Title

Flexible B12 ecophysiology of Phaeocystis antarctica due to a fusion B12-independent methionine synthase with widespread homologues.

Journal

Proceedings of the National Academy of Sciences, 121(6)

Authors

Rao, Deepa

Füssy, Zoltán

Brisbin, Margaret

et al.

Publication Date

2024-02-06

DOI

10.1073/pnas.2204075121

Peer reviewed

PNAS RESEARCH ARTICLE | MICROBIOLOGY

Flexible B_{12} ecophysiology of *Phaeocystis antarctica* due to a fusion B_{12}–independent methionine synthase with widespread homologues

Deepa Rao[a,b,1], Zoltán Füssy[c], Margaret M. Brisbin[b], Matthew R. McIlvin[b], Dawn M. Moran[b], Andrew E. Allen[c,d], Michael J. Follows[a], and Mak A. Saito[b,1]

Edited by Donald Canfield, Syddansk Universitet, Odense M., Denmark; received March 19, 2022; accepted November 13, 2023

Coastal Antarctic marine ecosystems are significant in carbon cycling because of their intense seasonal phytoplankton blooms. Southern Ocean algae are primarily limited by light and iron (Fe) and can be co-limited by cobalamin (vitamin B_{12}). Micronutrient limitation controls productivity and shapes the composition of blooms which are typically dominated by either diatoms or the haptophyte *Phaeocystis antarctica*. However, the vitamin requirements and ecophysiology of the keystone species *P. antarctica* remain poorly characterized. Using cultures, physiological analysis, and comparative omics, we examined the response of *P. antarctica* to a matrix of Fe-B_{12} conditions. We show that *P. antarctica* is not auxotrophic for B_{12}, as previously suggested, and identify mechanisms underlying its B_{12} response in cultures of predominantly solitary and colonial cells. A combination of proteomics and proteogenomics reveals a B_{12}-independent methionine synthase fusion protein (MetE-fusion) that is expressed under vitamin limitation and interreplaced with the B_{12}-dependent isoform under replete conditions. Database searches return homologues of the MetE-fusion protein in multiple *Phaeocystis* species and in a wide range of marine microbes, including other photosynthetic eukaryotes with polymorphic life cycles as well as bacterioplankton. Furthermore, we find MetE-fusion homologues expressed in metaproteomic and metatranscriptomic field samples in polar and more geographically widespread regions. As climate change impacts micronutrient availability in the coastal Southern Ocean, our finding that *P. antarctica* has a flexible B_{12} metabolism has implications for its relative fitness compared to B_{12}-auxotrophic diatoms and for the detection of B_{12}-stress in a more diverse set of marine microbes.

Phaeocystis antarctica | iron | proteogenomics | methionine synthase | B_{12}

Significance

The coastal Southern Ocean is a climate-critical region of atmospheric carbon absorption because of highly productive seasonal algal blooms. Antarctic phytoplankton growth depends on micronutrient availability, with iron and B_{12} primarily influencing microbial ecosystem structure and function. Through a combination of laboratory culture and comparative 'omics, we demonstrate that *Phaeocystis antarctica* can survive without B_{12}, unlike other keystone polar phytoplankton types, due to a B_{12}-independent methionine synthase fusion protein. This protein was identified via proteogenomics, and similar related proteins are present in a variety of phytoplankton species and expressed in laboratory and environmental samples. This flexibility of *P. antarctica* to grow independent of B_{12} availability has implications for regional ecosystems and nutrient cycles.

The availability of trace metals and vitamins limits primary productivity and impacts community composition in marine microbial ecosystems (1). Vitamin B_{12} (cobalamin) is required by many eukaryotic phytoplankton (2), but is only synthesized de novo by some marine bacteria and archaea (3). Although it has been known since the 1950s that most algal cultures require a B_{12} supplement, the genetic basis for auxotrophy (vitamin-requirement) and potential for limitation in natural marine communities was only confirmed in the 2000s (2, 4–7). These studies demonstrated that B_{12} can limit phytoplankton and alter community composition in temperate (8), coastal (9), and polar environments(6, 10–12). As a result of its biological source, rapid ecological cycling, and degradation under UV exposure, B_{12} is extremely scarce in surface waters (less than 10 pM), (7). These conditions are more extreme in the Southern Ocean, where seasonal conditions create the potential for a more limited B_{12} pool (6). The basis for marine organisms' B_{12} requirement is influenced by their catalog of B_{12}-requiring enzymes and corresponding B_{12}-independent alternatives. In particular, the biosynthesis of the essential amino acid methionine can occur via B_{12}-dependent and B_{12}-independent methionine synthase enzyme isoforms (MetH and MetE, respectively), affecting an organism's sensitivity to B_{12} scarcity (2, 13–16). Auxotrophic phytoplankton only contain MetH and depend on an external supply of B_{12} that is subject to ecological and environmental processes that control its availability. However, MetH is catalytically more efficient than MetE, and cells require relatively fewer copies of MetH than MetE to maintain methionine synthesis rates (17, 18). While there are costs to maintaining MetE, including increased MetE enzyme copies and the associated Zn requirement (19), the benefits include release from B_{12} dependence and flexibility in environments with variable cobalamin supply. Eukaryotes have three strategies for interacting with environmental B_{12}: 1) be

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: drao@alum.mit.edu or msaito@whoi.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2204075121/-/DCSupplemental.

Published February 2, 2024.

B_{12}-auxotrophic and completely dependent on an exogenous supply from bacterioplankton (associated or neighboring) or the environment, 2) contain only B_{12}-independent MetE, thereby reducing their vitamin dependence (e.g., red algae, fungi, and land plants), and 3) maintain both methionine synthase isoforms allowing metabolic flexibility. Curiously, while nearly all microalgae have MetH, not all maintain MetE (14, 15). B_{12}-auxotrophy appears to have evolved several, independent times in marine algae by loss of the *metE* gene in natural communities with sufficient B_{12} supply (2).

For algae, methionine is an essential amino acid for protein synthesis, folate recycling, and one-carbon and sulfur metabolism. Its functions include being the precursor to S-adenosylmethionine (AdoMet or SAM), a key methylating agent/methyl donor and radical source involved in DNA methylation; vitamin B_1 (thiamine) synthesis; and a precursor to dimethylsulfoniopropionate (DMSP) biosynthesis a precursor of dimethyl sulfide (DMS), a climate-active gas through its formation of cloud condensation nuclei. These marine microbial metabolic pathways—influenced by vitamin availability and community dynamics—mediate the biogeochemical cycling of B_{12}, cobalt (present at the center of the B_{12} molecule), sulfur, and carbon.

In the Southern Ocean, phytoplankton are primarily limited by light and iron (Fe) but can also be co-limited by B_{12} (6, 11, 20, 21). Nutrient amendment incubations with iron and B_{12} have revealed nutrient limitation patterns among the region's main blooming phytoplankton types—diatoms and the colonial haptophyte *Phaeocystis antarctica*. Several of the Antarctic diatoms responsive to these micronutrients are B_{12} auxotrophs that only contain MetH, whereas co-occurring *P. antarctica* were found to be responsive to just iron additions (6, 10, 22–24). The underlying processes influencing the *P. antarctica* blooms and their succession by other phytoplankton communities, as well as their impact on carbon export, have long been of interest in coastal Antarctic environments such as the Ross Sea Polynya (25). Differing nutritional needs (Fe and B_{12}) and resulting ecological niches may underpin the observed variations in bloom composition and will be discussed in the *Implications* section.

Despite their significance, the B_{12} requirements and ecophysiology of *P. antarctica* and other haptophytes have been largely unknown and the results have been contradictory. As a keystone Southern Ocean phytoplankton species, *P. antarctica* contributes to primary productivity, carbon export, ocean-atmosphere climate feedbacks, and ecosystem structuring via blooms that can last for weeks to months (25–27). While *P. antarctica* contains and uses MetH in B_{12}-replete conditions (28), B_{12} limitation assays have reported no change in growth rate (29). Other studies in a closely related species *Phaeocystis globosa* have reported strain-level differences with some containing just MetH or both MetH and MetE (15, 30). Recent bioinformatic surveys concluded that all haptophytes, including *Phaeocystis* are likely B_{12} auxotrophs, as inferred by the lack of a canonical *metE* in the 19 examined genomes and transcriptomes (28, 31). However, lab cultures studies which assessed the B_{12} requirement of the haptophyte *Emiliania huxleyi* found no difference in growth with or without B_{12}, leading to the inference that it is not auxotrophic for this vitamin, though bacterial contamination was not ruled out as a potential B_{12} source (14, 32–36). Moreover, prior studies have found strain-level variability of B_{12} requirements to be common in marine microbes (30, 37).

Constraining the micronutrient requirements and strategies of bloom-forming Antarctic phytoplankton groups—diatoms and *P. antarctica*—is essential for understanding regional population and community ecology, nutrient cycling, and broader ecosystem and biogeochemical feedbacks, especially as climate change alters these micronutrients dynamics in the Antarctic; for example, through the enhanced glacial iron inputs due to warming (38–40), potentially increasing demand for other key micronutrients such as B_{12}. In this study, we present the results of a detailed investigation of the B_{12} requirements of axenic *P. antarctica*. Using physiological and multi-omic analyses, we identified a previously unrecognized MetE-fusion protein that is produced under conditions of B_{12} scarcity, providing the underlying mechanism for *P. antarctica*'s B_{12}-sparing capability and context for its ecological success. Furthermore, we found homologues of MetE-fusion in a diverse phylogenetic and geographic range of marine protists. The widespread presence of MetE-fusion suggests an important aspect of B_{12} metabolism in marine ecosystems.

Results & Discussion

A. *P. antarctica* Physiology and Multi-Omics in Fe-B_{12} Experiments.

Physiology experiments. *P. antarctica* strain CCMP 1871 was acquired from the National Center for Marine Algae and Microbiota at the Bigelow Laboratory for Ocean Sciences, treated with antibiotics, and monitored for bacterial growth until axenic. This antibiotic preparation was important in order to prevent bacterial B_{12} sources from altering experiments. After extensive acclimation to low B_{12} conditions to deplete internal reserves, *P. antarctica* was grown in triplicate semi-continuous batch cultures in a factorial matrix of low and high iron (Fe) and a wide range of B_{12} concentrations spanning plausible in situ concentrations of 0, 1×10^{-14}, 1×10^{-13}, 1×10^{-12}, 1×10^{-11}, 1×10^{-10} M B_{12}. Standard algal media uses 1×10^{-10} M B_{12} in F/2 media, which is 100 times the observed typical marine B_{12} concentrations in the pM range. Low and high iron were 3 nM and 30 nM Fe added, resulting in 41 pM and 740 pM of dissolved inorganic iron species available (Fe′) (See *SI Appendix*, Table 1 for factorial matrix of Fe and B_{12} treatments). These iron concentrations were selected based on prior experiments demonstrating iron-replete and iron-stressed responses (not severely limited in order to obtain biomass) in this *P. antarctica* strain (24). Proteomic and transcriptomic analyses were conducted on biological triplicate cultures from each treatment condition, which were harvested in the late log phase (*SI Appendix*, Fig. S1).

P. antarctica has two main morphotypes that often coexist in natural populations: solitary, flagellated cells and colonies comprised of tens to thousands of non-motile cells (27). Iron condition determined the dominant morphotype, consistent with prior observations of this strain (24). Most cultures had a mix of solitary and colonial morphotypes, with varied ratios across the Fe-B_{12} treatments. Cultures were predominantly solitary cells under low iron and colonial under high iron (Fig. 1). Two low iron cultures were observed to have only solitary cells (1×10^{-12} and 1×10^{-11} M B_{12} cultures, Fig. 1*A*), as observed previously (24). The high iron cultures were predominantly in the colonial morphotype with coexisting single cells, as is typical of natural populations. Maximal growth rates varied between the low and high iron treatments across the tested B_{12} gradient. This was likely due to the differences between morphotypes, for example, different iron-responsive proteins are used by single cells versus colonial cells (24), and production of colonial mucilage could divert resources away from growth compared to flagellate

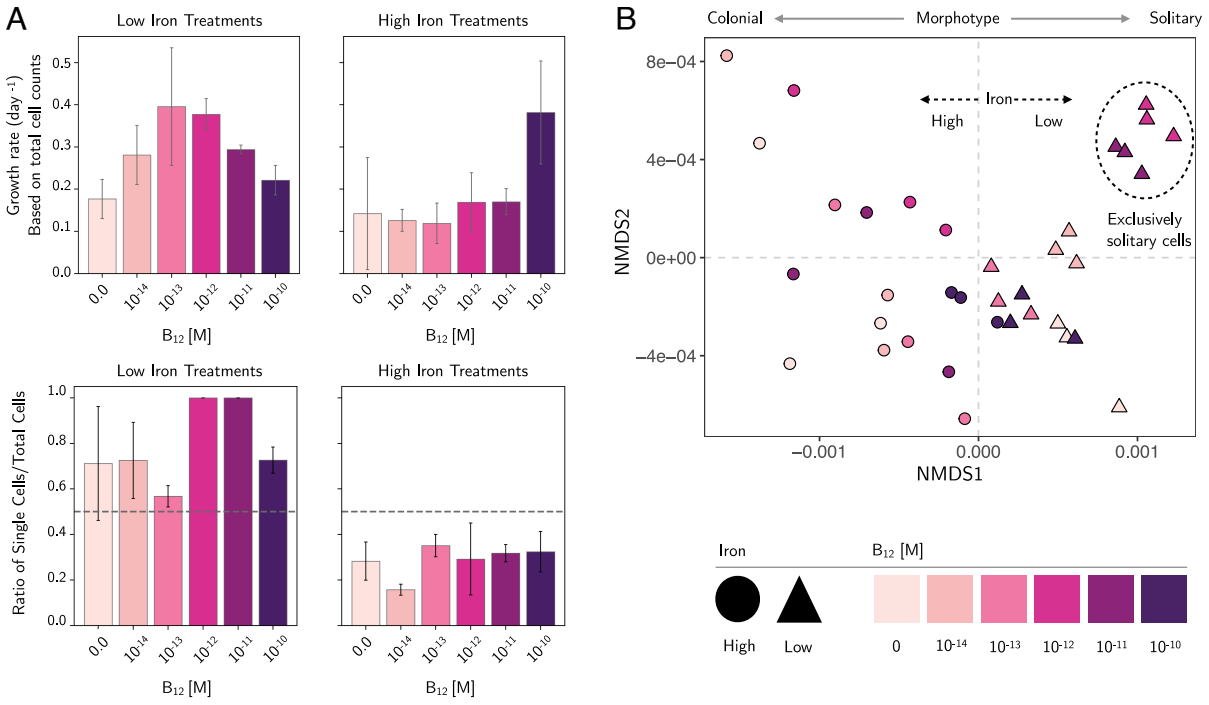

Fig. 1. (*A*) Physiological response of *P. antarctica* CCMP1871 to Fe-B_{12} limitation when grown under low and high Fe conditions with a gradient of B_{12} availability. *Top* panels show growth rates and *Bottom* panels show the morphotype ratio of single to total cells in culture at the time of harvest. (*Top*) Under low iron, the highest growth rates were in the middle ranges of B_{12} treatments, corresponding to seawater concentrations. Under high iron, growth rate was stable across B_{12} treatments, but there is a marked increase in the replete condition at 1×10^{-10} M B_{12}. (*Bottom*) The dashed horizontal line demarcates the 50% threshold of single cells to total cells counted. Error bars are the SD of the single cell ratio of three biological replicate culture samples. The low iron cultures were predominantly in the single cell morphotype, including two treatments that were purely solitary cells (1×10^{-11} and 1×10^{-12} M B_{12}), while the high iron cultures were predominantly colonial. (*B*) Results from a non-metric multidimensional scaling (NMDS) analysis using Euclidean distance between samples based on normalized expression patterns of 49,068 contigs in solitary and colonial *P. antarctica* cultures grown in a matrix of Fe-B_{12} conditions (k = 4; stress = 0.081). The shape corresponds to iron level (low, triangle; high, circle) and color to B_{12} concentration. Low and high iron treatments separate along the first NMDS axis, with low iron treatments on the left and high iron treatments on the right. Solitary cell culture replicates cluster separately along NMDS axis 1 (*Right*, 1×10^{-11} M and 1×10^{-12} M B_{12}) compared to mixed morphotype cultures and predominantly colonial cultures. Samples further separate by B_{12} concentration on the second NMDS axis; clear clustering by B_{12} concentration is visible among low-iron samples. Treatment replicates are more similar to each other than other samples across the entire experiment (ANOSIM R = 0.6357, P = 0.001, permutations = 999). A two-way PERMANOVA identified significant differences in transcriptome-wide gene expression patterns between treatments, with iron explaining 24.7% of the variation (P = 0.001, F = 12.3834) and B_{12} explaining 17.3% of the variation (P = 0.008, F = 1.7277) (see *SI Appendix*, Table 2 for PERMANOVA results). The authors' GitHub repository contains the transcriptome dataset and code for the NMDS, PERMANOVA, and ANOSIM analyses.

cells. Under low iron, growth rates were maximal within the observed environmental range of B_{12} concentrations 1×10^{-13} to 1×10^{-12} M and declined at lower and higher B_{12} concentrations, toward 0 and 1×10^{-10} M, respectively. Maximal growth rates under co-limitation conditions have been previously observed in other phytoplankton and suggest an ecophysiology adapted to typical in situ nutrient-limited conditions (41). In high iron, the growth rate was relatively stable, except for a roughly twofold increase under the highest B_{12} treatment (1×10^{-10} M). Cell density was markedly lower in low iron and low B_{12} (0 to 1×10^{-14} M B_{12}, notated as -Fe/-B_{12}) treatments, and increased at the intermediate vitamin concentrations (*SI Appendix*, Fig. S2).

NMDS and differential gene expression analysis. We used quantitative proteomic and transcriptomic methods to investigate the response of *P. antarctica* to cobalamin availability under low and high iron concentrations that promoted single cell and colonial morphotypes, respectively. A non-metric multidimensional scaling analysis (NMDS) was used to analyze gene expression patterns in the *P. antarctica* transcriptome (RNA-seq) under the tested range of Fe-B_{12} conditions, and resulted in a low stress value (<0.1, see Fig. 1*B*). Aligned with the NMDS, the clustering of biological replicates by Fe-B_{12} treatment is statistically significant according to an Analysis of Similarity (ANOSIM R = 0.6357, P = 0.001, number of permutations = 999); the majority of variation in RNA expression patterns is between treatments, consistent with observed Fe-B_{12} physiological responses (Fig. 1*A*). A two-way permutational multivariate analysis of variance (PERMANOVA) identified strong and significant transcriptome-wide differences in gene expression patterns between the experimental conditions with iron explaining 24.7% of the variation (r^2 = 0.24756, P = 0.001, F = 12.3834) and B_{12} explaining 17.3% of the variation (r^2 = 0.17269, P = 0.008, F = 1.7277) (see analysis results in *SI Appendix*, Table 2) (42, 43). NMDS Axis 1 corresponds to the interrelated iron conditions and morphotype ratios observed in Fig. 1*A*, with primarily colonial cells on the left and increasing to purely solitary, flagellated cells to the right, with a split between high vs. low iron conditions close to 0. Notably, the exclusively solitary cell cultures (1×10^{-11} M and 1×10^{-12} M B_{12} in low iron observed in Panel *A*) cluster together on the far right of NMDS Axis 1 away from other sample conditions, consistent with a distinct morphotype response to nutrient availability. That colonial and solitary cells' transcriptomes cluster separately has also been observed in *P. globosa* cultures (44) and is also consistent with *P. antarctica* morphological changes in response to iron condition (24). The remaining low iron

treatments (triangles, mixed morphotype ratio) are concentrated in the right half of the NMDS1 axis, with $-B_{12}$ and $+B_{12}$ samples clustered close together in both NMDS dimensions, in alignment with their similar growth rates and morphotype ratios (Fig. 1*A*). In the NMDS2 dimension, the predominantly colonial cells in high iron have a larger spread that followed increasing B_{12} availability for the lowest three B_{12} treatments and the highest three treatments migrating toward the center as growth rate increases. The highest B_{12} treatments cluster closely together (dark purple circles and triangles) under both high and low iron, suggesting a coherent transcriptomic response to high B_{12}, which is supported by an upregulation of MetH in both transcripts and proteins (*SI Appendix*, Fig. S3). Overall, the NMDS indicates that low-iron treatments are more similar to one another than to high-iron treatments, and vice versa. However, at the highest B_{12} concentration, both the low- and high-iron treatment replicates cluster closely and on the lower end of the NMDS2 axis, suggesting a B_{12} effect, as supported by changing MetE and MetH transcript expression and peptide abundance (*SI Appendix*, Fig. S3). Within the low-iron treatments, there are distinct and coherent clusters of replicates by B_{12} treatment along NMDS2 (relative to the predominantly colonial, high-iron cultures), in alignment with more variable growth rates along the tested B_{12} conditions (Fig. 1) and a larger set of significantly differentially expressed proteins between 0 M and 1×10^{-10} M B_{12} under low iron vs. high iron (see Fig. 2 and related discussion).

Discovery and regulation of methionine synthase in response to B_{12}. Prior to this study, only the B_{12}-requiring MetH was known to occur in *P. antarctica* and other haptophytes (31). However, comparative transcriptomics (RNA-seq) and proteomics (global and targeted MS peptides) revealed changes in methionine synthase isoforms across the gradient of B_{12} conditions. We observed a protein (and corresponding gene) of unknown function to be among the most significantly differentially abundant proteins (DAPs) in response to $-B_{12}$ /$+B_{12}$ (0 M, 1×10^{-10} M), being 2- to 4-fold more abundant in $-B_{12}$ treatments see Fig. 2, estimated FDR/FPR < 0.05, power law global error model [PLGEM], low iron ± B_{12} treatments (45, 46). Initially, two MetE proteins were annotated (MetE and MetE C-terminal), as a result of multiple assemblies in the proteomic database (*SI Appendix*, Fig. S4 and *Discussion*). We have determined that these are part of the same protein: a B_{12}-independent MetE-fusion (details in *Results* Section B).

In a comparison of the normalized transcript and protein abundances of MetH and two MetE-fusion contigs across matrix of Fe-B_{12} treatments, MetE was more abundant under low B_{12} conditions in both iron levels, with a steady pattern of decline as vitamin concentrations increased (see Fig. 2 for proteins and *SI Appendix*, Fig. S3 for transcripts). In the highest B_{12} treatments (1×10^{-10} M), there is a distinct increase in MetH and drop in MetE transcripts and proteins. The transcript and protein abundances reveal an inter-replacement of methionine synthase isoforms that depends on B_{12} availability, as observed in other MetE/MetH containing protists (*Phaeodactylum tricornutum*, *Pseudo-nitzschia granii*, *Chlamydomonas reinhardtii*) (16, 47, 48). This trade-off is a signature that the organism contains multiple isoforms of methionine synthase.

P. antarctica morphotypes appear to have different metabolic responses to B_{12} availability. Overall, *P. antarctica* cultures shared 143 significantly differentially abundant proteins (DAPs) expressed in $-/+$ B_{12} conditions (0 M and 1×10^{-10} M), across both low and high iron treatments (Fig. 2, DAPs determined with using PLGEM, FDR/FPR < 0.05). Low iron cultures had more DAPs (649) in response to $-/+$ B_{12} than high iron cultures (441) (Fig. 2 and *SI Appendix*, Figs. S6 and S7). The predominantly solitary cells in low iron steadily up-regulate *metH*/MetH transcripts and proteins in response to vitamin availability (*SI Appendix*, Fig. S3). The use of MetH over MetE proteins starts trading off near environmental concentrations ($\sim 1 \times 10^{-12}$ M B_{12}). Solitary cells may be more responsive to lower levels of B_{12} at which they up-regulate *metH*/MetH, as observed in both the transcript and proteomic data (Fig. 2). The predominantly colonial cells in high iron appear to need a higher B_{12} concentration to switch methionine synthase isoforms. This could be due to diffusive limitation of B_{12} into larger colonial cells and colonies (49).

While *metH* transcripts increased with increasing B_{12} availability from 1×10^{-13} M to 1×10^{-10} M, curiously, they were also potentially enhanced in the lowest B_{12} treatments (0 M) in both low and high iron experiments, despite high variability in the low iron treatments (see U-shaped responses in *SI Appendix*, Fig. S3). A similar observation of increased *metH* transcripts in -B_{12} conditions was observed in *P. granii*, where it was suspected that increased *metH* expression may compensate for reduced B_{12} to maintain levels of methionine synthesis (23). However, while that pattern is observed in *P. antarctica* transcripts, the proteomics data show an absence of MetH in $-B_{12}$. While the detection of the MetH enzyme is more challenging due to its lower abundance (as a result of higher enzymatic efficiency), the expression patterns taken together indicate a trade-off between isoforms. These results are consistent with prior studies in a chlorophyte and diatoms where vitamin B_{12} has a negative regulatory effect on *metE*/MetE gene and protein expression (6, 14, 15). Finally, the persistence of *metE*/MetE transcripts and proteins (*SI Appendix*, Figs. S3 and S5–S7), even at low abundance at high B_{12} concentrations, suggests an important ecophysiological role for this protein that could allow for flexibility under variable B_{12} conditions, such as involvement in allosteric regulation (50) and/or protein methylation in flagella (51, 52). In *Chlamydomonas*, MetE has been demonstrated to be an abundant protein within flagella that is used for protein methylation during flagellar assembly and disassembly (51, 52) and could have a similar role in *P. antarctica* as it changes morphotypes.

B_{12}-responsive proteins in P. antarctica suggest larger role in metabolism. We identified both an expected set of B_{12}-influenced proteins in *P. antarctica* that are involved in the methionine cycle (MetE/MetH/MetK/SAH1), heme biosynthesis, methyltransferase reactions, folate cycling (MetF), ubiquinone metabolism, and other systems (OmpA/MotB, CobW, MmsA) that await further exploration (some of which are highlighted in Fig. 2). Several significantly differentially abundant proteins from *P. antarctica* correspond to a set of B_{12}-bound proteins identified with a chemical probe in the heterotrophic bacterium *Halomonas* sp. HL-48 (50). Many of those proteins had prior unknown associations with B_{12}, but are involved in methionine, folate, and ubiquinone pathways, as well as DNA, RNA, and protein synthesis, expanding the potential role of B_{12} in cell metabolism (see *SI Appendix* for a discussion of shared DAPs in *P. antarctica* with those found in ref. 50). While it is known that MetH uses B_{12} as a cofactor, the B_{12}-probe was found bound to MetE in a *Halomonas* sp. (50). In bacteria, cobalamin-binding riboswitches can regulate MetE translation, but it was hypothesized that B_{12} may also be an allosteric regulator (50). The MetE-fusion transcript and protein abundances with varying B_{12} treatments described in this study align with this hypothesis. Nearly, the entire heme biosynthesis

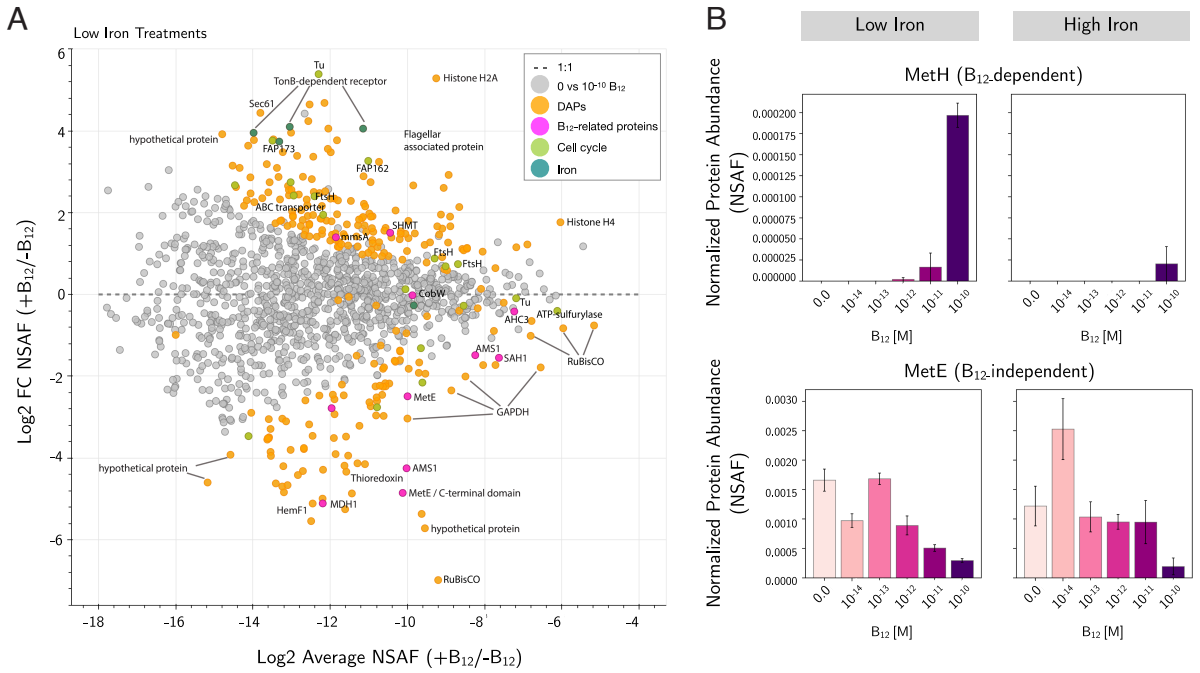

Fig. 2. *P. antarctica* CCMP1871 protein expression patterns across various B_{12} concentrations. (*A*) MA plots of proteomes in response to ± B_{12} (low Fe, 0 M and 1×10^{-10} M B_{12}). MA plots show the relationship between average protein abundance (log_2 NSAF) and fold-change (log_2 FC) between treatments. Each protein is represented by a gray dot. Colored dots were found to be significantly differentially abundant proteins (DAPs) (PLGEM, FDR/FPR < 0.05), with B_{12}-related proteins highlighted in pink, cell-cycle related proteins in green, and iron-related proteins in teal, all other DAPs are in yellow. Select proteins are annotated to facilitate discussion, for example, the methionine synthase isoforms (MetE and MetH). (*B*) Protein expression patterns of methionine synthase isoforms, B_{12}-dependent (MetH, contig_51544_3_2534_+) and B_{12}-independent (MetE, contig_21595_2_1159_-). *P. antarctica* inter-replaces MetE with MetH when B_{12} is available. See *SI Appendix* for an additional MA plot under high iron ± B_{12} and visit the authors' GitHub repository to examine proteins in an interactive figure of Panel *A* (low iron ± B_{12} DAPs).

pathway was significantly differentially abundant in response to B_{12}, suggesting a possible regulatory role in directing heme biosynthesis and its by-products like chlorophyll, phycobilins, and cytochromes, as suggested by ref. (50) (see *SI Appendix*, Fig. S8 for the expression patterns of heme pathway proteins). Finally, several differentially abundant proteins of unknown function in *P. antarctica* were also identified via the B_{12}-probe. Some of the highest fold-change proteins in +B_{12} and +Fe were the OmpA/MotB-like family proteins involved in cell-to-cell interactions (53). Similarly, CobW was responsive to +B_{12} and is hypothesized to be involved in coordination of the cobalt atom in the ring structure of B_{12} (54). Cell cycle and motility proteins like TuF and flagellar proteins (FAP 162 and FAP 173) were responsive to B_{12}. Furthermore, we found several of these genes and proteins in a field metatranscriptomic and metaproteomic sample of Ross Sea *P. antarctica* bloom [CORSACS, 2005 (24)], including PP-binding [PF00550], MetE C-terminal [PF01717], OmpA [PF00691], SOUL heme-binding protein [PF04832], AMS1 [PF02773], and ubiquitin family /thioredoxin [PF00240, PF00085], among others (data are provided in *SI Appendix*). This surprising coherence in B_{12}-related proteins in *P. antarctica* and *Halomonas*, and their observance in field samples in a known B_{12}-limited region, implies connections in B_{12} systems biology despite the large taxonomic differences.

B. Proteogenomic Identification and Sequence Characterization of a MetE-Fusion Protein.

Background on methionine synthases. Methionine synthases use different methyl donors to produce methionine by catalyzing the methylation of homocysteine: MetH (EC 2.1.1.13) uses cobalamin and MetE (EC 2.1.1.14) uses 5-methyl-THF-Glu_n. Canonical MetEs are smaller than MetH (735 vs. 1,227 AA in *E. coli*) and comprise N- and C-terminal domains connected by a linker region (55, 56). Structural analysis of MetE from diverse organisms like bacteria (*Escherichia coli* and *Thermotoga maritima*), fungi (*Neurospora crassa*), and protists (*Fragilariopsis cylindrus*) reveal common features: 1) a double-barrel conformation surrounding an active site (methyl-THF), 2) a conserved Zn^{2+}-binding site in the C-terminal barrel, 3) an N-terminal barrel that evolved via intra-gene duplication of the C-terminal domain (~ 340 AA), and 4) a likely conformational change that enables domain rearrangement prior to methyl transfer (56). Recently, a novel group of ancient B_{12}-independent methionine synthases, dubbed core-MetE, have been described in archaea with homologues in bacteria. Core-MetEs are approximately half the size of canonical MetEs and the sequences correspond to the C-terminal domain with the active site (57). These core-MetE enzymes obtain methyl groups from other corrinoids (e.g., methylcobalamin) instead of folate, indicating a greater diversity of potential MetE proteins than previously known (57, 58).

The MetE-fusion protein described below differs from canonical MetEs in only having one C-terminal domain maintaining the Zn^{2+}-binding site, with several additional conserved domains that together create a larger fusion protein, which likely obscured its prior identification. While the function of these domains is as yet undetermined, they are present in the MetE-fusion protein from *P. antarctica* and *P. globosa* as well as many other marine protists, suggesting an essential function. In the following sections, we describe the proteogenomic identification of MetE-fusion, its characteristics inferred from conserved domains, sequence alignments, and predicted structure that support this characterization.

Evidence for MetE-fusion protein in P. antarctica. As described above, the global proteomic and transcriptome analyses of B_{12}-limited cultures clearly identified peptides and transcripts

that contained a MetE domain in *P. antarctica*. Two contigs with similar expression patterns were annotated as MetE in the transcriptomic assembly of *P. antarctica* CCMP1871 (see *SI Appendix*, Fig. S4 for sequences and corresponding detected tryptic peptides). A BLASTp alignment confirmed that these two contigs mapped directly to a single contig in the transcriptomic assembly of *P. antarctica* CCMP1374 and in a Ross Sea metatranscriptome (24). These alignments suggest that the original assembly of the CCMP1871 transcriptome split the MetE-fusion into two smaller contigs. These findings informed the re-examination of the *P. antarctica* CCMP1374 draft genome (Phaant1 scaffold10) and the experiment transcriptome, resulting in the in silico reconstruction of the full-length coding DNA (CDS) and protein sequences. Henceforth, this protein sequence is referred to as MetE-fusion.

MetE-fusion conserved domains and predicted structure. A protein domain search using InterProScan revealed that this *P. antarctica* MetE-fusion is a multi-domain protein consisting of three conserved domains and is distinct from canonical MetEs. Based on a multiple sequence alignment, the MetE domain best matches the C-terminal conserved domain of canonical MetEs (URO-D/CIMS family) in other algae, fungi, protists, bacteria, and archaea (Fig. 3). However, the MetE conserved domain of MetE-fusion is roughly half the size (~ 320 to 380 AA) of canonical MetEs, corresponding to size of the more ancient C-terminal portion and core-MetEs. Compared to canonical MetEs, the MetE-fusion homologues are more similar to one another both in the MetE C-terminal domain region and broadly in their set and arrangement of conserved domains, see Fig. 4.

Notably, the MetE-fusion contains residues for the Zn^{2+}-binding domain at the active site ($HXCX_nC$) that are broadly conserved in all known MetE homologues (see Fig. 3 for cropped alignment with key residues highlighted in red and *SI Appendix* for the complete multiple-sequence alignment) (56, 57, 59, 61). This conserved motif is essential for L-homocysteine binding and activation (62). Some organisms like B_{12}-dependent green algae *Volvox carteri* and *Gonium pectorale* have mutated *metE* pseudogenes, that lack these conserved residues, resulting in loss of a functional enzyme (14). Along with the multi-omic expression patterns, the conserved $HXCX_nC$ residues suggest that this MetE-fusion is not a pseudogene. The MetE annotation and the Zn^{2+}-binding site were also supported by structural prediction using SWISS-MODEL: Matches were found to over 300 Protein Data Bank (PDB) templates (SWISS-MODEL January 2021), of which the top matches were to MetE, resulting in five models (PDB: 2nq5.1.A, 4ztx.1.A, 1ypx.1.A, 3t0c.1.A, and 4ap1.1.A, see *SI Appendix*). Ligand modeling identified complexed Zn^{2+} ions in some of the models bound by His, Cys, Cys, and Glu—a known feature of MetE proteins (56, 57). An example predicted protein structure corresponding to the MetE domain Zn^{2+}-binding site is included with Fig. 3, from a best match to B_{12}-independent methionine synthase from the model fungus *Neurospora crassa* (QMEAN = −4.90, GMQE = 0.19, Sequence Identity = 20.06%, Range = 192 to 584) (PDB No. 4ZTX).

Common conserved domains (CDs, determined via NCBI Conserved Domain Search and InterProScan) among MetE-fusion homologues include 1) an N-terminal domain involved in polyketide synthase/phosphopantetheine binding of the acyl carrier protein (PP-binding, ACP, PKS), 2) a middle domain that is the URO-D/Cobalamin-independent methionine synthase (CIMS/MetE), and 3) a C-terminal pyridine nucleotide-disulfide oxidoreductase that is described as FAD-dependent Pyr_redox_2 and as the putative oxidoreductase domain CzcO. This domain structure also appears to be conserved across taxa, for example

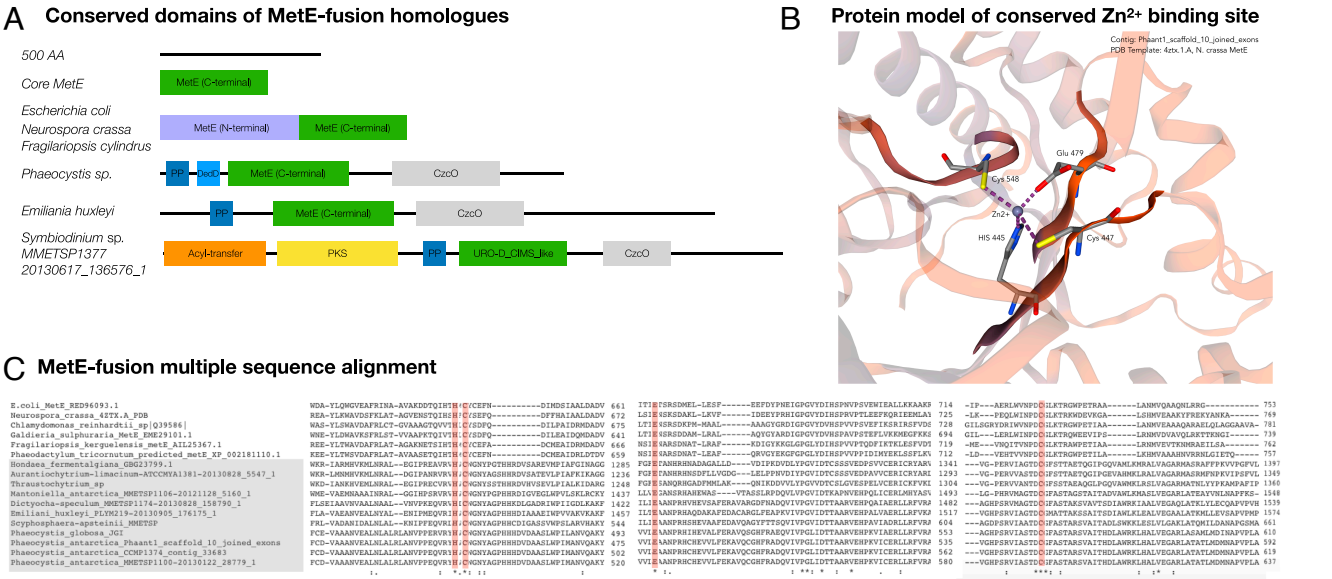

Fig. 3. Analysis of the MetE-fusion amino acid sequence via conserved domain search, multiple sequence alignment, and protein structure homology modeling. (*A*) Selected conserved domain diagrams of canonical MetEs from diverse organisms and MetE-fusion sequences from algae, including a 500 amino acid scale bar for reference. Full conserved domain arrangements and multiple sequence alignment of MetE and MetE-fusion homologues can be found in *SI Appendix*, Figs. S9 and S10. (*B*) Predicted 3D structure of computational protein model of *P. antarctica* MetE-fusion (Phaant1_scaffold10) to a template of MetE from *Neurospora crassa* (QMEAN = −4.90, Identity = 20.06%, PDB No. 4ZTX) rendered in cartoon representation. This figure is cropped to the region of the modeled metal active site structure with Zn^{2+}-binding residues His, Cys, Cys, and Glu, conserved in all known MetEs. (*C*) Multiple sequence alignment of known MetE and MetE-fusion amino acid sequences identified in this study (names highlighted in gray). For clarity, only a subset of representative species are shown. The alignment is cropped to focus on the region spanning the conserved Zn^{2+} binding residues for catalytic activity highlighted in red, ($HXCX_nC$) - [Cys_2HisGlu], found in bacterial, fungal, plant, core-MetE, and selected MetE-fusion homologues from haptophytes and other marine protists (56, 57, 59–61). Legend: Below each site (i.e., position) of the protein sequence alignment is a key denoting conserved sites (*), sites with conservative replacements (:), sites with semi-conservative replacements (.), and sites with non-conservative replacements ( ).

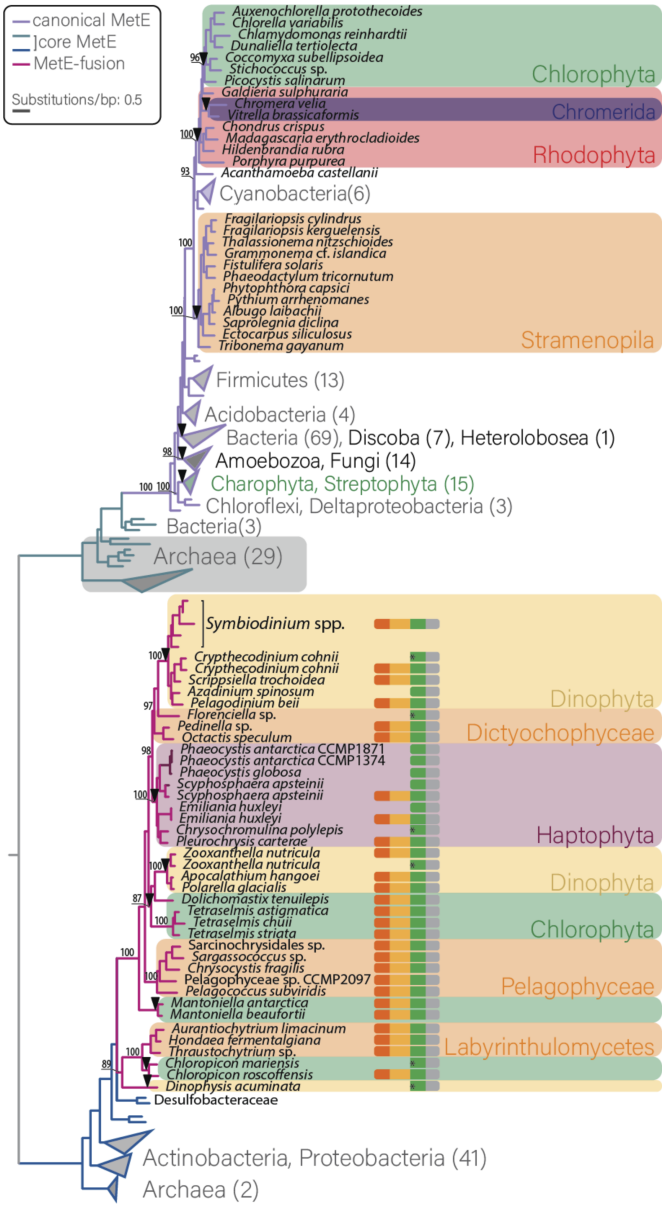

Fig. 4. Maximum Likelihood (ML) phylogeny of the MetE C-terminal domain from core, canonical, and MetE-fusion methionine synthases, shown having different branch colors (legend). These enzymes are found across domains and in a range of eukaryotes, suggesting multiple horizontal gene transfer events responsible for their distribution (marked by triangles). Canonical MetE evolved via intragenic duplication, whereas MetE-fusion combined an N-terminal acyltransferase (AT) domain, a polyketide synthase (PKS) domain, a phosphopantetheine-binding domain PP-binding), a MetE domain and a C-terminal CzcO oxidoreductase (presented in more detail in Fig. 4). The domain structure of each MetE-fusion sequence is shown (orange = AT; yellow = PKS; green = MetE C-terminal; gray = CzcO), highlighting that the N-terminal AT and PKS domains were perhaps lost in Haptophytes only, whereas the truncation in other taxa is probably artifactual (asterisk marks a missing N-terminal methionine). Notably, *E. huxleyi* encodes two MetE-fusion-like homologues, one possessing and one lacking this ACS/PKS extension, whereas in *Phaeocystis* spp. the ACS/PKS and MetE-fusion are encoded as two separate genes. Homologues were identified via searches against NCBI-nr, MMETSP, and genomic assemblies of haptophytes *P. antarctica* (draft) and *E. huxleyi* from JGI (Joint Genome Institute). The number of sequences constituting collapsed clades is shown in parentheses. Ultra-fast bootstrap support values are indicated at important branches (64). See *SI Appendix*, Fig. S10 for full phylogeny.

in closely related organisms like the temperate/tropical sister species *Phaeocystis globosa* and in more distantly related algae like *Symbiodinium* (Fig. 4 and *SI Appendix*, Figs. S9–S11).

While further biochemical inspection is required, the protein domain arrangement suggests the following possible related functions with the methionine synthase domain. The PP-binding domains are found in multi-domain enzymes polyketide synthases, where they act as a "swinging arm" for the attachment of vitamins, activated fatty acids, and amino acid groups. ACP and PKS enzymes have been previously identified in haptophytes, but not contiguous with MetE (28). The fusion protein contains additional motifs on the amino and carboxy sides, including CzcO (COG2072) that has been associated with ion transport, including for Zn and Co. There are multiple cysteine and histidine residues present in the CzcO motif in *P. antarctica*, although they do not align with those in COG2072. Given the Zn requirement of MetE, we speculate that this CzcO motif plays a role in Zn trafficking or insertion. Many other homologues contain this motif as well. The C-terminal Pyr_redox_2 domain also has known functions with iron-sulfur cluster binding and the glutamate biosynthesis process, suggesting a potential link between methionine synthase, iron, and sulfur metabolism in *P. antarctica*. Finally, because MetE-fusions lack the N-terminal folate-binding domain of canonical MetEs, it is likely that MetE-fusion uses alternative methyl donors like other core methionine synthases (58). Recent studies have identified folate-independent methionine synthases that utilize methylated nutrients like S-methylmethionine or glycine betaine, which could be potential targets for further analysis (58, 63).

C. Phylogeny and Distribution of MetE-Fusion Homologues in Protistan Cultures and the Environment. The MetE fusion protein identified from *P. antarctica* is present in a diverse range of marine microbes (protists and bacterioplankton) in global environmental metaproteomes and metatranscriptomes, implying its importance as a B_{12}-sparing mechanism in the ocean. A phylogenetic analysis was performed for homologues of the MetE-fusion protein identified in *P. antarctica* (Phaant1_scaffold_10) to examine its distribution and address its relation to canonical diatom MetEs (Fig. 4). A BLASTp search against the full NCBI-nr and Marine Microbial Eukaryotic Transcriptome Sequencing Project (MMETSP) databases returned homologues with best hits to a range of protistan and prokaryotic sequences with MetE-like domains. All identified MetE-fusion homologues contain a conserved domain region that is roughly ~ 320 to 380 AA long and resembles the C-terminal domain of MetE (see Fig. 3 and *SI Appendix* for conserved domain arrangements of homologues).

Searches against NCBI-nr returned a range of prokaryote MetE-like protein domains and few eukaryotic matches to MetE-like proteins preceded by PP-binding domains from *Emiliania huxleyi*, *Symbiodinium microadriaticum*, *Hondaea fermentalgiana*, and *Polarella glacialis*). Eukaryotic hits were often annotated as unnamed or hypothetical proteins or according to one of the other conserved domains, demonstrating the challenge of identifying MetE-fusion by automated annotation. Searches against MMETSP yielded numerous MetE-fusion matches to a wide range of marine protists, including dinophytes, heterokontophytes, chlorophytes, and labyrinthulids, and contained the multiple domains described above (MetE, PP-binding, and FAD/NAPH binding pyridine nucleotide-disulfide oxidoreductase domains), and mostly extended further on the N terminus, where they contain ketoacylsynthase and acyltransferase domains. Only haptophytes seem to encode MetE-fusion proteins that lack this N-terminal extension, for which we found support in the genomic data of *P. antarctica* and *E. huxleyi*. Canonical diatom

MetEs are distantly related to MetE-fusion proteins. There is a deep branch between prokaryotic and eukaryotic sequences, and inspection of the protein domain architecture reveals the reason for this divergence: The prokaryotic sequences align with the C-terminal MetE conserved domain, and do not contain the other CDs found in eukaryotic MetE-fusion proteins (Fig. 3).

The topology of the phylogenetic tree of the C-terminal domain of MetE-fusion and canonical MetEs reveals that both arose via horizontal gene transfer to eukaryotes from different core-MetEs. The canonical MetE evolved via intragenic duplication in Bacteria, resulting in the well-characterized double-barrel structure with a linker region (56). This "double-domain" MetE was likely introduced into eukaryotes multiple times, at least in red algae, later transferred to Chromerids and Chlorophytes; in Charophytes and Streptophytes; in Stramenopiles; and among heterotrophic clades in Discoba, Heterolobosea, and Opisthokonta. The MetE-fusion evolved in eukaryotes employing a different core-MetE domain, then likely passed between Stramenopile lineages (Labyrinthulids, Pelagophytes, Dictyochophytes) and lineages of green algae, dinoflagellates, and haptophytes (horizontal gene transfer events suggested in Fig. 4).

MetE-fusion in Phaeocystis sp. and other polar protists. The MetE-fusion protein has homologues in *Phaeocystis* species, including *P. antarctica* strains (CCMP1871 and CCMP1374) and *P. globosa*. In addition to multi-omic analysis in this study, the MetE-fusion is found expressed in another laboratory experiment of *P. antarctica* CCMP1374 (see Wu et al. 2019 supplemental materials of Mn-Fe limitation experiment cultures, contig_33683_53_3002_+) (65). In addition to *P. antarctica* strains, MetE-fusion homologues were found in several polar species: *Apocalathium* (*Scrippsiella*) *hangoei*, *Pelagophyceae* CCMP2097, *Pedinella* sp. CCMP2098, *Mantoniella antarctica*, and *Polarella glacialis*—a recently sequenced polar dinoflagellate (NCBI-nr) (64). There is no apparent evolutionary connection among these species, as has been noted for organisms that contain canonical MetE (66). However, there may be a biogeographical basis for increased retention and expression of MetE-fusion genes and proteins in regions of scarce B_{12} like the Southern Ocean, as has been observed in polar diatoms (like *F. cylindrus*) (23) and corroborated in MetE-fusion gene searches in the TARA Ocean Gene Atlas (*SI Appendix*, Fig. S12 and discussion in *SI Appendix*).

There are two key physiological characteristics in common among several protist species on this phylogenetic tree of the MetE-fusion protein: most have a complex polymorphic life cycle (multiple life stages) and nearly all have a flagellated motile stage. The MetE-fusion sequences were identified in other marine microeukaryotes, including haptophytes (*Phaeocystis*, *Emiliania*, *Scyphosphaera*, *Pleurochrysis*, and *Chrysochromulina*), dinoflagellates (*Symbiodinium*, *Crypthecodinium*, *Scrippsiella*, *Brandtodinium*, *Pelagodinium*, and *Azadinium*), and chlorophytes (*Tetraselmis*). More distant matches on a separate branch were found to members of labyrinthulids (*Stramenopila*), though these proteins cluster most closely to *Desulfobacteraceae bacterium* epoxyalkane coenzyme M transferase. Some of these protists transition between haploid and diploid stages (*Phaeocystis* and *Emiliania*) or have vegetative stages. Several genera are known to have colony-forming species (*Phaeocystis, Tetraselmis, Pelagococcus, and Aurantiochytrium*). Others are known to have symbiotic associations (*Symbiodinium microadriaticum*, *Pelagodinium beii* with foraminifera, and *Brandtodinium nuctuiluca* with radiolarians). In addition, many of these organisms are known DMSP producers (e.g., *P. antarctica, E. huxleyi, Tetraselmis*) and prior studies have observed connections between methionine synthase activity and DMSP production (67, 68). Finally, prior studies have identified MetE in *Chlamydomonas reinhardtii*, where it is localized and most abundant in regenerating and resorbing flagella (51, 52). Given *P. antarctica*'s polymorphic life cycle and the potential challenges of trafficking B_{12} to flagella or among multiple colonial cells, maintaining MetH and a MetE-fusion could provide additional advantages. These are surprising and intriguing patterns among the eukaryotic protein homologue results that warrant closer inspection into the phylogenetic relationships among these groups and possible related metabolic pathways, especially concerning B_{12}, cell cycle regulation, and DMSP.

Global distribution of MetE-fusion in ocean metatranscriptomes. We scanned the MATOU (Marine Atlas of Tara Ocean Unigenes) metatranscriptomic occurrences database to explore broad size- and depth-related distribution of MetE-fusion homologues in the ocean, along with their taxonomy and occurrence within samples (69). Over 2,000 sequences with homology to the MetE-fusion protein were identified in the TARA Ocean Gene Atlas (OGA) (*SI Appendix*, Fig. S12). Many of the sequences are multi-domain proteins containing a similar set of conserved domains. The transcripts are predominantly found in the surface water and deep chlorophyll maximum depth samples, corresponding to the habitat of photosynthetic eukaryotes, and are widely distributed in the marine environment, with higher relative abundance in the Southern Ocean stations (*SI Appendix*, Fig. S12). Homologue transcripts were found in size fractions corresponding to solitary *Phaeocystis* cells and larger microalgal types (0.8 to 5, 5 to 20, 20 to 180, 180 to 2,000 μm). The taxonomy of protein homologues was 51% *Collozoum* (a colonial radiolarian genus), 39% Dinophyceae (dinoflagellates), and approximately 9% Haptophyceae (*Phaeocystis* spp. and others) (*SI Appendix*, Fig. S12). This taxonomic distribution is notable because *Phaeocystis* spp. are known symbionts with radiolaria and acantharia, and are prey to dinoflagellates, including a kleptoplastic relationship in the Ross Sea (70, 71). Such intertwined lifestyles can influence each species' evolution, as exemplified by *Phaeocystis* and *Acantharia*, whose symbiotic relationships have been inferred to be flexible—depending more on biogeography than on taxonomic specificity (70).

The widespread presence of MetE-fusion sequences in marine protists and bacterioplankton and meta-omic datasets provides evidence that there exists an overlooked group of B_{12}-independent methionine synthases in marine ecosystems. Given the large changes in MetE and MetH across B_{12} gradients observed in this study (Fig. 2), measurements of these proteins and transcripts could potentially be deployed as biomarkers for B_{12} stress, either separately or as a ratio to each other. While B_{12}-auxotrophic diatoms may be the most responsive to cobalamin amendment experiments, the MetE-fusion can give an indication of the B_{12}-condition of *P. antarctica* in the Southern Ocean and offer insights to its ecophysiology and relative fitness in the community.

Implications

Our study of the vitamin B_{12}-mediated regulation of gene and protein expression in *P. antarctica* has provided key insights into the nutritional requirements of this keystone species, and in the process revealed a previously overlooked group of putative B_{12}-independent methionine synthases. *P. antarctica* has a metabolic and phenotypic plasticity to iron and B_{12} that is a key aspect of its ecophysiology in the Southern Ocean. Its ability to grow without B_{12}—unlike other co-occurring Southern Ocean diatoms—likely allows it to persist and thrive in massive blooms through extreme seasonal conditions. Specifically, the

progression from early spring blooms of solitary, flagellated to colonies of *P. antarctica* to bacterioplankton and then diatoms (25–27), align with the current understanding of inputs and biosynthesis of iron and B_{12}, respectively, and the biological demand for each (*SI Appendix*, Fig. S13) (6, 20, 38–40, 72–74). Our finding that solitary *P. antarctica* cells have optimal growth rates under reduced Fe-B_{12} (0.1 pM to 1 pM) (Fig. 1), indicate a niche for this morphotype in the early spring when these micronutrients can become depleted with increased biological demand (6, 20). Maximal growth rates under co-limitation have also been observed in other phytoplankton (e.g., *Trichodesmium*) (41). By regulating its methionine synthase isoforms in response to B_{12} availability (among a suite of other proteomic changes), *P. antarctica* has a potential advantage to bloom in the early austral spring, when bacterial production is relatively low (75) and inputs of trace metals may increase with glacial and sea ice melt (38–40, 76). Increased bacterial growth and thereby B_{12} availability, may be enhanced by the growth of *P. antarctica* colonies and their associated dissolved organic carbon (19, 75). Our results show that *P. antarctica* can switch to using MetH when B_{12} is abundant, potentially further stimulating population growth. As the austral summer progresses, increased bacterioplankton stocks (75) likely increase the availability of B_{12} likely contributes to diatom blooms of fast-growing B_{12}-auxotrophic species that are relieved of vitamin limitation, as observed in Ross Sea experiments of natural algal communities, where auxotrophic *Pseudo-nitzschia subcurvata* increased in abundance in response to iron and B_{12} additions, while non-auxotrophic (*metE*-containing) *P. antarctica* and *F. cylindrus* did not increase in biomass under +Fe+B_{12} treatments compared to +Fe alone (6). In addition to increased ambient B_{12}, associated bacteria in a diatom's phycosphere or inside *Phaeocystis* colonies may provide a direct flux of B_{12}, as observed in lab studies (77).

More generally, our results refute recent claims that most haptophytes are B_{12}-auxotrophs (31). The detection of the MetE-fusion proteins and transcripts in existing marine environmental datasets implies that this protein is both abundant and actively used by natural phytoplankton populations (*SI Appendix*, Fig. S12 and *Discussion*). The sensitivity of this MetE-fusion to B_{12} availability makes it a potentially useful biomarker for vitamin-stress in a range of phytoplankton species. Using a combination of axenic culture experiments under B_{12}-scarcity paired with multi-omic analyses, we detected the MetE-fusion protein. In retrospect, previous assessments of marine microbial B_{12} requirements—and especially key eukaryotic phytoplankton groups such as the haptophytes and other protists described herein—were stymied by a combination of factors, including 1) a lack of a clear phylogenetic pattern to B_{12}-auxotrophy in plankton and a propensity for strain-level differences (15, 30); 2) bioinformatic evaluation for B_{12} auxotrophy utilizing experimental transcriptomes of cultures (due to lack of available haptophyte genomes) grown in B_{12}-replete media or with possible bacterial contaminants that are not expressing a B_{12} response; and 3) *metE* gene loss due to B_{12} availability, as supported by a lack of taxonomic pattern for *metE* presence (66), *metE* pseudogenes in *V. carteri* and *G. pectorale* (14), and experimental evolution experiments demonstrating the loss of *metE* when grown in consistent B_{12} supply (16).

The detection of MetE-fusion genes in diverse marine protists implies the extent of B_{12}-auxotrophy has been overestimated and is in need of revision. Moreover, it implies an importance of B_{12} in microbial ecology beyond microbes that are auxotrophic for B_{12}. Previously, organisms that are purely auxotrophic (only containing MetH) were considered most likely to be impacted by B_{12} scarcity. However, the results of this study imply that a much broader swath of marine microbes may regulate their proteome in response to environmental B_{12}. Because MetH is so much more efficient than MetE, there appear to be ecological niches for the use of both isoforms, including in regions when Zn-sparing may also be more important (19). While organisms using MetH may be responding to more replete nutrient conditions, many of those same organisms now appear to also maintain the B_{12}-sparing capability of MetE, inferring a significant ecological advantage in regions or periods of B_{12} scarcity. In this sense, the notion of a simple obligate B_{12} auxotrophic dependency in the oceans may be rarer than previously thought and, instead, a dynamic metabolism that responds to B_{12} may be the rule rather than the exception. Moreover, while B_{12} has been shown to be an important micronutrient in the ecology of coastal Antarctic regions, the broad taxonomic and geographic distribution of MetE-fusions implies that B_{12} nutrition may be an important ecological factor throughout the coastal oceans.

Materials and Methods

Detailed methods are provided in *SI Appendix*. This includes details of the algal cultures, Fe-B_{12} limitation experiment, RNA and protein extraction, de novo transcriptomic assembly, protein mass spectrometry, peptide and protein identification, data processing, statistical and differential expression analysis, metatranscriptomic and metaproteomic searches, and phylogenetic analysis. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (*Data Availability*).

Data, Materials, and Software Availability. The *P. antarctica* CCMP1871 mass spectrometry culture proteomics data and transcriptome-derived FASTA files have been deposited with ProteomeXchange consortium through PRoteomics IDEntifications Database (PRIDE) repository under project accession number PXD031524 (https://doi.org/10.6019/PXD31524) (78, 79). Code associated with Non-metric MultiDimensional Scaling (NMDS), Permutational multivariate analysis of variance (PERMANOVA), and Analysis of Similarities (ANOSIM) on transcriptomic data and interactive figures of the significantly differentially abundant protein analyses are included in the relevant figure captions, supporting information, and the authors' GitHub repository at https://github.com/maksaito/Phaeo-Fe-B12 (80).

ACKNOWLEDGMENTS. We would like to acknowledge S. Dutkiewicz and A. Worden for helpful comments on the proteomics analysis, and R. Banerjee for useful conversations regarding the related functions of the conserved domains. This work was supported by research grants NSF-ANT 1643684, NSF-OCE 2123055, Center for Chemical Currencies of a Microbial Planet (C-COMP) NSF-STC 2019589, the Gordon and Betty Moore Foundation Grant 3782, and Simons Foundation Grant 1038971 to M.A.S.; Grants NSF-ANT-1043671, NSF-OCE-1756884, Gordon and Betty Moore Foundation Grant 3828, and Simons Foundation Grant 970820 to A.E.A.; and Simons Collaboration on Ocean Processes and Ecology Award 329108 and Simons Collaboration on Computational Biogeochemical Modeling of Marine Ecosystems Award 549931 to M.J.F. This is C-COMP publication number 036. Portions of the paper were developed from the thesis of D.R.

Author affiliations: [a]Earth Atmospheric Planetary Sciences Department, Massachusetts Institute of Technology, Cambridge, MA 02139; [b]Marine Chemistry and Geochemistry Department, Woods Hole, MA 02543; [c]Microbial and Environmental Genomics Department, J.C. Venter Institute, La Jolla, CA 92037; and [d]Integrative Oceanography Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92037

Author contributions: D.R. and M.A.S. designed research; D.R., M.R.M., D.M.M., and A.E.A. performed research; M.R.M. contributed new reagents/analytic tools; D.R., Z.F., M.M.B., M.R.M., A.E.A., and M.A.S. analyzed data; and D.R., Z.F., M.M.B., M.R.M., D.M.M., M.J.F., and M.A.S. wrote the paper.

The authors declare no competing interest.

1. C. Moore *et al*., Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* **6**, 701–710 (2013).
2. M. T. Croft, A. D. Lawrence, E. Raux-Deery, M. J. Warren, A. G. Smith, Algae acquire vitamin B_{12} through a symbiotic relationship with bacteria. *Nature* **438**, 90–93 (2005).
3. D. A. Rodionov, A. G. Vitreschak, A. A. Mironov, M. S. Gelfand, Comparative genomics of the vitamin B_{12} metabolism and regulation in prokaryotes. *J. Biol. Chem.* **278**, 41148–41159 (2003).
4. M. Droop, A pelagic marine diatom requiring cobalamin. *J. Marine Biol. Assoc. U.K.* **34**, 229–231 (1955).
5. A. Carlucci, S. Silbernagel, P. McNally, Influence of temperature and solar radiation on persistence of vitamin B_{12}, thiamine, and biotin in seawater 1. *J. Phycol.* **5**, 302–305 (1969).
6. E. M. Bertrand *et al*., Vitamin B_{12} and iron colimitation of phytoplankton growth in the Ross Sea. *Limnol. Oceanogr.* **52**, 1079–1093 (2007).
7. S. A. Sañudo-Wilhelmy *et al*., Multiple B-vitamin depletion in large areas of the coastal ocean. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14041–14045 (2012).
8. S. Sañudo-Wilhelmy, C. Gobler, M. Okbamichael, G. Taylor, Regulation of phytoplankton dynamics by vitamin B_{12}. *Geophys. Res. Lett.* **33** (2006).
9. C. J. Gobler, C. Norman, C. Panzeca, G. T. Taylor, S. A. Sañudo-Wilhelmy, Effect of B-vitamins (B1, B_{12}) and inorganic nutrients on algal bloom dynamics in a coastal ecosystem. *Aquat. Microbial. Ecol.* **49**, 181–194 (2007).
10. F. Koch *et al*., The effect of vitamin B_{12} on phytoplankton growth and community structure in the Gulf of Alaska. *Limnol. Oceanogr.* **56**, 1023–1034 (2011).
11. C. Panzeca *et al*., B vitamins as regulators of phytoplankton dynamics. *Eos Trans. Am. Geophys. Union* **87**, 593–596 (2006).
12. E. M. Bertrand *et al*., Phytoplankton-bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9938–9943 (2015).
13. M. T. Croft, M. J. Warren, A. G. Smith, Algae need their vitamins. *Eukaryotic Cell* **5**, 1175–1183 (2006).
14. K. E. Helliwell, G. L. Wheeler, K. C. Leptos, R. E. Goldstein, A. G. Smith, Insights into the evolution of vitamin B_{12} auxotrophy from sequenced algal genomes. *Mol. Biol. Evol.* **28**, 2921–2933 (2011).
15. K. E. Helliwell, G. L. Wheeler, A. G. Smith, Widespread decay of vitamin-related pathways: Coincidence or consequence? *Trends. Genet.* **29**, 469–478 (2013).
16. K. E. Helliwell *et al*., Fundamental shift in vitamin B12 eco-physiology of a model alga demonstrated by experimental evolution. *ISME J.* **9**, 1446–1455 (2015).
17. J. C. González, K. Pearlso, J. E. Penner-Hahn, R. G. Matthews, Cobalamin-independent methionine synthase from *Escherichia coli*: A zinc metalloenzyme. *Biochemistry* **35**, 12228–12234 (1996).
18. C. W. Goulding, R. G. Matthews, Cobalamin-dependent methionine synthase from *Escherichia coli*: Involvement of zinc in homocysteine activation. *Biochemistry* **36**, 15749–15757 (1997).
19. E. M. Bertrand *et al*., Methionine synthase interreplacement in diatom cultures and communities: Implications for the persistence of B12 use by eukaryotic phytoplankton. *Limnol. Oceanogr.* **58**, 1431–1450 (2013).
20. P. N. Sedwick *et al*., Early season depletion of dissolved iron in the ross sea polynya: Implications for iron dynamics on the Antarctic continental shelf. *J. Geophys. Res.: Oceans* **116**, C12019 (2011).
21. E. M. Bertrand *et al*., Iron limitation of a springtime bacterial and phytoplankton community in the Ross Sea: Implications for vitamin B_{12} nutrition. *Front. Microbiol.* **2**, 160 (2011).
22. F. Koch *et al*., Vitamin B1 and B_{12} uptake and cycling by plankton communities in coastal ecosystems. *Front. Microbiol.* **3**, 363 (2012).
23. K. A. Ellis, N. R. Cohen, C. Moreno, A. Marchetti, Cobalamin-independent methionine synthase distribution and influence on vitamin B_{12} growth requirements in marine diatoms. *Protist* **168**, 32–47 (2017).
24. S. J. Bender *et al*., Colony formation in *Phaeocystis antarctica*: Connecting molecular mechanisms with iron biogeochemistry. *Biogeosciences* **15**, 4923–4942 (2018).
25. K. R. Arrigo *et al*., Phytoplankton community structure and the drawdown of nutrients and CO_2 in the Southern Ocean. *Science* **283**, 365–367 (1999).
26. G. DiTullio *et al*., Rapid and early export of *Phaeocystis antarctica* blooms in the Ross Sea, Antarctica. *Nature* **404**, 595–598 (2000).
27. W. O. Smith Jr, M. R. Dennett, S. Mathot, D. A. Caron, The temporal dynamics of the flagellated and colonial stages of *Phaeocystis antarctica* in the Ross Sea. *Deep Sea Res Part II: Top. Stud. Oceanogr.* **50**, 605–617 (2003).
28. A. E. Koid *et al*., Comparative transcriptome analysis of four prymnesiophyte algae. *PLoS One* **9**, e97801 (2014).
29. F. Bolinesi *et al*., On the relationship between a novel prorocentrum sp. and colonial *Phaeocystis antarctica* under iron and vitamin B12 limitation: Ecological implications for Antarctic waters. *Appl. Sci.* **10**, 6965 (2020).
30. Y. Z. Tang, F. Koch, C. J. Gobler, Most harmful algal bloom species are vitamin B1 and B_{12} auxotrophs. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20756–20761 (2010).
31. C. Nef *et al*., How haptophytes microalgae mitigate vitamin B_{12} limitation. *Sci. Rep.* **9**, 1–11 (2019).
32. L. Provasoli, J. McLaughlin, M. Droop, The development of artificial media for marine algae. *Archiv. Mikrobiol.* **25**, 392–428 (1957).
33. J. C. Lewin, R. R. Guillard, Diatoms. *Annu. Rev. Microbiol.* **17**, 373–414 (1963).
34. R. R. Guillard, B12 specificity of marine centric diatoms 1, 2. *J. Phycol.* **4**, 59–64 (1968).
35. A. Carlucci, P. M. Bowes, Production of vitamin B12, thiamine, and biotin by phytoplankton 1. *J. Phycol.* **6**, 351–357 (1970).
36. K. C. Haines, R. R. Guillard, Growth of vitamin B12-requiring marine diatoms in mixed laboratory cultures with vitamin B12-producing marine bacteria. *J. Phycol.* **10**, 245–252 (1974).
37. L. Peperzak, W. Gieskes, R. Duin, F. Colijn, The vitamin B requirement of *Phaeocystis globosa* (Prymnesiophyceae). *J. Plankton Res.* **22**, 1529–1537 (2000).
38. H. Planquette, R. M. Sherrell, S. Stammerjohn, M. P. Field, Particulate iron delivery to the water column of the Amundsen sea, Antarctica. *Marine Chem.* **153**, 15–30 (2013).
39. R. Sherrell, M. Lagerström, K. Forsch, S. Stammerjohn, P. Yager, Dynamics of dissolved iron and other bioactive trace metals (Mn, Ni, Cu, Zn) in the Amundsen Sea Polynya, Antarctica. *Elementa* **3**, 000071 (2015).
40. M. Van Manen *et al*., The role of the Dotson ice shelf and circumpolar deep water as driver and source of dissolved and particulate iron and manganese in the Amundsen Sea Polynya, Southern Ocean. *Marine Chem.* **246**, 104161 (2022).
41. N. G. Walworth *et al*., Mechanisms of increased trichodesmium fitness under iron and phosphorus co-limitation in the present and future ocean. *Nat. Commun.* **7**, 1–11 (2016).
42. J. Oksanen *et al*., Community ecology package. R. Package Version 2:5-2 (2018).
43. M. J. Anderson, "Permutational multivariate analysis of variance (PERMANOVA)" in *Wiley StatsRef: statistics reference online*, M. Davidian et al., Eds. (Wiley, 2014), pp. 1–15.
44. M. Mars Brisbin, S. Mitarai, Differential gene expression supports a resource-intensive, defensive role for colony production in the bloom-forming haptophyte, *Phaeocystis globosa*. *J. Eukaryot. Microbiol.* **66**, 788–801 (2019).
45. N. Pavelka *et al*., A power law global error model for the identification of differentially expressed genes in microarray data. *BMC Bioinf.* **5**, 203 (2004).
46. N. Pavelka *et al*., Statistical similarities between transcriptomics and quantitative shotgun proteomics data. *Mol. Cell. Proteom.* **7**, 631–644 (2008).
47. E. M. Bertrand, A. E. Allen, Influence of vitamin B auxotrophy on nitrogen metabolism in eukaryotic phytoplankton. *Front. Microbiol.* **3**, 375 (2012).
48. N. R. Cohen *et al*., Iron and vitamin interactions in marine diatom isolates and natural assemblages of the northeast pacific ocean. *Limnol. Oceanogr.* **62**, 2076–2096 (2017).
49. H. Ploug, W. Stolte, B. B. Jørgensen, Diffusive boundary layers of the colony-forming plankton alga *Phaeocystis* sp.-implications for nutrient uptake and cellular growth. *Limnol. Oceanogr.* **44**, 1959–1967 (1999).
50. M. F. Romine *et al*., Elucidation of roles for vitamin B_{12} in regulation of folate, ubiquinone, and methionine metabolism. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E1205–E1214 (2017).
51. M. J. Schneider, M. Ulland, R. D. Sloboda, A protein methylation pathway in chlamydomonas flagella is active during flagellar resorption. *Mol. Biol. Cell* **19**, 4319–4327 (2008).
52. R. D. Sloboda, L. Howard, Protein methylation in full length chlamydomonas flagella. *Cell Motil. Cytoskeleton* **66**, 650–660 (2009).
53. N. R. Movva, K. Nakamura, M. Inouye, Gene structure of the OmPa protein, a major surface protein of *Escherichia coli* required for cell-cell interaction. *J. Mol. Biol.* **143**, 317–328 (1980).
54. D. Heldt *et al*., Aerobic synthesis of vitamin B12: Ring contraction and cobalt chelation. *Biochem. Soc. Trans.* **33**, 815–819 (2005).
55. S. Y. Tsuji, N. Wu, C. Khosla, Intermodular communication in polyketide synthases: Comparing the role of protein-protein interactions to those in other multidomain proteins. *Biochemistry* **40**, 2317–2325 (2001).
56. R. Pejchal, M. L. Ludwig, Cobalamin-independent methionine synthase (MetE): A face-to-face double barrel that evolved by gene duplication. *PLoS Biol.* **3**, e31 (2004).
57. D. Deobald, R. Hanna, S. Shahryari, G. Layer, L. Adrian, Identification and characterization of a bacterial core methionine synthase. *Sci. Rep.* **10**, 1–13 (2020).
58. M. N. Price, A. M. Deutschbauer, A. P. Arkin, Four families of folate-independent methionine synthases. *PLoS Genet.* **17**, e1009342 (2021).
59. R. W. Wheatley, K. K. Ng, M. Kapoor, Fungal cobalamin-independent methionine synthase: Insights from the model organism, *Neurospora crassa*. *Archiv. Biochem. Biophys.* **590**, 125–137 (2016).
60. J. C. Gonzalez, R. V. Banerjee, S. Huang, J. S. Sumner, R. G. Matthews, Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from *Escherichia coli*: Two solutions to the same chemical problem. *Biochemistry* **31**, 6045–6056 (1992).
61. M. Koutmos *et al*., Metal active site elasticity linked to activation of homocysteine in methionine synthases. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3286–3291 (2008).
62. R. G. Matthews, C. W. Goulding, Enzyme-catalyzed methyl transfers to thiols: The role of zinc. *Curr. Opin. Chem. Biol.* **1**, 332–339 (1997).
63. R. Caspi *et al*., The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **38**, D473–D479 (2010).
64. T. G. Stephens *et al*., Genomes of the dinoflagellate *Polarella glacialis* encode tandemly repeated single-exon genes with adaptive functions. *BMC Biol.* **18**, 1–21 (2020).
65. M. Wu *et al*., Manganese and iron deficiency in southern ocean *Phaeocystis antarctica* populations revealed through taxon-specific protein indicators. *Nat. Commun.* **10**, 1–10 (2019).
66. S. J. Giovannoni, Vitamins in the sea. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 13888–13889 (2012).
67. T. Gröne, G. Kirst, The effect of nitrogen deficiency, methionine and inhibitors of methionine metabolism on the DMSP contents of *Tetraselmis subcordiformis* (stein). *Marine Biol.* **112**, 497–503 (1992).
68. C. Nef *et al*., Cobalamin scarcity modifies carbon allocation and impairs DMSP production through methionine metabolism in the haptophyte microalgae *Tisochrysis lutea*. *Front. Marine Sci.* **7** (2020).
69. Q. Carradec *et al*., A global ocean atlas of eukaryotic genes. *Nat. Commun.* **9**, 1–13 (2018).
70. J. Decelle *et al*., An original mode of symbiosis in open ocean plankton. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18000–18005 (2012).
71. R. J. Gast, D. M. Moran, M. R. Dennett, D. A. Caron, Kleptoplasty in an Antarctic dinoflagellate: Caught in evolutionary transition? *Environ. Microbiol.* **9**, 39–45 (2007).
72. O. Mangoni *et al*., *Phaeocystis antarctica* unusual summer bloom in stratified Antarctic coastal waters (Terra Nova Bay, Ross Sea). *Marine Environ. Res.* **151**, 104733 (2019).
73. R. F. Strzepek, M. T. Maldonado, K. A. Hunter, R. D. Frew, P. W. Boyd, Adaptive strategies by southern ocean phytoplankton to lessen iron limitation: Uptake of organically complexed iron and reduced cellular iron requirements. *Limnol. Oceanogr.* **56**, 1983–2002 (2011).
74. R. F. Strzepek, K. A. Hunter, R. D. Frew, P. J. Harrison, P. W. Boyd, Iron-light interactions differ in southern ocean phytoplankton. *Limnol. Oceanogr.* **57**, 1182–1200 (2012).
75. H. Ducklow *et al*., The seasonal development of the bacterioplankton bloom in the Ross Sea, Antarctica, 1994–1997. *Deep Sea Res. Part II: Top. Stud. Oceanogr.* **48**, 4199–4221 (2001).
76. A. E. Noble, D. M. Moran, A. E. Allen, M. A. Saito, Dissolved and particulate trace metal micronutrients under the McMurdo sound seasonal sea ice: Basal sea ice communities as a capacitor for iron. *Front. Chem.* **1**, 25 (2013).
77. B. P. Durham *et al*., Cryptic carbon and sulfur cycling between surface ocean plankton. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 453–457 (2015).
78. E. W. Deutsch *et al*., The proteomeXchange consortium in 2020: Enabling "big data" approaches in proteomics. *Nucleic Acids Res.* **48**, D1145–D1152 (2020).
79. Y. Perez-Riverol *et al*., The pride database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).
80. M. Saito, D. Rao, Code for data analysis and visualizations. Github. https://github.com/maksaito/Phaeo-Fe-B12. Deposited 15 October 2023.