptide encoded by both allelic copies, was more abundant (Fig. 2*C*). This is therefore an example of canonical gene expression in a diploid genome, where allelic copies display similar expression patterns.

RNA-seq analysis revealed that CBA1 transcript abundance patterns were similar to those observed for the corresponding proteins via the two proteomic approaches described above. Much higher CBA1 transcript abundance (4- to 30-fold) was observed under low cobalamin availability in both *T. pseudonana* and *P. tricornutum* (Fig. 3*B*). Together, these analyses reveal that CBA1 protein and transcripts display coordinated behavior under cobalamin scarcity (Fig. 2).

CBA1 has a clear N-terminal signal peptide sequence for secretion [Cello and SignalP- predicted (33, 34)] and no transmembrane domains. It contains a partial conserved domain that is weakly similar to the periplasmic component of a bacterial iron hydroxamate ATP-binding cassette (ABC) transport system (FepB; N-terminal end is truncated, Pt48322 BLASTp search E-value 1.33e-4), but the protein otherwise lacks characterized domains. There appear to be homologous versions of CBA1 encoded in all currently sequenced diatom genomes as well as in those from other members of the stramenopile lineage, *Ectocarpus siliculosus* and *Aureococcus anophagefferens*. Outside of the stramenopile lineage, we find no sequences that clearly encode CBA1 within available algal genome (*SI Appendix*, Table S3) or transcriptome datasets, including extensive dinoflagellate transcriptome sequencing. However, because these sequences are not highly conserved and show some similarities to a class of bacterial proteins, rigorous analysis of the functional similarity of these proteins between algal lineages will depend on the future identification of residues essential for CBA1's functionality.

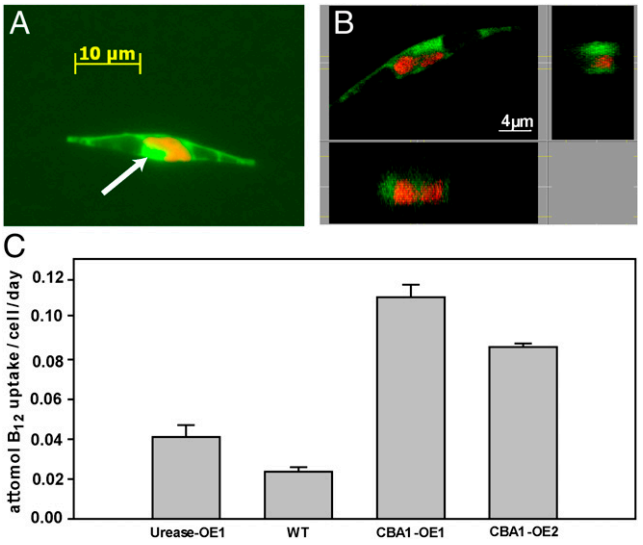

Fig. 4. Protein CBA1 is directly implicated in cobalamin acquisition. Protein CBA1 appears to be localized to the outside of the cell and likely the ER, and is directly implicated in cobalamin acquisition. Epifluorescent (*A*) and confocal (*B*) micrographs of protein CBA1 fused to YFP and overexpressed in *P. tricornutum*. YFP fluorescence is false-colored green, whereas chlorophyll a fluorescence is false-colored red. The side panels of the confocal image show the fluorescence distribution in the cross-sections of the central image indicated by the light yellow lines. (*C*) Cobalamin uptake rates by wt *P. tricornutum* and transgenic *P. tricornutum* cell lines overexpressing CBA1 (CBA1-OE1, CBA1-OE2) or Urease (Urease-OE1) measured over 24 h in exponential growth phase under vitamin B_{12}-replete conditions (n = 3). The growth rate over the 24-h experiment for the wt (WT) was 0.72 ± 0.07; for Urease-OE1, it was 1.01 ± 0.02; for CBA1-OE2, it was 1.10 ± 0.03; and for CBA1-OE1, it was 1.08 ± 0.03, given as mean of measurements on biological triplicate cultures ± 1 SD.

Overexpression of CBA1, Subcellular Localization, and Phenotypic Characterization. We examined the subcellular localization of CBA1 through overexpression of the *P. tricornutum* isoform (Pt48322) in the native host as an YFP fusion construct. Epifluorescent microscopy showed that the YFP signal was associated with the outer membrane and also with an intracellular organelle-like structure adjacent to the chloroplast (Fig. 4*A*, white arrow). YFP fluorescence was not observed throughout the cytoplasm or within the mitochondria, chloroplast, peroxisome, or nucleus. Intracellular localization around, but not within, the chloroplast was verified using confocal microscopy (Fig. 4*B*) and is similar to fluorescence microscopy of other proteins localized to the diatom endoplasmic reticulum (ER), which envelopes the chloroplast in red lineage algae (35). The most well-characterized eukaryotic protein secretion pathway begins with translocation of proteins with signal peptides across the ER membrane into the lumen, where they are retained for processing and folding and are subject to quality control. After ER processing, these proteins are exported to the Golgi network and then relocated to the outside of the cell via vesicle transport (36). The likely ER lumen localization detected here for CBA1 may thus be attributable to processing of this protein before secretion. This is consistent with the strong prediction for a signal peptide (D = 0.86, SignalP 3.0) and observed additional targeting to the periphery of the cell.

We characterized the phenotypic response of this overexpression in *P. tricornutum* by measuring cobalamin uptake rates in two cell lines overexpressing this protein (CBA1-OE1 and CBA1-OE2) and comparing them with uptake rates in the wild type (wt) and a line overexpressing an unrelated protein, Urease (Urease-OE1). To repress native CBA1 expression, uptake rates were measured in cultures grown in cobalamin-replete conditions. In the transgenic diatoms, the promoter for a light harvesting complex protein (FcpB), which is highly expressed under exponential growth, controls CBA1 overexpression. Overexpression of CBA1 enhanced cell-specific, radiolabeled cobalamin uptake rates in exponentially growing *P. tricornutum* cells two- to threefold (Fig. 4*C*). This enhanced uptake rate directly implicates CBA1 in cobalamin acquisition and, along with its localization to the outside of the cell, suggests that CBA1 may bind cobalamin and, through association with additional unidentified proteins, participate in its transport into the cell. This finding is significant in that CBA1 is, to our knowledge, the sole identified protein in any marine eukaryotic microbe to be directly linked to cobalamin acquisition.

CBA1 in Environmental Sequencing Datasets. We identified transcripts that likely encode CBA1 in cDNA libraries generated from natural phytoplankton communities (Fig. 5 and *SI Appendix*, Fig. S7). CBA1 transcripts were detected in each library examined. These libraries were generated from diverse marine locations, including sea ice and water column samples from the Ross Sea of the Southern Ocean and water column samples from the North Pacific Ocean, Monterey Bay, and Puget Sound, suggesting that

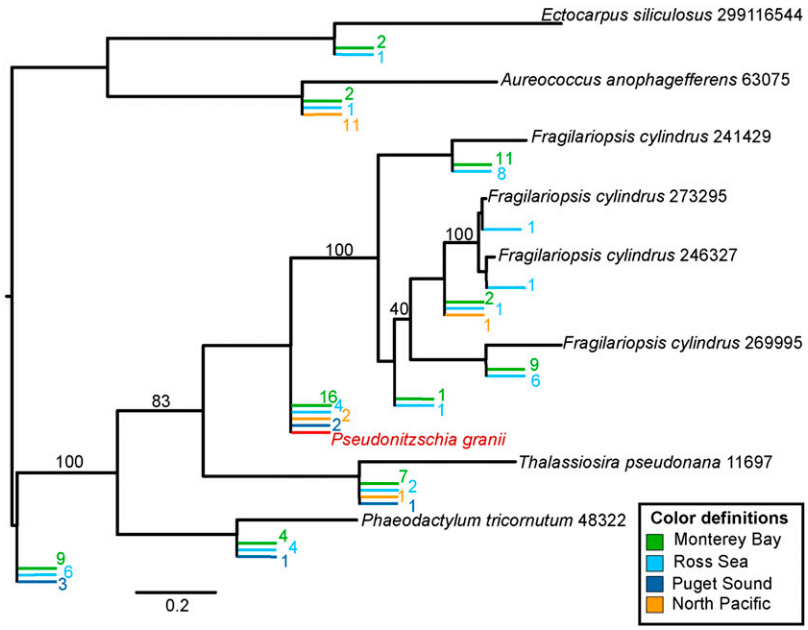

Fig. 5. Transcripts encoding CBA1 are expressed in diverse marine environments. Phylogenetic tree containing CBA1 sequences from 454 metatranscriptomic (cDNA) libraries from the Ross Sea of the Southern Ocean, Monterey Bay, Puget Sound, and North Pacific Ocean. Reference sequences from *P. tricornutum*, *F. cylindrus*, *T. pseudonana*, *A. anophagefferens*, and *E. siliculosus* genomes were used to construct these trees (37) and are shown in black. CBA1-like sequences from environmental samples are shown in color, as described in the key. CBA1 transcripts were detectable in diverse marine environments, suggesting that cobalamin acquisition is an important component of diatom molecular physiology.

CBA1 is widely distributed within the marine environment. These nucleic acid sequences are displayed as a phylogenetic tree that uses CBA1 sequences from available genomes of cultured organisms to construct a reference tree onto which metatranscriptomic sequences are placed (37) (Fig. 5). Although CBA1 is not among the most abundant diatom transcripts detected, its expression level is comparable to that of transcripts encoding other essential well-characterized proteins (*SI Appendix*, Fig. S7). This suggests that CBA1 and the micronutrient cobalamin play important roles in the molecular physiology and ecology of natural diatom communities. In metagenomic data, DNA sequences most similar to CBA1 were detected only in the larger size fraction (>3 μm) samples, suggesting that this protein is restricted to larger phytoplankton. The majority of environmentally derived transcripts encoding CBA1 are most similar to those encoded by cultured diatoms (Fig. 5), suggesting that diatoms are likely the major reservoir of this protein in the marine environment.

Other Molecular Responses to Low Cobalamin Availability. Identification of CBA1 and its abundance patterns in culture suggests that diatoms adjust their molecular physiology to increase capacity for cobalamin acquisition in the face of cobalamin deprivation. We can consider other cobalamin-sensitive transcripts and proteins to identify additional molecular responses to low vitamin availability.

Included in the small pool of *P. tricornutum* gene products changing under cobalamin starvation is the cobalamin-independent methionine synthase MetE (28056), which was much more abundant under low cobalamin availability (Figs. 1*D* and 2 and Table 2). This suggests that *P. tricornutum* expresses MetE to replace MetH when cobalamin is scarce, consistent with transcript abundance patterns observed previously in this diatom (21). RNA-seq results revealed that an adjacent two-component histidine kinase sensor appears to be coregulated with *metE*, and thus may play a role in the *P. tricornutum* response to cobalamin starvation (Fig. 6). The cobalamin-dependent methionine synthase MetH was not detected in our proteome study, possibly attributable to low abundance, which is expected because MetH has much higher catalytical activity compared with MetE (24). MetH (Pt 23399, Tp 693) transcripts were detected here via RNA-seq and did not show significant changes in abundance as a function of cobalamin availability in *P. tricornutum* but were more abundant under low cobalamin in *T. pseudonana* (Fig. 2*C* and Dataset S2).

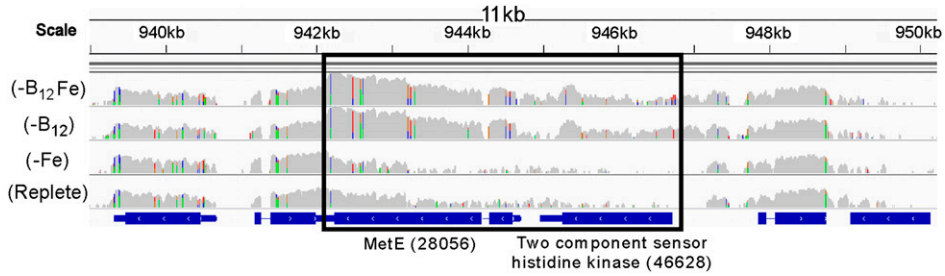

Fig. 6. RNA-seq analysis reveals coregulation of MetE and a two-component sensor. RNA-seq coverage for an 11-kb region of the *P. tricornutum* genome is shown. Individual tracks are shown for each treatment, cobalamin and iron starvation, cobalamin starvation, iron starvation, and the replete control. The *x* axis shows the position in the genome, and the *y* axis (gray shading) shows the relative coverage of transcript data. Vertical color lines represent areas in the coverage mapping where there were mismatches of the reads to the reference genome (A = green, C = blue, G = yellow, T = red). The bottom track shows the gene models from the Joint Genome Institute 2.0 genome project. Transcripts mapping to cobalamin-independent methionine synthase (*metE*) are much more abundant under cobalamin scarcity and with low cobalamin and low iron. In addition, a two-component histidine kinase sensor appears to be coregulated with *metE*, and may thus play a role in the *P. tricornutum* response to cobalamin starvation.

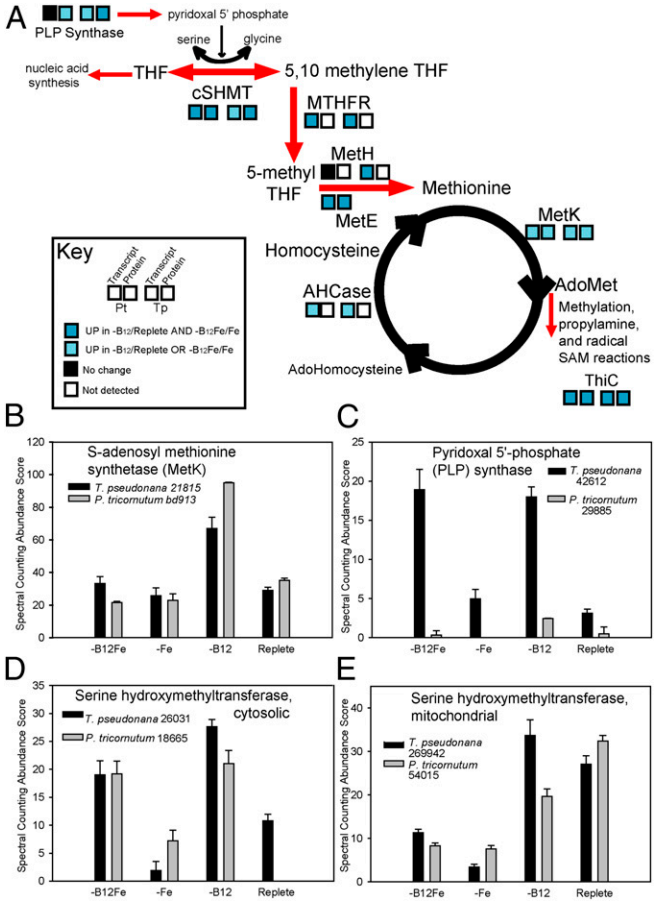

Fig. 7. Interconnections between methionine, folate, PLP metabolism, and cobalamin availability. (*A*) Schematic diagram describes the connections between PLP, folate (THF), methionine, and thiamine metabolism in two diatom species, displayed with supporting protein abundance data. The gene products involved in these pathways and their responses to cobalamin scarcity are shown for each diatom, as denoted in the key. The behavior of both transcripts and proteins are shown, with Pt indicating *P. tricornutum* (*Left*) and Tp indicating *T. pseudonana* (*Right*). Dark blue indicates that the gene product is more abundant under −vitamin B_{12} vs. replete conditions and −vitamin B_{12}Fe vs. −Fe conditions, and lighter blue indicates that the gene product was more abundant under one of those conditions. Black denotes that there was no change observed between these conditions, and white indicates that the product was not detected. AHCase, adenosylhomocysteinase; cSHMT, cytosolic serine hydroxymethyltransferase; MTHFR, methylenetetrahydrofolate reductase. (*B*–*E*) Abundance patterns for select proteins included in the schematic above are displayed. Bar graphs of spectral counting abundance scores for proteins of interest are given for each of four treatments in both diatoms, where bars are means of technical triplicate measurements and error bars are 1 SD about the mean. Overall, these patterns suggest that there are interconnections between methionine, folate, PLP, and thiamine metabolism and cobalamin availability in diatoms.

Another potential use for cobalamin in diatom cells is as a cofactor for methylmalonyl CoA mutase (MmcM; Pt 51830, Tp 33685). This enzyme's function remains unclear in diatoms, although it may be important in the citric acid cycle as well as in propionate and fatty acid metabolism (20). MmcM uses adenosylcobalamin as a cofactor, which could be produced via an adenosylcobalamin transferase enzyme encoded in these diatom genomes (CblB; Pt 45992, Tp 263198). MmcM encoding transcripts did not respond to low cobalamin, but those encoding CblB were more abundant under cobalamin scarcity (Fig. 3*C* and Dataset S2). Because this protein may be responsible for generating adenosylcobalamin, a biologically active form of the vitamin that is not used by MetH, the reason for this response remains unclear and leaves open the possibility that there are additional, unrecognized cobalamin-dependent metabolisms in diatoms.

Another potentially important consequence of low cobalamin availability to diatom cells is methyl folate trapping (14, 38). Under conditions of reduced methionine synthase activity, 5-methyltetrahydrofolate (MeTHF) buildup can prevent efficient folate cycling, and thus have an impact on other folate-dependent metabolisms, such as DNA synthesis (28). We find molecular evidence for this phenomenon in these diatoms, as summarized in Fig. 7. A protein involved in folate one-carbon metabolism, cytosolic serine hydroxymethyltransferase (SHMT) (36, 37), is more abundant under low cobalamin (Figs. 2 and 7 and Dataset S2). SHMT is pyridoxal 5′ phosphate (PLP, vitamin B_6)-dependent and catalyzes the interconversion of THF and 5,10-methylene tetrahydrofolate [5,10 MTHF (39)]. The 5,10 MTHF can then be converted irreversibly to MeTHF, which accumulates under low methionine synthase activity and leads to folate trapping in other organisms (28). The increase in cytosolic SHMT abundance under cobalamin scarcity suggests that diatoms may increase their capacity for THF and 5,10 MTHF interconversion under low cobalamin conditions, perhaps in an effort to mediate folate cycling. This is consistent with assertions that SHMT mediates the partitioning of one-carbon units between DNA synthesis and methionine cycling in other organisms (40). Additionally, an enzyme putatively involved in PLP synthesis is much more abundant under low cobalamin in *T. pseudonana* (Figs. 2 and 7). This increase is consistent with a higher demand for PLP under low cobalamin, potentially for use by the PLP-dependent SHMT enzymes. Taken together, these data suggest that folate, cobalamin, and PLP metabolism are linked in diatoms, as observed previously in metazoans (41).

These results also suggest that AdoMet starvation may be an important consequence of cobalamin deprivation in diatoms (Fig. 7). Methionine adenosyltransferase (MetK), which converts methionine to AdoMet, was more abundant under low cobalamin in both diatoms (Fig. 7 and *SI Appendix*, Table S1). In addition to many other cellular functions, AdoMet is responsible for reductive repair and remethylation of cobalamin in MetH (42). Increased MetK levels may enhance AdoMet production, leading to more efficient repair of oxidized cobalamin in MetH. It is also possible that enhanced MetK abundance results from attempts to meet cellular AdoMet demand despite methionine scarcity. ThiC, a thiamine biosynthesis protein, which uses an AdoMet-dependent radical reaction to form the nonsulfur-containing branch of thiamine, 4-amino-5-hydroxymethyl-2-methylpyrimidine (43), was more abundant under low cobalamin (Figs. 2 and 7). Because the other proteins involved in thiamine biosynthesis were not observed to be more abundant under cobalamin scarcity, it may be that ThiC is up-regulated in response to AdoMet deprivation (Fig. 7), supporting the notion that AdoMet starvation is an important consequence of low cobalamin availability.

Although some proteins that display cobalamin-responsive abundance patterns have predicted cellular functions as described above, more than half of them play unknown roles (Table 2 and *SI Appendix*, Table S2). Three proteins of unknown function in *T. pseudonana* (24346, 23556, and 1896; Table 2) do not have homologs in *P. tricornutum*, were more abundant under low cobalamin availability, and did not increase in abundance under low iron. These unknown proteins may be involved in the *T. pseudonana* response to cobalamin scarcity and warrant further evaluation of their potential role in cobalamin acquisition.

Implications for Cobalamin Biochemistry and Marine Biogeochemistry. Our results suggest that diatoms use at least three strategies in response to cobalamin scarcity, including efforts to increase cobalamin acquisition machinery, reduce cobalamin demand, and mitigate damage induced by reduced methionine synthase activity

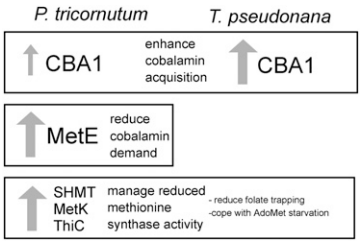

Fig. 8. Diatoms display three primary responses to cobalamin scarcity. Schematic representation of the three primary responses to cobalamin starvation in two diatoms. Both diatoms enhanced CBA1 production, likely in an effort to enhance cobalamin acquisition. The magnitude of the increase in CBA1 protein and transcripts was larger for *T. pseudonana*, likely because it has an absolute cobalamin requirement. *P. tricornutum* enhanced MetE production to reduce cobalamin demand. Both diatoms also appeared to conduct cellular rearrangements to cope with reduced methionine synthase activity, including enhanced cytosolic SHMT, MetK, and radical AdoMet (SAM) enzyme ThiC abundance under low cobalamin availability.

(Fig. 8). These results implicate enhanced CBA1, MetE, and SHMT abundance as well as altered folate and PLP metabolism in the acclimation of diatom cells to low cobalamin availability and suggest that AdoMet starvation is an important consequence of cobalamin deprivation in diatoms. Future metabolomic mapping of the folate, PLP, homocysteine, and methionine-derived compounds involved in these processes could yield important additional insight into these dependencies.

Because CBA1 is a component of a previously undescribed cobalamin acquisition pathway that differs from the well-characterized bacterial and metazoan mechanisms, its identification may advance efforts to understand the biochemistry and evolution of cobalamin acquisition. Further analyses are required to elucidate the specific role served by CBA1 in cobalamin acquisition and to identify the other proteins involved in this process. Subsequent comparative studies of different cobalamin acquisition pathways may then yield valuable insight into cobalamin biochemistry and utilization through time.

Previous work revealed that diatoms and other algae exude a protein into their growth media that binds vitamin B_{12} very strongly (44, 45). In *T. pseudonana*, this protein was shown to be a distinct, large, multimeric complex (46), but the identity of its components remains unknown. It is possible that CBA1 represents a part of this multimer, although this is unlikely because overexpression of CBA1 without any other components of a complex enhances cobalamin uptake, whereas the exuded vitamin B_{12} binder is thought to restrict vitamin uptake (44, 45). It is also possible that CBA1 interacts with this complex to move cobalamin from the growth media into the cell (47). Future study will be required to elucidate the relationship between CBA1 and the strong exuded vitamin B_{12} binders produced by algae.

Detection of CBA1 transcripts in existing marine environmental datasets (Fig. 5 and *SI Appendix*, Fig. S7) implies that this protein is abundant and used by natural phytoplankton populations, and that cobalamin acquisition is thus an important component of diatom molecular physiology in the natural environment. This, along with the large-scale proteome and transcriptome rearrangement induced by cobalamin scarcity in diatoms, supports including cobalamin availability and its production and consumption dynamics in efforts to predict and model marine primary productivity. Future studies mapping the abundance of CBA1 and other cobalamin-responsive gene products identified here, along with traditional measures of productivity and community composition, may prove useful for elucidating the major oceanic controls on cobalamin starvation and utilization in phytoplankton communities and could yield insights into the influence of cobalamin availability on marine primary productivity and microbial community composition.

Methods

Axenic cultures of *T. pseudonana* CCMP 1335 *and P. tricornutum* CCMP 632 were grown in f/2-type media with a natural seawater base, with modified vitamin, EDTA, and trace metal concentrations (48). For the *P. tricornutum* experiment, cells were acclimated with 5 nM added total iron (10^{-11} M Fe′) and 0.5 pM added vitamin B_{12} for four transfers, allowing at least three doublings per transfer. Cells were then inoculated (3.2% vol/vol) into twelve 2.2-L bottles containing 1.8 L of media (4 treatments, biological triplicates). The media was as described above except for variable iron and vitamin B_{12} concentrations, which were chosen based on previous studies to result in cultures experiencing replete, iron-limited, low cobalamin and low cobalamin with iron limitation conditions (49, 50). The low iron treatment had 2.5 nM Fe_{total} ($10^{-11.3}$ M Fe′) and 100 pM added vitamin B_{12}, the low vitamin B_{12} treatment had no added vitamin B_{12} and 100 nM Fe_{total} ($10^{-9.69}$ M Fe′), and the low vitamin B_{12} and low iron treatment had no added vitamin B_{12} and 2.5 nM Fe_{total} ($10^{-11.3}$ mol L Fe′), whereas the replete treatment had 100 pM added vitamin B_{12} and 100 nM Fe_{total} ($10^{-9.69}$ M Fe′). Samples were taken daily for fluorescence and cell counts, and were harvested for protein and RNA after 6 d for the high iron conditions and after 9 d for the low iron condition (Fig. 1 *A* and *B*, arrows). Just after the harvest time point, each remaining culture was split in four and resupplied with either nothing, 100 pM vitamin B_{12}, 100 nM Fe_{total}, or both vitamin B_{12} and Fe, and growth was monitored via in vivo fluorescence.

The *T. pseudonana* experiment was conducted as above except with different vitamin B_{12} and iron concentrations chosen due to known differences in cobalamin and iron requirements for these diatoms (49–51). Acclimation cultures had 1 pM added vitamin B_{12} and 65 nM added Fe_{total} ($10^{-9.88}$ M Fe′). The low iron treatment had 50 nM Fe_{total} ($10^{-10.0}$ M Fe′) and 100 pM added vitamin B_{12}, the low vitamin B_{12} treatment had 0.3 pM added vitamin B_{12} and 400 nM Fe_{total} ($10^{-9.09}$ M Fe′), and the low vitamin B_{12} and low iron treatment had 0.3 pM added vitamin B_{12} and 50 nM Fe_{total} ($10^{-10.0}$ M Fe′), and the replete treatment had 100 pM added vitamin B_{12} and 400 nM Fe_{total} ($10^{-9.09}$ M Fe′). Just after the harvest time point, each culture was split in four and resupplied with either nothing, 100 pM vitamin B_{12}, 400 nM Fe_{total}, or both vitamin B_{12} and Fe, and growth was monitored via in vivo fluorescence.

Protein was extracted, digested, and analyzed via shotgun MS as previously described, with minor modifications (30, 31, 52). Briefly, protein digests from one of the three replicates were analyzed in technical triplicate via liquid chromatography MS using a Paradigm MS4 HPLC system (Michrom) with reverse phase chromatography, a Michrom ADVANCE source, and a Thermo Scientific LTQ ion trap mass spectrometer. The mass spectra were searched using SEQUEST (Thermo, Inc.). Database search results were further processed using the PeptideProphet statistical model (53) and spectral counting abundance scoring within Scaffold 3.0 (Proteome Software, Inc.). Proteins discussed as "differentially expressed" in pairwise comparisons were determined by the Fisher exact test ($P < 0.01$) (54).

Targeted MS was conducted via SRM as previously described (30) for two tryptic peptides found to be unique to CBA1 in *P. tricornutum*, using a Thermo Vantage TSQ Triple Quadrupole Mass Spectrometer (Thermo Scientific). Isotopically labeled versions of each tryptic peptide (Sigma–Aldrich) (55) were used as internal standards. Method details are given in *SI Appendix*, Table S2.

RNA from duplicate samples from each treatment was purified, amplified, and then used to prepare SOLiD Total RNA-Seq Kit (Life Technologies) libraries according to the instructions of the manufacturer. The raw SOLiD sequence data were mapped and aligned against the reference genome, and reads per kilobase of exon model (RPKM) values for each sample were assigned. These RPKM values were pooled for replicates, and differentially abundant transcripts were identified as those that had a $\log_2$-fold change in RPKM values greater than 1 or less than −1 between the different treatments.

For CBA1 overexpression, full-length *P. tricornutum* 48322 cDNA was PCR-amplified and cloned into a TOPO pENTR (Invitrogen) and then subjected to Gateway (Invitrogen) recombination with a diatom C-terminal YFP pDONR vector (56), which was transformed into *P. tricornutum* via particle bombardment (57). For localization, epifluorescent microscopy was performed on a Zeiss Axioscope. Confocal microscopy was performed on a Leica TCS SP5 spectral system.

Vitamin B_{12} uptake rate assessments were conducted on wt and overexpression lines similar to previously described methods, using ^{57}Co-labeled cobalamin, monitored via gamma counting (58).

Metatranscriptomic analyses used to construct the tree in Fig. 5 were conducted on multiple samples collected from various marine locations. RNA was purified and amplified linearly, and cDNA was synthesized and prepared

for sequencing on the 454 platform (Roche Diagnostics) according to manufacturer protocols. Orthologs to CBA1 were retrieved from the cDNA sequence data by TBLASTn. For comparative abundance analyses (*SI Appendix*, Fig. S7), cDNA was prepared from six Ross Sea samples for sequencing on the Illumina platform. The resulting reads were assembled de novo and compared with available sequence databases for functional and phylogenetic annotation. Diatom ORFs assigned to each domain of interest were identified, and read counts for each ORF were summed across all six libraries. RPKM values were calculated for each ORF and were then summed across ORFs that contained a domain of interest, effectively normalizing transcript abundance to coding sequence length and abundance of all diatom-assigned transcripts.

ACKNOWLEDGMENTS. We thank Dawn Moran, Abigail Heithoff, Louie Wurch, Matt McIlvin, and Vladimir Bulygin for technical assistance; Abigail Noble, Jeff Hoffman, and Jeff McQuaid for Antarctic sample collection; and Mauricio Arriagada for contributions to metatranscriptomic analyses. We acknowledge Cathy Drennan, Sonya Dyhrman, Dianne Newman, and Ben Van Mooy for helpful discussions. We are grateful for comments provided by anonymous reviewers. This work was supported by National Science Foundation Awards ANT 0732665, OCE 0752291, and OCE 1031271 and Gordon and Betty Moore Foundation funding (to M.A.S.); by National Science Foundation Graduate Research Fellowship 2007037200 and Environmental Protection Agency STAR Fellowship F6E20324 (to E.M.B.); and by National Science Foundation Awards ANT 0732822, ANT 1043671, MCB 1024913, OCE 0727997, and OCE 1136477 and Department of Energy Award DE SC0006719 (to A.E.A.).

1. Falkowski PG, et al. (2004) The evolution of modern eukaryotic phytoplankton. *Science* 305:354–360.
2. Nelson DM, Treguer P, Brzezinski MA, Leynaert A, Queguiner B (1995) Production and dissolution of biogenic silica in the oceans: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochem Cycles* 9:359–372.
3. Boyd PW, et al. (2007) Mesoscale iron enrichment experiments 1993-2005: Synthesis and future directions. *Science* 315:612–617.
4. Moore JK, Doney SC, Lindsay K (2004) Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. *Global Biogeochem Cycles* 18(GB4028): 10.1029/2004GB002220.
5. Panzeca C, et al. (2006) B vitamins as regulators of phytoplankton dynamics. *EOS* 87: 593–596.
6. Bertrand EM, et al. (2007) Vitamin B_{12} and iron co-limitation of phytoplankton growth in the Ross Sea. *Limnol Oceanogr* 52:1079–1096.
7. Gobler CJ, Norman C, Panzeca C, Taylor GT, Sanudo-Wilhelmy SA (2007) Effect of B-vitamins and inorganic nutrients on algal bloom dynamics in a coastal ecosystem. *Aquat Microb Ecol* 49:181–194.
8. Koch F, et al. (2011) The effect of vitamin B_{12} on phytoplankton growth and community structure in the Gulf of Alaska. *Limnol Oceanogr* 56:1023–1034.
9. Cowey CB (1956) A preliminary investigation of the variation of vitamin B-12 in oceanic and coastal waters. *J Mar Biol Assoc U.K.* 35:609–620.
10. Droop MR (1957) Vitamin B_{12} in marine ecology. *Nature* 180:1041–1042.
11. Menzel DW, Spaeth JP (1962) Occurrence of vitamin B_{12} in the Sargasso Sea. *Limnol Oceanogr* 7:151–154.
12. Hutchinson GE (1961) The paradox of the plankton. *Am Nat* 95:137–145.
13. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS (2003) Comparative genomics of the vitamin B_{12} metabolism and regulation in prokaryotes. *J Biol Chem* 278: 41148–41159.
14. Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG (2005) Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature* 438:90–93.
15. Droop MR (2007) Vitamins, phytoplankton and bacteria: Symbiosis or scavenging? *J Plankton Res* 29:107–113.
16. Karl DM (2002) Nutrient dynamics in the deep blue sea. *Trends Microbiol* 10:410–418.
17. Azam F, Malfatti F (2007) Microbial structuring of marine ecosystems. *Nat Rev Microbiol* 5:782–791.
18. Bertrand EM, Saito MA, Jeon YJ, Neilan BA (2011) Vitamin B_{12} biosynthesis gene diversity in the Ross Sea: The identification of a new group of putative polar B_{12} biosynthesizers. *Environ Microbiol* 13:1285–1298.
19. Panzeca C, et al. (2009) Distributions of dissolved vitamin B12 and Co in coastal and open-ocean environments. *Estuar Coast Shelf Sci* 85:223–230.
20. Croft MT, Warren MJ, Smith AG (2006) Algae need their vitamins. *Eukaryot Cell* 5: 1175–1183.
21. Helliwell KE, Wheeler GL, Leptos KC, Goldstein RE, Smith AG (2011) Insights into the evolution of vitamin B_{12} auxotrophy from sequenced algal genomes. *Mol Biol Evol* 28: 2921–2933.
22. Banerjee RV, Matthews RG (1990) Cobalamin-dependent methionine synthase. *FASEB J* 4:1450–1459.
23. Goulding CW, Postigo D, Matthews RG (1997) Cobalamin-dependent methionine synthase is a modular protein with distinct regions for binding homocysteine, methyltetrahydrofolate, cobalamin, and adenosylmethionine. *Biochemistry* 36: 8082–8091.
24. González JC, Banerjee RV, Huang S, Sumner JS, Matthews RG (1992) Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from Escherichia coli: Two solutions to the same chemical problem. *Biochemistry* 31:6045–6056.
25. Stefels JP (2000) Physiological aspects of the production and conversion of DMSP in marine algae and higher plants. *J Sea Res* 43:183–197.
26. Sunda W, Kieber DJ, Kiene RP, Huntsman S (2002) An antioxidant function for DMSP and DMS in marine algae. *Nature* 418:317–320.
27. Charlson RJ, Lovelock JE, Andreae MO, Warren SG (1987) Oceanic phytoplankton, atmospheric sulfur, cloud albedo and climate. *Nature* 326:655–661.
28. Scott JM, Weir DG (1981) The methyl folate trap. A physiological response in man to prevent methyl group deficiency in kwashiorkor (methionine deficiency) and an explanation for folic-acid induced exacerbation of subacute combined degeneration in pernicious anaemia. *Lancet* 2:337–340.
29. Marchetti A, et al. (2012) Comparative metatranscriptomics identifies molecular basis for physiological responses of phytoplankton to varying iron availability. *Proc Natl Acad Sci USA* 109:1828–1829.
30. Saito MA, et al. (2011) Iron conservation by reduction of metalloenzyme inventories in the marine diazotroph Crocosphaera watsonii. *Proc Natl Acad Sci USA* 108:2184–2189.
31. Wurch LL, Bertrand EM, Saito MA, Van Mooy BA, Dyhrman ST (2011) Proteome changes driven by phosphorus deficiency and recovery in the brown tide-forming alga *Aureococcus anophagefferens*. *PLoS ONE* 6:e28949.
32. Dyhrman ST, et al. (2012) The Transcriptome and Proteome of the Diatom Thalassiosira pseudonana Reveal a Diverse Phosphorus Stress Response. *PLoS ONE* 7: e33768.
33. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* 10: 1–6.
34. Yu CS, Chen YC, Lu CH, Hwang JK (2006) Prediction of protein subcellular localization. *Proteins* 64:643–651.
35. Apt KE, et al. (2002) In vivo characterization of diatom multipartite plastid targeting signals. *J Cell Sci* 115:4061–4069.
36. Ellgaard L, Molinari M, Helenius A (1999) Setting the standards: Quality control in the secretory pathway. *Science* 286:1882–1888.
37. Matsen FA, Kodner RB, Armbrust EV (2010) pplacer: Linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 11:538.
38. Selhub J (2002) Folate, vitamin B_{12} and vitamin B_6 and one carbon metabolism. *J Nutr Health Aging* 6:39–42.
39. Snell K, et al. (2000) The genetic organization and protein crystallographic structure of human serine hydroxymethyltransferase. *Adv Enzyme Regul* 40:353–403.
40. Herbig K, et al. (2002) Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide and S-adenosylmethionine biosyntheses. *J Biol Chem* 277:38381–38389.
41. Selhub J, Jacques PF, Wilson PW, Rush D, Rosenberg IH (1993) Vitamin status and intake as primary determinants of homocysteinemia in an elderly population. *JAMA* 270:2693–2698.
42. Drennan CL, Matthews RG, Ludwig ML (1994) Cobalamin-dependent methionine synthase: The structure of a methylcobalamin-binding fragment and implications for other B12-dependent enzymes. *Curr Opin Struct Biol* 4:919–929.
43. Chatterjee A, et al. (2008) Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. *Nat Chem Biol* 4:758–765.
44. Droop MR (1968) Vitamin B12 and marine ecology. IV. The kinetics of uptake, growth and inhibition in *Monochrysis lutheri*.. *J Mar Biol Assoc U K* 48:689–733.
45. Pintner IJ, Altmeyer VL (1979) Vitamin B_{12}-binder and other algal inhibitors. *J Phycol* 15:391–398.
46. Sahni MK, Spanos S, Wahrman MZ, Sharma GM (2001) Marine corrinoid-binding proteins for the direct determination of vitamin B12 by radioassay. *Anal Biochem* 289: 68–76.
47. Davies AG, Leftley JW (1985) Vitamin B_{12} binding my microalgal ectocrines: Dissociation constant of the vitamin-binder complex using an ultrafiltration technique. *Mar Ecol Prog Ser* 21:267–273.
48. Sunda W, Huntsman SA (1995) Cobalt and zinc interreplacement in marine phytoplankton: Biological and geochemical implications. *Limnol Oceanogr* 40:1404–1417.
49. Allen AE, et al. (2008) Whole-cell response of the pennate diatom Phaeodactylum tricornutum to iron starvation. *Proc Natl Acad Sci USA* 105:10438–10443.
50. Kustka AB, Allen AE, Morel FMM (2007) Sequence analysis and transcriptional regulation of iron acquisition genes in two marine diatoms. *J Phycol* 43:715–729.
51. Swift DG, Taylor WR (1972) Growth of vitamin B12-limited cultures: *Thalassiosira pseudonana*, *Monochrysis*, and *Isochrysis galbana*. *J Phycol* 10:385–391.
52. Lu X, Zhu H (2005) Tube-gel digestion: A novel proteomic approach for high throughput analysis of membrane proteins. *Mol Cell Proteomics* 4:1948–1958.
53. Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 74:5383–5392.
54. Zhang B, et al. (2006) Detecting differential and correlated protein expression in label-free shotgun proteomics. *J Proteome Res* 5:2909–2918.
55. Stemmann O, Zou H, Gerber SA, Gygi SP, Kirschner MW (2001) Dual inhibition of sister chromatid separation at metaphase. *Cell* 107:715–726.
56. Siaut M, et al. (2007) Molecular toolbox for studying diatom biology in Phaeodactylum tricornutum. *Gene* 406:23–35.
57. Falciatore A, Casotti R, Leblanc C, Abrescia C, Bowler C (1999) Transformation of nonselectable reporter genes in marine diatoms. *Mar Biotechnol (NY)* 1:239–251.
58. Bertrand EM, et al. (2011) Iron limitation of a springtime bacterial and phytoplankton community in the Ross Sea: Implications for vitamin B_{12} nutrition. *Frontiers in Aquatic Microbiology* 2:160.