The ISME Journal (2012) 6, 1883–1900

www.nature.com/ismej

ORIGINAL ARTICLE

A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters

Timothy J Williams[1], Emilie Long[1,2], Flavia Evans[1], Mathew Z DeMaere[1], Federico M Lauro[1], Mark J Raftery[3], Hugh Ducklow[4], Joseph J Grzymski[5], Alison E Murray[5] and Ricardo Cavicchioli[1]

[1]*School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, New South Wales, Australia;* [2]*Département de formation 'Biologie Cellulaire et Moléculaire', UFR 927, Université Pierre et Marie Curie (UPMC) Paris VI, Paris, France;* [3]*Bioanalytical Mass Spectrometry Facility, The University of New South Wales, Sydney, New South Wales, Australia;* [4]*The Ecosystems Center, Marine Biological Laboratory, Woods Hole, MA, USA and* [5]*Desert Research Institute, Division of Earth and Ecosystem Sciences, Reno, NV, USA*

A metaproteomic survey of surface coastal waters near Palmer Station on the Antarctic Peninsula, West Antarctica, was performed, revealing marked differences in the functional capacity of summer and winter communities of bacterioplankton. Proteins from Flavobacteria were more abundant in the summer metaproteome, whereas winter was characterized by proteins from ammonia-oxidizing Marine Group I Crenarchaeota. Proteins prevalent in both seasons were from SAR11 and Rhodobacterales clades of Alphaproteobacteria, as well as many lineages of Gammaproteobacteria. The metaproteome data were used to elucidate the main metabolic and energy generation pathways and transport processes occurring at the microbial level in each season. In summer, autotrophic carbon assimilation appears to be driven by oxygenic photoautotrophy, consistent with high light availability and intensity. In contrast, during the dark polar winter, the metaproteome supported the occurrence of chemolithoautotrophy via the 3-hydroxypropionate/4-hydroxybutyrate cycle and the reverse tricarboxylic acid cycle of ammonia-oxidizing archaea and nitrite-oxidizing bacteria, respectively. Proteins involved in nitrification were also detected in the metaproteome. Taurine appears to be an important source of carbon and nitrogen for heterotrophs (especially SAR11), with transporters and enzymes for taurine uptake and degradation abundant in the metaproteome. Divergent heterotrophic strategies for Alphaproteobacteria and Flavobacteria were indicated by the metaproteome data, with Alphaproteobacteria capturing (by high-affinity transport) and processing labile solutes, and Flavobacteria expressing outer membrane receptors for particle adhesion to facilitate the exploitation of non-labile substrates. TonB-dependent receptors from Gammaproteobacteria and Flavobacteria (particularly in summer) were abundant, indicating that scavenging of substrates was likely an important strategy for these clades of Southern Ocean bacteria. This study provides the first insight into differences in functional processes occurring between summer and winter microbial communities in coastal Antarctic waters, and particularly highlights the important role that 'dark' carbon fixation has in winter.

The ISME Journal (2012) **6,** 1883–1900; doi:10.1038/ismej.2012.28; published online 26 April 2012
Subject Category: integrated genomics and post-genomics approaches in microbial ecology
Keywords: marine microorganisms; metaproteomics; antarctic microbiology; southern ocean microbiology

Introduction

The Southern Ocean encircles the Antarctic continent, and influences the climate of the entire Earth. It is a major sink for global carbon dioxide emissions, accounting for around 30% of global ocean uptake of CO_2 (AASSP, 2011). Global thermohaline circulation is driven by the formation of deep water masses in the Southern Ocean that, in turn, drive global nutrient cycling throughout the oceans. As the southernmost part of the global ocean

Correspondence: R Cavicchioli, School of Biotechnology and Biomolecular Sciences, The University of New South Wales, BABS, Sydney, New South Wales 2052, Australia.
E-mail: r.cavicchioli@unsw.edu.au or
AE Murray, Desert Research Institute, Division of Earth and Ecosystem Sciences, 2215 Raggio Parkway, Reno, NV 89512, USA.
E-mail: alison.murray@dri.edu
Received 23 October 2011; revised 23 February 2012; accepted 24 February 2012; published online 26 April 2012

network, this polar region undergoes extreme seasonal variations in sea-ice cover and illumination, which greatly influences the primary productivity of the indigenous marine microorganisms. During the austral summer, Antarctic surface water is exposed to continuous high solar irradiation causing sea-ice and glacier-ice melt, and the waters are highly productive. In contrast, the prolonged cold and darkness during the austral winter permits sea-ice expansion and minimal or no photosynthetic primary production.

Although microorganisms that depend upon oxygenic photosynthesis decline during the winter season, some groups of microorganisms increase in abundance (Murray and Grzymski, 2007). Shifts in microbial community composition between summer and winter have been documented (Murray *et al.*, 1998; Massana *et al.*, 2001; Murray and Grzymski, 2007; Piquet *et al.*, 2011; Ghiglione and Murray, 2012; Grzymski *et al.*, 2012). However, little is known about global gene expression of Southern Ocean microbial communities as relevant studies (for example, metaproteomics, metatranscriptomics and stable isotope metagenomics) have not been performed.

By determining which proteins have been synthesized by microorganisms present at the time of sampling, metaproteomics enables the reconstruction of microbial processes and metabolic pathways that are central to the functioning of the ecosystem. Marine metaproteomic investigations have been performed on the ocean gyre of the Sargasso Sea (Sowell *et al.*, 2009), ocean gyre and coastal sites in the South Atlantic (Morris *et al.*, 2010), and a coastal region near Oregon (Sowell *et al.*, 2011). In Antarctic environments, metaproteomics has been applied to two marine-derived lakes (Ng *et al.*, 2010; Lauro *et al.*, 2011; Yau *et al.*, 2011). These studies have been facilitated by an expanding database of genomic and metagenomic data that provides a framework for the identification of proteins by mass spectrometry (Thomas *et al.*, 2007).

In the current study, we used metaproteomics to elucidate the predominant metabolic processes performed by bacterioplankton present in surface coastal waters of the West Antarctic Peninsula (WAP). This site, in the coastal waters of Anvers Island, is in the study region of the Palmer Long Term Ecological Research (LTER) program, where a >20-year research program has focused on the Antarctic pelagic marine ecosystem. This region is experiencing one of the most rapid warming events on Earth (Turner *et al.*, 2005). Using a customized metagenomic database that included metagenome data for summer and winter communities from the WAP (Grzymski *et al.*, 2012), metagenome sequence from six libraries sequenced from the Southern Ocean off East Antarctica and genome sequences of numerous marine microorganisms, we identified proteins from Antarctic Peninsula summer and winter metaproteome samples collected during the International Polar Year. By examining expressed proteins in the samples, we were able to infer the major trophic pathways and processes and consider the significance of the differences between summer and winter bacterioplankton.

Materials and methods

Sample collection

Seawater samples were collected from coastal surface waters off of Anvers Island, Western Antarctic Peninsula. Samples for metaproteomic analysis (200 l each) were collected near LTER Station B on six occasions in the austral summer of 2008 (23 February, 26 February, 28 February, 1 March and 4 March), and on three dates in austral winter of 2008 (28 July, 12 August and 19 August) at LTER Station B on the first two samples and at the seawater intake for the third sample. The depth of the water column at Station B is ~70 m and ~6 m at the seawater intake. Physiochemical data collected coincident with metaproteome samples were averaged across sample dates in 2008 to provide an environmental context for the conditions between summer and winter. Seawater physical, chemical and biological properties were determined using standard LTER methods (http://pal.lternet.edu/publications/documents/protocols/). ^{3}H-Leucine incorporation rates were performed as described previously (Kirchman *et al.*, 1985). Dissolved organic carbon (DOC) was measured with a Shimadzu TOC-V high temperature combustion total carbon analyzer as described previously (Sharp *et al.*, 1993; Farmer and Hansell, 2007). Seawater was processed by prefiltering through sequential 5.0- and 2.5-μm filters (Polygard, Millipore, Billerica, MA, USA) and cells concentrated using a tangential flow filtration system and the cells harvested on 25-mm 0.2-μm Supor filters (Millipore). Samples were immersed in a protease inhibitor solution (Ng *et al.*, 2010), which included 2.5 mM EGTA, 2.5 mM EDTA, 0.1 mM Tris-EDTA (pH 8), 1 mM PMSF (freshly prepared) and 50 μm protease inhibitor cocktail VI (Calbiochem, San Diego, CA, USA), and the sample tubes placed into liquid N_2 and stored at −80 °C until extraction.

Metaproteomics

Protein extraction from filter membranes followed the method of Ng *et al.* (2010), using 20-ml lysis buffer containing 10 mM Tris-EDTA (pH 8.0; Univar, Sydney, Australia), 20 μl of protease inhibitor cocktail VI (Calbiochem), 0.1% sodium dodecyl sulfate (Univar) and 1 mM dithiothreitol (Sigma-Aldrich, Sydney, Australia). The only departure from the method was that five sonications were carried out at the following settings: 40 s intervals, 0.5 s pulse on/0.5 s pulse off, 20% amplitude and that 3-kDa rather than 5-kDa Amicon filter units were used.

One-dimensional SDS–PAGE and in-gel trypsin digestion also followed Ng *et al.* (2010), except gels were stained with Coomassie blue. Digested peptides were separated by nano-LC using an Ultimate 3000 HPLC and autosampler system (Dionex, Amsterdam, Netherlands). Samples (2.5 μl) were concentrated and de-salted onto a micro C18 precolumn (500 μm × 2 mm, Michrom Bioresources, Auburn, CA, USA) with H_2O:CH_3CN (98:2, 0.1% TFA) at 15 μl min^{-1}. After a 4-min wash the precolumn was switched (Valco 10 port valve, Dionex) into line with a fritless nano column (75 μ × ~10 cm) containing C18 media (5 μ, 200 Å Magic, Michrom Bioresources). Peptides were eluted using a linear gradient of H_2O:CH_3CN (98:2, 0.1% formic acid) to H_2O:CH_3CN (64:36, 0.1% formic acid) at 250 nl min^{-1} over 30 min. High voltage (2000 V) was applied to low volume tee (Upchurch Scientific, Oak Harbor, WA, USA) and the column tip positioned ~0.5 cm from the heated capillary ($T = 280\,^{\circ}C$) of an Orbitrap Velos (Thermo Electron, Bremen, Germany) mass spectrometer. Positive ions were generated by electrospray and the Orbitrap operated in data-dependent acquisition mode. A survey scan *m/z* 350–1750 was acquired in the Orbitrap (Resolution = 30 000 at *m/z* 400, with an accumulation target value of 1 000 000 ions) with lockmass enabled. Up to the 10 most abundant ions (>5000 counts) with charge states > +2 were sequentially isolated and fragmented within the linear ion trap using collisionally induced dissociation with an activation $q = 0.25$ and activation time of 30 ms at a target value of 30 000 ions. *m/z* ratios selected for MS/MS were dynamically excluded for 30 s. Peak lists were generated using Mascot Daemon/extract_msn (Matrix Science, Thermo, London, UK) using the default parameters, and submitted to Mascot 2.1 (Matrix Science). An Antarctic metaproteome sequence database, AntComb, was constructed from: fosmid libraries (IMG Acc: 2008193000, 2008193001, 2012990003, 2040502005 and 2040502004), individual marine microbial genomes (RefSeq Project ID: 202, 58903, 54247, 58401, 59427, 57855, 54575, 54169, 51877, 54583, 54265, 54403, 54255, 54577, 54163, 54227, 54185, 58403, 54207, 54205, 54623, 54467, 58597, 54183, 57863, 52598, 54259, 58183, 13044 and 51609) and Antarctic (Southern Ocean off East Antarctica) metagenome samples (SRA Acc: SRX024734, SRX024735, SRX024799, SRX025108, SRX024736 and SRX024800). Gene predictions for fosmid data were taken from IMG. Fosmid end, or fully sequenced fosmids, were annotated through the IMG, which incorporates a full annotation system including BlastP to the RefSeq database (Markowitz *et al.*, 2012). Predictions for Celera WGS assembled metagenome data were produced in-house with SHAP (DeMaere *et al.*, 2011). The combined set of translated gene predictions were reduced to a non-redundant set with cdhit-est using a global sequence identity cutoff of 100%. MS/MS data analysis and validation of protein identifications were performed as previously described (Ng *et al.*, 2010) except that the databases used were NR and AntComb. Mascot searches with a false discovery rate >5% were rejected. All other Mascot results were combined and validated using a MudPit analysis in Scaffold 3.0 (Proteome Software Inc., Portland, OR, USA) using the following parameters: minimal probability of peptide identification, 95%; minimal probability of protein identification, 99%. Protein matches were only accepted if they were identified by a minimum of two unique peptides. All proteins were manually annotated with the aid of BLASTP (Allen *et al.*, 2009), and the protein hit that showed the highest sequence identity was recorded, including the organism name (Supplementary Table S1). To determine if individual phylogenetic groups and inferred metabolic functions exhibited seasonal variation, spectral counts were analyzed statistically by using the resampling method implemented in the script ALL_scrambler.pl (Allen *et al.*, 2009). Briefly, a subset ($n = 20\,000$) of the spectra from summer and winter was selected at random, with replacement, classified functionally or phylogenetically, and compared. This process was repeated 10 000 times and the median difference between groups was recorded. The statistically significant differences were then assessed at a 99% confidence level as described previously (Rodriguez-Brito *et al.*, 2006).

Results and discussion

Physicochemical data for the summer and winter samples

Chlorophyll *a* values were 1.2–3.8 μg l^{-1} for summer and 0.037–0.171 μg l^{-1} for winter. DOC values for summer and winter samples were 48.8 and 41.8 μmol l^{-1}, respectively. Bacterial production in summer was much higher as indicated by leucine incorporation rates: average 12.7 ± 1.6 vs 2 ± 0.4 pmol leucine h^{-1}. Average water temperature was 1.4 °C and −0.92 °C in summer and winter, respectively. Nutrient levels averaged 1.53 ± 0.04 μM PO_4, 55.22 ± 6.15 μM $Si(OH)_4$ and 30.45 0.58 μM $NO_2^- + NO_3^-$ for the last three summer samples (29 Febuary–6 March) and 2.03 ± 0.04 μM PO_4, 51.75 ± 7.40 μM $Si(OH)_4$ and 30.56 ± 0.16 μM $NO_2^- + NO_3^-$ over the three winter sample dates.

Phylogenetic diversity represented in the Antarctic metaproteome

A total of 1061 proteins were detected (310 unique to the austral summer, 349 unique to the austral winter and 402 that were detected in both data sets); the full list of proteins is in Supplementary Table S1. Of the 1061 identified proteins, 755 proteins were from the AntComb database, 323 were from NR and 17 were common to both. This suggests that protein sequence data were far better represented in fosmids and metagenomic data than in

microbial genomes contained in NR. Of the 1061 proteins, these included 831 proteins with highest sequence identity to bacterial proteins, 189 archaeal proteins consisting mainly of Marine Group I Crenarchaeota (MGI; 188), 32 eukaryotic proteins and seven phage proteins. Within the bacterial subset, the majority had the best match (highest sequence identity) to proteins from members of the Alphaproteobacteria (404), followed by Gammaproteobacteria (322), Bacteroidetes (56), Betaproteobacteria (18), Nitrospirae (4) and Cyanobacteria (4). Matches to the Bacteroidetes clade (= *Cytophaga-Flavobacterium-Bacteroides* or CFB group) were mainly to Flavobacteria (54 proteins). Our data are consistent with the phylogenetic diversity obtained by previous gene-based analyses of coastal Southern Ocean waters, which identified the relative abundance of MGI, Alphaproteobacteria, Gammaproteobacteria and Flavobacteria (Murray and Grzymski, 2007; Ghiglione and Murray, 2012; Grzymski *et al.*, 2012). Cyanobacteria have been reported to be a negligible component of the marine picoplankton in Antarctic waters (Ghiglione and Murray, 2012). Nevertheless, we detected proteins with best matches to cyanobacterial viruses (cyanophages), including capsid protein Gp23 from Myoviruses (Supplementary Table S1). These cyanophage can exceed 200 nm in length (Sullivan *et al.*, 2005) allowing them to be preferentially captured on the filters. Consistent with this, low numbers of genes with matches to *Prochlorococcus spp., Synechococcus spp.* and their phages have been detected from large-scale shotgun metagenomics analyses of pelagic Southern Ocean waters from East Antarctica (Wilkins D *et al.*, unpublished data), but not in the Antarctic Peninsula metagenomes reported by Grzymski *et al.* (2012).

Within the Alphaproteobacteria, nearly half (199) of the matches had highest identity to proteins from the SAR11 clade (including *Candidatus* Pelagibacter ubique); other proteins had the best match to members of Rhodobacterales (129), Sphingomonadales (37), Rhizobiales (21) and uncultivated SAR116 cluster (9). Within Rhodobacterales, proteins affiliated with the *Roseobacter* clade predominated (105). The majority of Sphingomonadales proteins had 100% sequence identity to the marine oligotroph *Sphingopyxis alaskensis*. Within the Gammaproteobacteria, most of the proteins had the best match to members of four groups: Alteromonadales (83), gammaproteobacterial sulfur oxidizer EOSA-1 (GSO–EOSA-1) complex (Walsh *et al.*, 2009; Swan *et al.*, 2011) (72), Oceanospirillales (63) and the Oligotrophic Marine Gammaproteobacteria (OMG) group (Cho and Giovannoni, 2004) (39). Alteromonadales were represented by proteins with 100% sequence identity to *Pseudoalteromonas haloplanktis*, *Moritella* sp. PE36 and bacterium TW-7. Matches to the GSO–EOSA-I complex are indicative of the presence of the uncultivated sister groups SUP05 and/or ARCTIC96BD-19 (see section *GSO–EOSA-1 bacteria ARCTIC96BD-19/SUP05* below). Within Oceanospirillales, the most common matches were to proteins with the highest identity to *Neptuniibacter caesariensis* or *Marinomonas* spp. Proteins from the Bacteroidetes had the best match to proteins from Flavobacteria, including *Polaribacter* spp., *Zunongwangia profunda*, *Psychroflexus torquis*, and the uncultivated marine bacterium MS024-2A.

All the archaeal proteins matched best to members of MGI (detected exclusively in winter), except for two that matched to the Antarctic methanogen *Methanococcoides burtonii*. MGI proteins had the best match to the ammonia-oxidizing archaea (AOA) *Nitrosopumilus maritimus* or *Cenarchaeum symbiosum* (Hallam *et al.*, 2006a, b; Walker *et al.*, 2010). The dominance of proteins for MGI is consistent with the minimal phylogenetic diversity previously observed for the *Archaea* in surface waters of the Southern Ocean compared with the diversity of the *Bacteria* (Massana *et al.*, 2001; Murray and Grzymski, 2007; Grzymski *et al.*, 2012).

Matches to eukaryotic phytoplankton were more abundant in summer with 32 and 8 total matches in summer vs winter, respectively. Spectral counts were also significantly higher for eukaryotic proteins in summer. These proteins included photosystem proteins and ribulose bisphosphate carboxylase/oxygenase (RuBisCO) from algae, consistent with algae being the major drivers of CO_2 fixation in summer (Murray and Grzymski, 2007). Phytoplankton is undoubtedly the dominant agent of light-mediated carbon fixation in summer for waters off the Antarctic Peninsula. Nanoplankton and larger cells would be expected to be captured by the 1.6- or 2.5-μm prefilters that were used for sampling (see Materials and methods). The eukaryotic proteins present on the 0.2-μm filters are likely to derive from organismal turnover, or from sufficiently small unicellular algae such as *Ostreococcus tauri*. Although bacterial RuBisCO was not detected in the metaproteome, a chemolithoautotroph-related form I RuBisCO was present in several fosmids sequenced in the WAP metagenome (Grzymski *et al.*, 2012), and the winter metaproteome did contain a bacterial CbbQ protein (Supplementary Table S1), which may have a role in post-translation activation of RuBisCO (Hayashi *et al.*, 1997).

Phylogenetic differences between summer and winter

The protein assignments illustrated that specific taxa of bacteria and archaea had different seasonal abundances (Figure 1). Spectra assigned to MGI, SAR11 and GSO–EOSA-1 were significantly more abundant in the winter metaproteome, whereas spectra assigned to Bacteroidetes, Sphingomonadales, Oceanospirillales, Alteromonadales and OMG were significantly more abundant in the summer metaproteome. By contrast, the contribution of SAR11 proteins to inferred phylogenetic diversity was

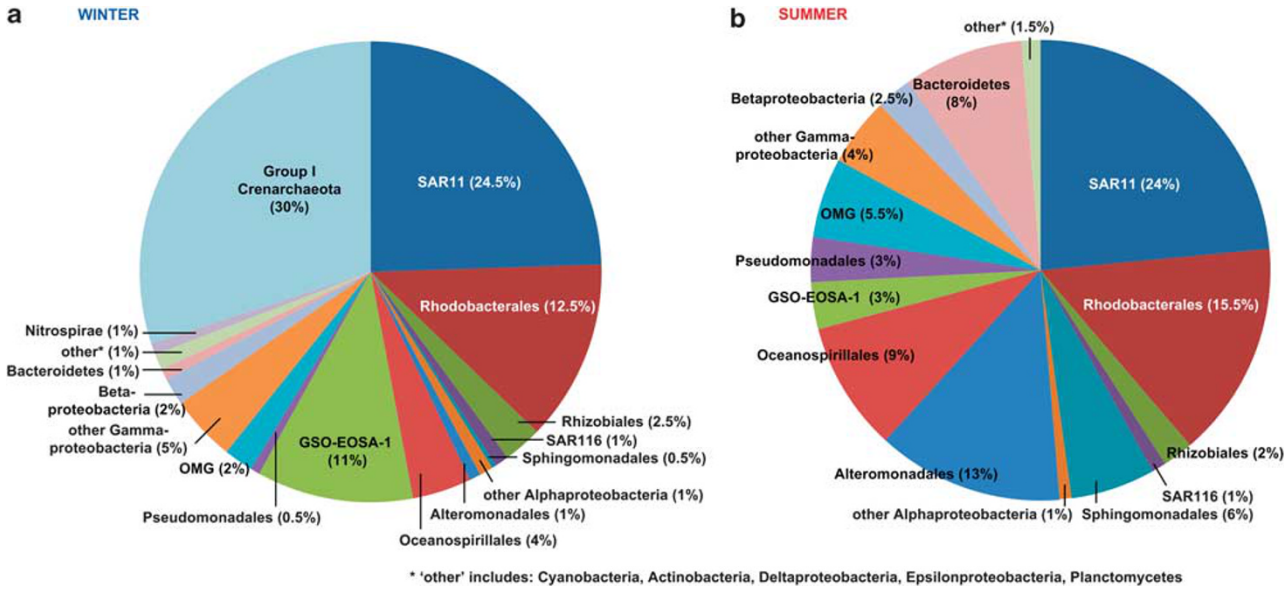

Figure 1 Phylogenetic assignment of bacterial and archaeal proteins detected in the (**a**) winter and (**b**) summer metaproteome. GSO–EOSA-1, gammaproteobacterial sulfur oxidizer EOSA-1.

constant between summer and winter (Figure 1). This suggests that individual SAR11 proteins increased in abundance in winter (leading to higher spectral counts), possibly as a consequence of decreased availability of soluble substrates, which in turn required increased production of solute-binding proteins.

The phylogenetic differences often translated to different metabolic capacities in summer and winter, especially for carbon assimilation (Figure 2). Bacteroidetes proteins were present mainly in summer (Tables 1 and 2 and Supplementary Table S1), consistent with previous studies identifying Flavobacteria as particularly abundant during periods of high primary production (Abell and Bowman, 2005; Murray and Grzymski, 2007). Flavobacteria directly associate with algae (including diatoms) (Grossart *et al.*, 2005), and growth of proteorhodopsin (PR)-containing Flavobacteria can be promoted by light (Gómez-Consarnau *et al.*, 2007). Proteins from *S. alaskensis*, *P. haloplanktis* and Alteromonadales strains TW-7 and *Moritella* sp. PE36 were all unique to the summer metaproteome (except for one *S. alaskensis* GroEL protein common to both seasons; Supplementary Table S1), and Rhodobacterales, SAR11 and the GSO–EOSA-1 complex were well represented in both summer and winter, although proteins matching GSO–EOSA-1 predominated in the winter metaproteome (Table 1 and Supplementary Table S1, Figure 1).

Overall, the differences between functional groups of proteins were not as pronounced between summer and winter (Figure 3). Exceptions include TonB-dependent receptors (TBDRs) and other outer membrane proteins (OMPs), which comprised a higher proportion of individual proteins in summer, due predominantly to higher number of Bacteroidetes proteins, and the higher number of hypothetical proteins in winter, which was due largely to the prevalence of MGI proteins, many of which have yet to be characterized.

Proteins from MGI were exclusive to winter, with proteins from AOA (including best matches to *N. maritimus*) making up to 30% of the winter metaproteome (Figure 1). This is consistent with the abundance of MGI being reported to inversely correlate with chlorophyll *a* concentration and comprising a very low percentage of bacterioplankton cells in summer waters (Murray *et al.*, 1998; Church *et al.*, 2003). Proteins with the highest identity to nitrite-oxidizing bacteria were also found only in winter (Table 2 and Supplementary Table S1), although this was not significant, owing to low spectral counts. In the Atlantic and Pacific Oceans, MGI are abundant in the mesopelagic waters below the photic zone (for example, Karner *et al.*, 2001; Schattenhofer *et al.*, 2009; Canfield *et al.*, 2010; Stewart *et al.*, 2012). The prevalence of this group in surface coastal waters of the WAP in winter is likely to be due to wind-driven mixing, which brings deeper waters to the surface and eliminates the summer surface water layer until it forms again the following summer. Nevertheless, the metaproteomic data indicate that MGI are abundant and metabolically active at the surface during winter.

Transport

The most prevalent bacterial and archaeal proteins in the metaproteome were components of ATP-binding cassette (ABC) transporters (145 proteins; 13.7% of the total metaproteome), ribosomal subunits (113; 10.7%) and various classes of protein chaperones (76; 7.1%; see section *Chaperones and*

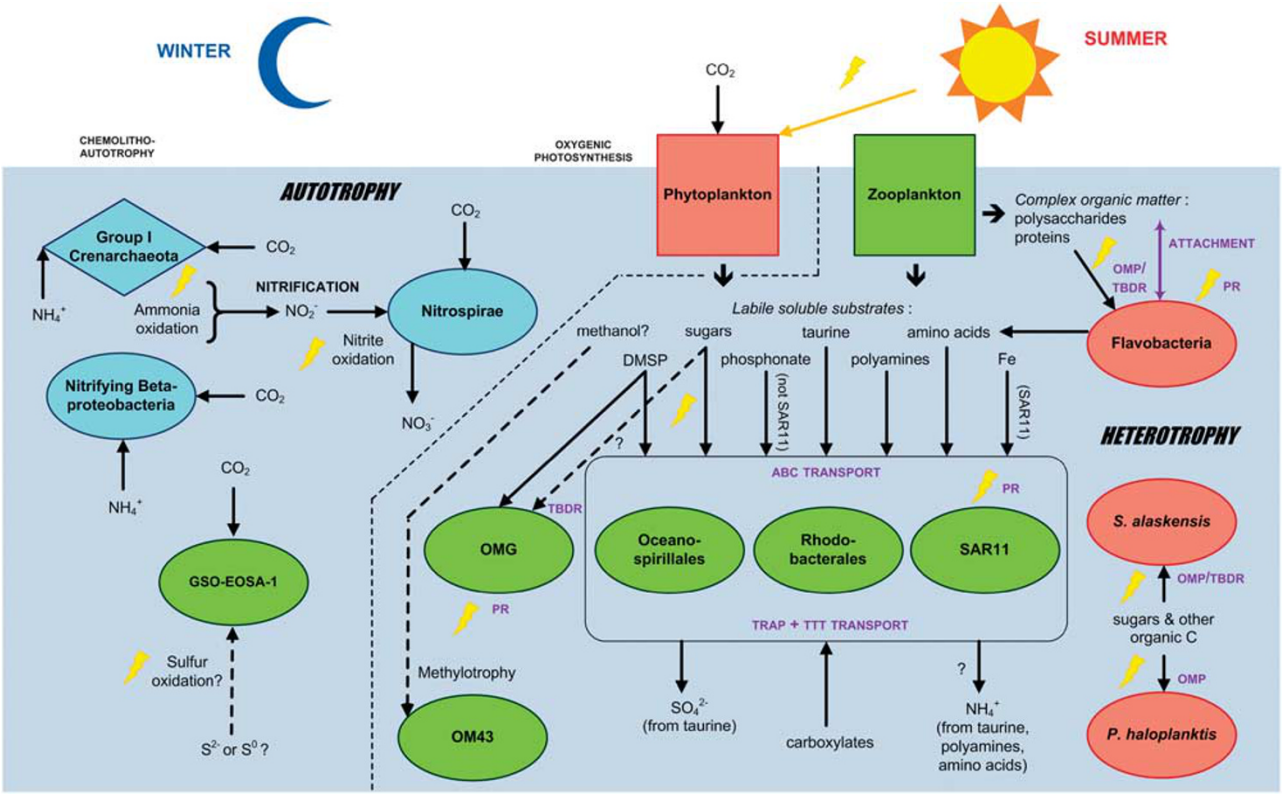

Figure 2 Depiction of the metabolic characteristics of winter and summer communities inferred from the metaproteome. Individual groups were inferred to be present based on best protein matches. Eukaryotic groups are indicated by squares, bacterial clades by ellipses and the sole shown archaeal clade by a diamond. Groups are shaded based on their seasonal distribution: predominantly found in winter (blue) vs summer (red), or abundant in both seasons (green). The microbial processes that appear to be the most important in the cycling of carbon, nitrogen and sulfur, and for energy metabolism are shown. For convenience, dissolved inorganic carbon is shown as CO_2. Pathways or processes that are particularly important for energy generation are signified using a yellow lightning bolt. ABC, ATP-binding cassette; DMSP, dimethylsulfoniopropionate; OMP, outer membrane proteins; PR, proteorhodopsin; TBDR, TonB-dependent receptor; TRAP, tripartite ATP-independent periplasmic transporter; TTT, tripartite tricarboxylate transporter.

quality control below) (Table 1 and Supplementary Table S1). Proteins involved in high-affinity transport were highly represented, similar to findings from metaproteomic studies of the Sargasso Sea (Sowell *et al.*, 2009) and Oregon coast (Sowell *et al.*, 2011).

The majority of ABC transport proteins were periplasmic-binding proteins (PBPs), which were much more abundant (138 of the 145 total) than the membrane-associated ATPase and permease components (Supplementary Table S1). Other high-affinity transporters include tripartite ATP-independent periplasmic (TRAP) transporters (32 proteins, including TAXI) and tripartite tricarboxylate transporters (TTT, 5 proteins), and both types were represented in the metaproteome solely by PBPs (Supplementary Table S1).

The relatively high abundance of PBPs relative to ATPase and/or permease components can be attributed to high expression levels to enhance the frequency of solute capture. Achieving conversion of the imported substrates also requires a relatively minimal number of cytoplasmic enzymes relative to PBPs. Membrane proteins are also more difficult to extract and solubilize from biomass thereby potentially reducing their frequency of detection relative to inherently soluble proteins such as PBPs (Burg *et al.*, 2010; Morris *et al.*, 2010; Williams *et al.*, 2010). However, despite not using methods to specifically enrich for membrane proteins, we detected several classes of OMPs including 73 TBDRs and cell membrane proteins such as PR (3), ammonium transporters (3) and sodium/solute symporters (4; Table 1 and Supplementary Table S1).

ABC, TRAP and TTT transporter proteins were identified in summer and winter samples, and the majority (85%) matched to marine Alphaproteobacteria (Table 1 and Supplementary Table S1). A large number of PBPs of ABC transporters matched to members of SAR11 (54) or *Roseobacter* (43) clades. Based on sequence identity, these were inferred to be specific for polyamines, amino acids, carbohydrates, taurine/sulfonates, oligopeptides, glycerol-3-phosphate, phosphonate or iron. Based on spectral counts, ABC transporter PBPs (for general amino acids, branched-chain amino acids, glycine betaine and sugars) and TRAP transporter PBPs were significantly more abundant in winter than in summer.

The large number of ABC transporter PBPs for amino acids and polyamines (putrescine/spermidine) that appear to derive from members of the

Table 1 Proteins detected in the metaproteome that are discussed in the main text

Protein	*SAR11*	*Roseobacter clade*	*Other Rhodobacterales*	*SAR116*	*Sphingomonadales*	*GSO–EOSA-1 (SUP 05/ARC TIC96 BD-19)*	*OMG*	*Other Gammaproteobacteria*	*Bacteroidetes (CFB)*	*Marine Group I Crenarchaeota*
ABC transport system, general amino-binding protein (YhdW)	8	—	2	1	—	—	—	—	—	—
ABC transport system, glycine betaine-binding protein (OpuAC)	8	—	—	—	—	—	—	—	—	—
ABC transport system, proline/glycine betaine-binding protein (ProX)	3	—	5	—	—	—	—	1	—	—
ABC transport system, branched-chain amino acids-binding protein (LivJ/LivK/BraC)	7	5	2	1	—	—	—	—	—	—
ABC transport system, glutamate/glutamine/aspartate/asparagine-binding protein (BztA)	—	2	2	—	—	—	—	—	—	—
ABC transport system, taurine/sulfonate-binding protein (TauA)	4	1	1	—	—	—	—	—	—	—
ABC transport system, spermidine/putrescine-binding protein (PotD)	9	5	1	1	—	—	—	—	—	—
ABC transport system, sugar-binding protein (MalK)	6?	5	5	1	—	—	—	4	—	—
ABC transport system, oligopeptide-binding protein (OppA)	1	4	5	—	—	—	—	1	—	—
ABC transport system, glycerol-3-PO_4-binding protein (UgpB)	?	1	3	—	—	—	—	—	—	—
ABC transport system, phosphonate-binding protein (PhnD)	—	1	—	—	—	—	—	1	—	—
ABC transport system, iron-binding protein (SfuC)	2	—	—	—	—	—	—	—	—	—
TRAP transporter, binding subunit (DctP)	16	5	1	2	—	—	—	2	—	—
TTT transporter, binding subunit (TctC)	3	2	—	—	—	—	—	—	—	—
Ammonium transporter (Amt)	—	2	—	—	—	—	—	—	—	1
Glutamate dehydrogenase (GdhA)	—	—	—	—	—	—	—	—	—	2
Glutamate synthase, large subunit (GltA)	1	1	—	—	—	—	—	—	—	—
Glutamate synthase, small subunit (GltB)	3	—	1	—	—	—	—	—	—	—
Glutamine synthetase (GlnA/GlnT)	3	—	1	—	—	1	—	1	—	—
Aconitase (AcnA/AcnB)	1	—	—	—	1	—	—	3	—	1
Isocitrate dehydrogenase (Idh)	—	—	—	—	—	—	—	—	—	1
Succinyl-CoA synthetase, beta subunit (SucC)	—	—	—	—	1	—	—	2	—	1
Succinate dehydrogenase (SdhA)	—	—	—	—	—	—	—	—	—	1
Fumarase (FumC)	—	—	—	—	—	—	—	—	1	—
Malate dehydrogenase (Mdh)	1	—	—	—	—	—	—	4	—	—
Isocitrate lyase (AceA)	1	—	1	—	—	1	—	3	—	—
Malate synthase (AceB)	—	—	—	—	—	—	—	1	—	—
Phosphoenolpyruvate carboxykinase (PckA)	—	—	—	—	—	—	—	2	—	1
Enolase (Eno)	—	—	—	—	—	—	—	1	—	1
Phosphoglycerate mutase (ApgM)	—	—	—	—	—	—	—	—	—	1
Glyceraldehyde-3-PO_4 dehydrogenase (GapA)	—	—		—		1	—	2	—	—
Fructose-1,6-bisphosphatase (GlpX)	—	—	—	—	1		—	—	—	1
TonB-dependent receptor: CirA/BtuB/SusC/RagA, etc	—		1	—	2	—	10	16	38	—
Flagellin	—	—	2	—	1	—	—	—	—	—
Aspartate aminotransferase (AspC)	1	—	—	—	—	1	—	—	—	—

Table 1 (Continued)

Protein	*SAR11*	*Roseobacter clade*	*Other Rhodobacterales*	*SAR116*	*Sphingomonadales*	*GSO–EOSA-1 (SUP 05/ARC TIC96 BD-19)*	*OMG*	*Other Gammaproteobacteria*	*Bacteroidetes (CFB)*	*Marine Group I Crenarchaeota*
Serine–pyruvate aminotransferase (Agxt)	—	—	—	—	—	—	—	—	—	1
Taurine–pyruvate aminotransferase (Tpa)	1	1	—	—	—	—	—	—	—	—
Aspartokinase (YclM)	—	—	—	—	1	—	—	—	—	1
Phosphoglycerate dehydrogenase (SerA)	—	—	—	—	1	—	—	—	—	2
Serine hydroxymethyltransferase (GlyA)	—	—	—	—	1	—	—	—	—	—
Glycine cleavage system, L-protein (Lpd)	—	—	—	—	—	—	—	1	—	—
DMSP demethylase (DmdA)	1	—	—	—	—	—	1	—	—	—
3-Methylmercaptopropionyl-CoA ligase (DmdC)	—	—	—	—	—	—	1	1	—	—
Ketol-acid reductoisomerase (IlvC)	1	—	—	—	1	1	—	—	—	1
Dihydroxy-acid dehydrogenase (IlvD)	1	—	—	—	—	—	—	—	—	—
Acetolactate synthase (IlvI)	—	—	—	—	—	1	—	—	—	1
3-Isopropylmalate dehydrogenase (LeuB)	1	—	—	—	—	—	—	—	—	—
Methylmalonate-semialdehyde dehydrogenase (MmsA)	—	—	—	—	—	—	—	2	—	—
3-Hydroxyisobutyrate dehydrogenase (MmsB)	1	—	—	—	—	—	—	—	—	—
Dihydrodipicolinate synthase (DapA)	—	—	—	—	—	1	—	—	—	—
Methionine synthase (MetH)	—	—	—	—	—	—	1	—	—	2
Methionine adenosyltransferase (MetK)	—	—	—	—	—	—	—	—	—	1
O-Acetylhomoserine (thiol)-lyase (MetY)	3	—	—	—	—	—	—	—	—	—
N-Acetyl-gamma-glutamyl-phosphate reductase (ArgC)	—	—	—	—	—	—	—	—	—	2
Chorismate mutase/prephenate dehydrogenase (TyrA)	—	—	—	—	—	1	—	—	—	1
Cyclohexadienyl dehydratase (PheC)	3	—	—	—	—	—	—	—	—	—
Sulfoacetaldehyde acetyltransferase (Xsc)	1	1	1	—	—	—	—	1	—	—
Adenylylsulfate reductase, alpha subunit (AprA)	1	—	—	—	—	2	—	—	—	—
Cysteine desulfurase (IscS)	—	—	—	—	—	—	—	—	—	1
Sulfite reductase, beta subunit (ferredoxin) (CysI)	—	—	—	—	—	—	—	—	—	1
Sarcosine oxidase (SoxA)	1	—	1	—	—	—	1	—	—	—
Dimethylglycine dehydrogenase (Dmg)	1	—	—	—	—	—	—	—	—	—
Formate dehydrogenase, alpha subunit (FdhA)	—	1	1	—	—	—	—	—	—	—
Formate dehydrogenase, beta subunit (FdhB)	2	—	—	—	—	—	—	—	—	—
Formate–tetrahydrofolate ligase (Fhs)	1	1	—	—	—	—	—	—	—	—
Adenosylhomocysteinase (AhcY)	1	—	—	—	2	3	—	1	—	2
Adenylosuccinate synthetase (PurA)	—	—	—	—	—	—	—	—	—	1
Adenylosuccinate lyase (PurB)	1	—	—	—	1	—	—	—	—	1
Proteorhodopsin	2	—	—	—	—	—	1	—	—	—
ATP synthase, subunit alpha, F_0F_1 (AtpA)	1	2	1	—	—	1	1	6	—	—
ATP synthase, subunit beta, F_0F_1 (AtpD)	1	1	1	—	1	1	—	3	—	—
ATP synthase, subunit gamma, F_0F_1 (AtpG)	1	—	—	—	—	1	—	1	—	
ATP synthase, subunit alpha, V-type (AtpA)	—	—	—	—	—	—	—	—	—	4
ATP synthase, subunit beta, V-type (AtpB)	—	—	—	—	—	—	—	—	—	2
Cytochrome c oxidase subunit II (CoxB)	1	—	—	—	—	1	—	—	—	—

Proteins are arranged according to the phylogenetic grouping, with the number of detected orthologs given for each. Boxes are shaded according to the season in which proteins were detected: winter only (blue), summer only (scarlet), and both summer and winter (green). A'?' denotes a questionable assignment.

Table 2 Proteins inferred to belong to carbon fixation and nitrification pathways identified in the winter metaproteome

Carbon fixation and nitrification	
Archaea	N. maritimus *homolog*
Ammonia oxidation	
Ammonium transporter	Nmar_1698 (100%)
Ammonia monooxygenase, subunit B	Nmar_1503 (100%)
Ammonia monooxygenase, subunit B	Nmar_1503 (84%)
3-Hydroxypropionate/4-hydroxybutyrate cycle	
Acetyl-CoA/propionyl-CoA carboxylase (ATP-binding subunit)	Nmar_0273 (86%)
Acetyl-CoA/propionyl-CoA carboxylase (ATP-binding subunit)	Nmar_0273 (74%)
Acetyl-CoA/propionyl-CoA carboxylase (carboxyltransferase subunit)	Nmar_0272 (84%)
Malonic semialdehyde reductase	Nmar_1043 (77%)
Malonic semialdehyde reductase	Nmar_1043 (77%)
3-Hydroxypropionyl-CoA synthetize	Nmar_1309 (84%)
Acryloyl-CoA reductase	Nmar_1622 (81%)
Acryloyl-CoA reductase	Nmar_1622 (80%)
Methylmalonyl-CoA mutase, large subunit	Nmar_0954 (87%)
Succinic semialdehyde reductase	Nmar_1110 (90%)
4-Hydroxybutyryl-CoA synthetase (AMP-forming)	Nmar_0206 (90%)
4-Hydroxybutyryl-CoA dehydratase	Nmar_0207 (97%)
4-Hydroxybutyryl-CoA dehydratase	Nmar_0207 (91%)
Crotonyl-CoA hydratase	Nmar_1308 (100%)
3-Hydroxybutyryl-CoA dehydrogenase (NAD+)	Nmar_1028 (87%)
Acetoacetyl-CoA beta-ketothiolase	Nmar_1631 (88%)
Energy conservation (electron transport chain)	
Multicopper oxidase (NirK-like)	Nmar_1667 (87%)
Blue (Type 1) copper domain protein	Nmar_0185 (72%)
Plastocyanin-family protein	Nmar_0343 (63%)
Bacteria	*Closest homolog*
Ammonia oxidation	
Ammonia monooxygenase, subunit B	AmoB, *Nitrosospira* sp. NpAV (Betaproteobacteria) (100%)
Reverse tricarboxylic acid cycle	
Pyruvate:ferredoxin oxidoreductase, gamma subunit	PorG, *Candidatus* Leptospirillum rubarum (Nitrospirae) (100%)
Nitrite oxidation	
Nitrite oxidoreductase, alpha subunit	NxrA, *Candidatus* Nitrospira defluvii (Nitrospirae) (69%)
Nitrite oxidoreductase, alpha subunit	NarG, *Candidatus* Kuenenia stuttgartiensis (Planctomycetes) (78%)

The closest match of each protein is shown for the ammonia-oxidizing archaeon *N. maritimus*, ammonia-oxidizing Betaproteobacteria, and nitrite-oxidizing Nitrospirae and Planctomycetes. For some proteins, multiple orthologs were identified in the metaproteome. Percentages in parentheses represent percentage amino-acid identity.

Roseobacter clade (Supplementary Table S1) are consistent with genomic studies implicating marine *Roseobacter* spp. as using these organic solutes as sources of carbon and nitrogen (Moran *et al*., 2007). The *Roseobacter* clade is a dominant constituent of bacterioplankton, especially in coastal waters, demonstrating metabolic versatility and a capacity to associate with phytoplankton blooms as a means of deriving metabolites from exudates (Buchan *et al*., 2005; Moran *et al*., 2007). Nevertheless, unlike heterotrophic Bacteroidetes and Gammaproteobacteria groups, based on spectral counts, Rhodobacterales was not significantly more abundant in summer than in winter. ABC transporter proteins were also identified in the winter metaproteome for phosphonate transport, matching to members of the Rhodobacterales and Gammaproteobacteria (Table 1 and Supplementary Table S1).

The ABC transporter PBPs that best matched to SAR11 included those specific for polyamines (spermidine/putrescine), general amino acids, branched-chain amino acids, glycine betaine, taurine/sulfonates, iron and possibly sugars (Table 1 and Supplementary Table S1). The expression of putative sugar transporters accords with the ability of coastal SAR11 isolates to utilize energy-rich carbohydrates (Schwalbach *et al*., 2010). In other metaproteome studies of SAR11, summer populations off the Oregon coast were similarly dominated by transport proteins for carbon- and nitrogen-containing compounds, rather than phosphate (Sowell *et al*., 2011). In contrast, a study of the ocean gyre of the Sargasso Sea identified abundant SAR11 phosphate and phosphonate ABC transporters (Sowell *et al*., 2009).

Unlike ABC transporters, TRAP and TTT transporters utilize a proton motive force rather than ATP hydrolysis to drive solute uptake (Forward *et al*., 1997; Winnen *et al*., 2003). The expression of these indicates that carboxylic acids are likely to be transported and metabolized for growth, predominantly by the *Roseobacter* and SAR11 clades (Moran and Zepp, 1997; Moran *et al*., 2007).

Many of the SAR11 proteins in the metaproteome that are implicated in carbon and nitrogen substrate uptake are the same as those detected for SAR11 in the Sargasso Sea: ABC transporters for polyamines, amino acids, and taurine and TRAP dicarboxylate transporters (Sowell *et al*., 2009). However, the Sargasso Sea SAR11 metaproteome was notable for the absence of iron transport proteins, and this was attributed to high levels of iron in Sargasso Sea summer surface waters due to eolian input from the Saharan desert (Sowell *et al*., 2009). The expression of *sfuC* (iron-binding PBP ABC transport protein), *cspL* (RNA-binding protein implicated in controlling SfuC translation in response to iron scarcity) and *groES* was elevated during iron-limited growth of SAR11, and SfuC has been suggested to be a useful ecological marker for iron limitation (Smith *et al*., 2010). In the current metaproteome, all three proteins (SfuC, CspL and GroES) from SAR11 were detected in the summer (Supplementary Table S1), when iron limitation may be more acute due to more intense competition for iron with phytoplankton

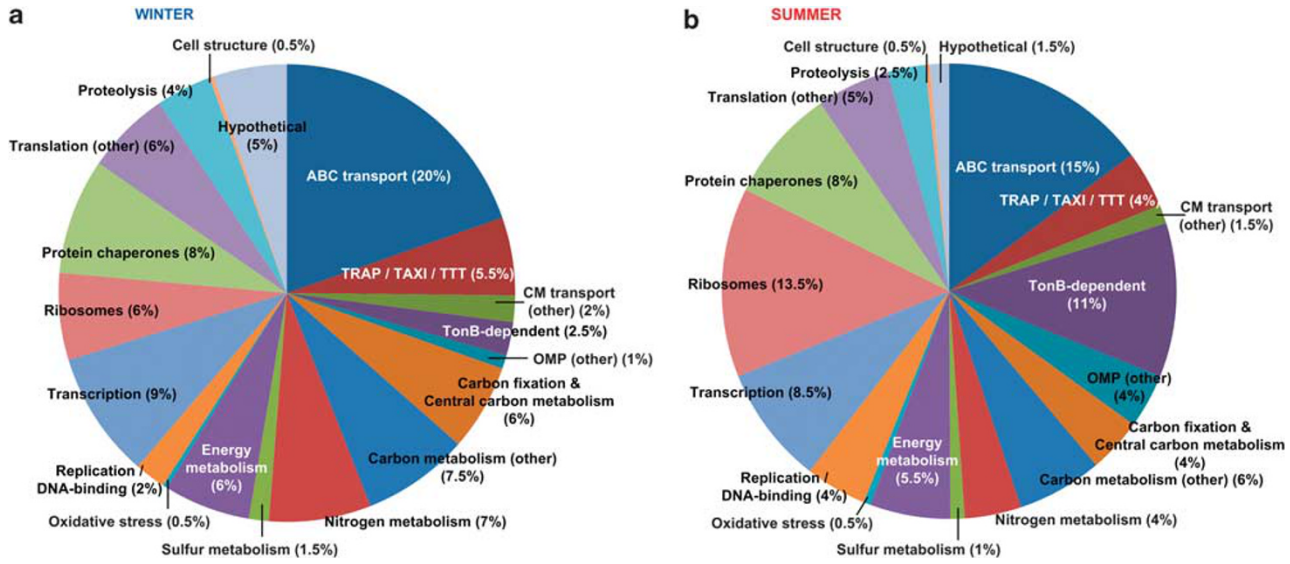

Figure 3 Functional assignment of bacterial and archaeal proteins detected in the (**a**) winter and (**b**) summer metaproteome. CM, cell membrane; OMP, outer membrane protein; TAXI, TRAP-associated extracytoplasmic immunity; TRAP, tripartite ATP-independent periplasmic transporter; TTT, tripartite tricarboxylate transporter.

and other organisms; however, spectral counts revealed no significant difference between seasons.

TBDRs are involved in proton motive force-dependent outer membrane transport, and in many Gram-negative bacteria these transporters are involved in the binding and uptake of iron–siderophore complexes, cobalamin (Koebnik, 2005) or carbohydrates (Blanvillain *et al.*, 2007). TBDRs were matched to Bacteroidetes (40 total, 37 summer only, 2 winter only and 1 common to both seasons); Gammaproteobacteria (27 total, 18 summer, 6 winter and 3 both seasons) mostly to Alteromonadales and OMG; and Alphaproteobacteria (6 total, 5 summer and 1 winter) with most matches to *S. alaskensis* (Table 1 and Supplementary Table S1). Spectra referable to TBDRs were significantly more abundant in summer than in winter. The identification of a large number of Bacteroidetes TBDRs is consistent with most (78%) Bacteroidetes protein identifications being predicted OMPs. Many of the TBDRs that matched to Bacteroidetes appeared to be involved in the acquisition and uptake of polymeric substrates mediated by cellular attachment (see section *Attachment* below). For marine Gammaproteobacteria (for example, OMG) and Alphaproteobacteria (for example, *S. alaskensis*), some TBDRs may be deployed for scavenging iron, or carbohydrates that are free in seawater or associated with particulate organic matter (Blanvillain *et al.*, 2007; Ting *et al.*, 2010). Although DOC levels were comparable in the winter and summer samples, bacterial production rates were around six times higher in summer compared with winter, as indicated by leucine incorporation rates (Materials and methods). Chlorophyll *a* levels in the summer were much higher than in winter, and we hypothesize that TBDR-mediated uptake is adapted for the acquisition of biopolymers and other nutrients derived from phytoplankton. Overall, the data indicate a greater reliance on TBDR-mediated uptake by Bacteroidetes and certain Gammaproteobacteria and Alphaproteobacteria, especially in the summer, whereas SAR11 and Rhodobacterales rely more on labile solute capture all-year-round by PBP-dependent high-affinity uptake systems.

Flagellin proteins were detected for *P. haloplanktis*, *Moritella* sp. PE36 and *S. alaskensis* (Table 1 and Supplementary Table S1) suggesting mobility is advantageous for these bacteria in the Southern Ocean. *P. haloplanktis* is a copiotroph that specializes in the exploitation of short-lived nutrient patches (Stocker *et al.*, 2008), and the oligotrophic lifestyle of *S. alaskensis* has nevertheless been linked to the pursuit and utilization of nutrient pulses (Lauro *et al.*, 2009; Williams *et al.*, 2009).

Attachment

The proteins matching Bacteroidetes were mostly cell surface proteins associated with particle adhesion and nutrient acquisition (Supplementary Table S1). In contrast, only a few transcription and translation proteins and only one metabolic enzyme (fumarase) were detected (Supplementary Table S1). Frequently detected OMPs included a gliding motility protein (GldL) involved in the exploration of solid surfaces, and components of the TBDR and SusCD/RagAB complexes (Table 1 and Supplementary Table S1). SusCD is associated with the binding of exogenous polysaccharides, which are then degraded and imported into the periplasm (Reeves *et al.*, 1997; Gilbert, 2008). RagAB is involved in binding exogenous proteins, and subsequent uptake of their degradation products as peptides (Nagano *et al.*, 2007). When sampled, the planktonic Flavobacteria that were captured on

filters may have had minimal metabolic activity, but were poised to detect and exploit particulate matter. The metaproteomic data corroborate the view that marine Flavobacteria are specialized for attachment and growth on detrital particles (especially associated with the decay phase of phytoplankton blooms), and rely mostly on biopolymers for carbon and energy, in preference to the labile solutes favored by heterotrophic Proteobacteria (Cottrell and Kirchman, 2000; Abell and Bowman, 2005; González *et al.*, 2008).

Ammonia oxidation and carbon fixation

Two chemolithoautotrophic pathways are indicated in the Antarctic winter metaproteome: the 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HB) cycle and the reverse tricarboxylic acid cycle (Table 2). The former is represented by proteins assigned to the 3HP/4HB cycle, a uniquely archaeal pathway for carbon fixation present in AOA such as *N. maritimus*. The full pathway has yet to be characterized, but homologs of proteins referred to this pathway in MGI and other archaeal species that use the 3HP/4HB cycle (Berg *et al.*, 2007; Kockelkorn and Fuchs, 2009; Walker *et al.*, 2010; Ramos-Vera *et al.*, 2011) were found exclusively in the winter metaproteome (Table 2 and Supplementary Table S2). These data are consistent with the high abundance of MGI in coastal surface Antarctic waters, especially in winter (Murray *et al.*, 1998; Church *et al.*, 2003; Murray and Grzymski, 2007; Grzymski *et al.*, 2012).

For AOA, ammonia oxidation is intimately linked to the 3HP/4HB cycle because it provides the energy necessary for carbon fixation. We identified proteins from Crenarchaeota for ammonia transport and ammonia oxidation; a previous metaproteomic study detected ammonia monooxygenase subunits from *N. maritimus* at coastal upwelling sites in the South Atlantic, providing evidence for archaeal nitrification at the surface in summer (Morris *et al.*, 2010). Our study also detected components of the proposed respiratory pathway for AOA, including NADH dehydrogenase, multicopper oxidase, and other copper-containing proteins and ferredoxin involved in electron transfer reactions required for energy metabolism (Table 2 and Supplementary Table S2) (Könneke *et al.*, 2005; Walker *et al.*, 2010). Our data support the hypothesis that MGI are important in nitrification and carbon fixation in Antarctic coastal surface waters during the winter months.

The presence of an ammonia monooxygenase subunit from a nitrifying member of the Betaproteobacteria in the winter metaproteome indicates that ammonia-oxidizing bacteria (AOB) are also present in winter surface waters. Notably, two ammonia-oxidizing Betaproteobacteria-associated RuBisCO genes were detected in the winter metagenome (Grzymski *et al.*, 2012). Ammonia oxidation by AOA and AOB leads to nitrite, which is subsequently oxidized to nitrate to complete the process of nitrification. Two nitrite oxidoreductase alpha subunits were detected in the winter metaproteome, which showed the highest identities to proteins from Nitrospirae (Lücker *et al.*, 2010) and anammox Planctomycetes (Strous *et al.*, 2006) (Table 2 and Supplementary Table S1). The presence of pyruvate:ferredoxin oxidoreductase, also from the Nitrospirae clade (Table 2 and Supplementary Table S1), provides the sole evidence for operation of the reverse tricarboxylic acid cycle for carbon fixation in our metaproteome data. This chemoautotrophic pathway has been documented for the nitrite-oxidizing bacteria *Ca.* Nitrospira defluvii and *Ca.* Leptospirillum rubarum (both Nitrospirae) with the potential for mixotrophic growth inferred for the latter (Goltsman *et al.*, 2009; Lücker *et al.*, 2010).

Other detected MGI proteins include those belonging to an oxidative Krebs cycle, and gluconeogenesis (including fructose-1,6-bisphosphatase), with phosphoenolpyruvate carboxykinase used to generate phosphoenolpyruvate from oxaloacetate (Table 1 and Supplementary Table S1) (Danson *et al.*, 2007). MGI biosynthetic pathways also identified in the metaproteome include enzymes of individual amino-acid biosynthesis pathways, and the biosynthesis of nucleotides and various cofactors/coenzymes, including cobalamin (Supplementary Table S2).

A potential for mixotrophic growth has been suggested for *N. maritimus*, based on the genomic presence of genes that encode transporters for organic compounds (amino acids, oligopeptides and taurine; Walker *et al.*, 2010). Mixotrophic growth has also been inferred for the soil AOA *Nitrosphaera viennensis*, based on stronger growth when provided with pyruvate, albeit carbon assimilation is driven predominantly by CO_2 (as bicarbonate) fixation (Tourna *et al.*, 2011). No unambiguous ABC transporters from Crenarchaeota were detected in the winter metaproteome (Supplementary Table S2). It is possible that the proteases and amino-acid degradation enzymes that were detected were associated with protein turnover by the proteasome. An archaeal cysteine desulfurase was detected in the metaproteome, indicating that cysteine can serve as a source of sulfur for biosynthetic purposes, such as iron–sulfur cluster proteins, which were especially abundant in the metaproteome and attributable to MGI (Table 1 and Supplementary Table S1).

Ammonia assimilation

The identification of proteins associated with the uptake and incorporation of ammonia (ammonium transporter, glutamine synthetase (GS), glutamate synthase (GOGAT) and glutamate dehydrogenase (GDH)) indicate that the Antarctic coastal bacteria and archaea that were sampled were metabolically

poised to assimilate inorganic reduced nitrogen (Table 1 and Supplementary Table S1). However, the metaproteome data also suggest that these bacteria and archaea adopt divergent pathways for ammonia assimilation, consistent with proposed niche partitioning in marine habitats (Martens-Habbena *et al.*, 2009). The bacteria (including SAR11 and GSO–EOSA-1) appear to employ the GS–GOGAT cycle, whereas AOA utilize GDH (Table 1 and Supplementary Table S1). An extremely high specific affinity for dissolved ammonia has been demonstrated for *N. maritimus* strain SCM1, providing this AOA with a competitive advantage under oligotrophic conditions (Martens-Habbena *et al.*, 2009). One suggestion is that the extremely high capacity for ammonia uptake by AOA is due to overexpression of ammonium transporters (Stewart *et al.*, 2012). If so, then the resulting intracellular accumulation could allow for ammonia assimilation via the lower-affinity enzyme GDH rather than the higher-affinity GS.

Taurine degradation

Taurine (2-aminoethanesulfonic acid) is an organosulfonate found in the tissues of marine invertebrates and certain algae, and is a potential source of carbon, nitrogen and sulfur. Further, growth of SAR11 cultures on taurine has previously been experimentally demonstrated (Schwalbach *et al.*, 2010). Proteins that could be assigned to a function in the taurine degradation pathway were identified in the metaproteome, including taurine–pyruvate aminotransferase and sulfoacetaldehyde acetyltransferase (Table 1 and Supplementary Table S1; Figure 4). Complete taurine degradation yields acetyl-CoA for carbon assimilation, and ammonia that can be assimilated as a nitrogen source via the GS–GOGAT cycle. The use of acetate as a carbon source for cellular biomass production is consistent with glyoxylate cycle enzymes also being detected (isocitrate lyase and malate synthase) for a range of heterotrophs, including SAR11 (Table 1 and Supplementary Table S1). Acetate is also a product of dimethylsulfoniopropionate (DMSP) degradation (see section *Methionine and DMSP degradation* below). The detection of taurine degradation enzymes from a range of bacteria, combined with the presence of ABC-type taurine/sulfonate transporters, suggests that this organosulfonate is a favored substrate in coastal Antarctic Peninsula waters. SAR11 taurine transporters were also detected in the Sargasso Sea (Sowell *et al.*, 2009).

Taurine desulfonation by sulfoacetaldehyde acetyltransferase releases sulfite anion (Figure 4), which is highly reactive and can cause damage to DNA and proteins (Kappler, 2011). SAR11 is unable to reductively assimilate oxidized sulfur, possibly because of the bioenergetic costs required for the import and subsequent reduction of sulfate (Tripp *et al.*, 2008). The ability to oxidize sulfite to sulfate is suggested by the presence of adenylylsulfate (APS) reductase (AprAB) and a putative membrane anchor protein (AprM) in SAR11 genomes; the SAR11 APS reductase alpha subunit (AprA) was detected in the metaproteome (Table 1 and Supplementary Table S1). This suggests that sulfite is oxidatively bound to AMP to generate the activated intermediate APS, as a mechanism for the intracellular detoxification of sulfite (Meyer and Kuever, 2007).

There is no evidence in SAR11 genomes of ATP sulfurylase or APS adenylyltransferase (APAT) for substrate-level phosphorylation to ATP or ADP. Nevertheless, immediately adjacent to the *aprABM* gene cluster is a gene that encodes a HIT family hydrolase, homologous to adenylylsulfatase. We posit that SAR11 is capable of enzymatic hydrolysis of APS to AMP and sulfate, or, owing to the inherent instability of the phosphosulfur anhydride bond, hydrolysis of APS is spontaneous (Figure 4). According to our hypothesis, AMP would be regenerated and sulfate subsequently released from the cell. Efficient detoxification of sulfite requires available intracellular AMP. Therefore, as an oligotroph this mechanism could be advantageous to SAR11 as AMP would be recycled for subsequent sulfite detoxification. If taurine is metabolized in this way by SAR11 and other heterotrophic bacteria, the breakdown of taurine to produce inorganic sulfur compounds could represent an important route by which these sulfur species (sulfate, sulfite and thiosulfate) are produced at the ocean surface.

Glycine betaine degradation and one-carbon (C1) metabolism

Glycine betaine is a compatible solute and a potentially important cryoprotectant in the Antarctic environment. The large number of glycine betaine transport proteins identified from both seasons (Table 1 and Supplementary Table S1) suggests that bacteria are importing this osmolyte. Glycine betaine can also be catabolized to glycine; this may be of particular importance for SAR11, which has been found to be functionally auxotrophic for glycine (Tripp *et al.*, 2009). Two of the three proteins involved in the degradation pathway of glycine betaine to glycine (Sun *et al.*, 2011) were detected in the metaproteome (dimethylglycine dehydrogenase and sarcosine oxidase), with best matches to SAR11 (Table 1 and Supplementary Table S1; Figure 5). Both of these oxidation steps generate formaldehyde as a byproduct, which would spontaneously degrade to formate. Formate could then be oxidized to CO_2 by formate dehydrogenase and serve as a source of reductant, with the CO_2 released from the cell (Sowell *et al.*, 2008). Alternatively, the formate could be activated by formyl-tetrahydrofolate ligase and directed to C1 metabolism for biosynthesis (Figure 5). Proteins pertaining to both possible fates were detected in the metaproteome (Supplementary Table S1; Figure 5).

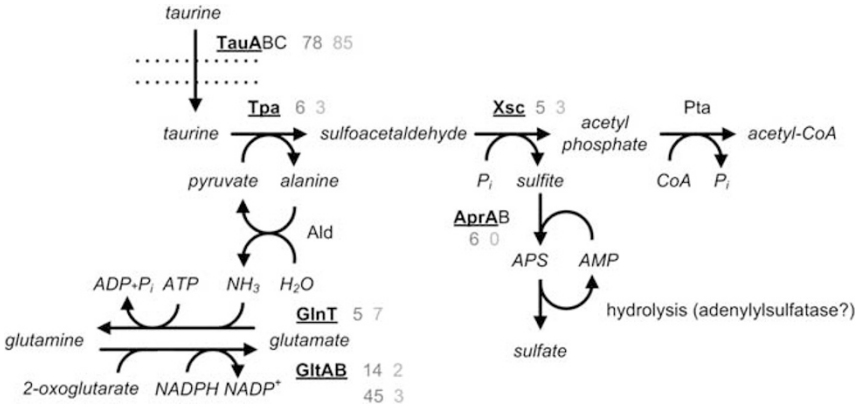

Figure 4 Putative pathway for the degradation of taurine in SAR11 and associated detoxification of sulfite. Genes for all proteins were found in SAR11 genomes; bolded and underlined proteins were detected in the metaproteome; and spectral counts per protein are shown for winter (first number after the protein) and summer (second number after the protein). Ald, alanine dehydrogenase; AprAB, adenylylsulfate (APS) reductase; GlnT, glutamine synthetase; GltAB, glutamate synthase; Pta, phosphotransacetylase; TauABC, taurine/sulfonate ABC transporter; Tpa, taurine-pyruvate aminotransferase; Xsc, sulfoacetaldehyde acetyltransferase.

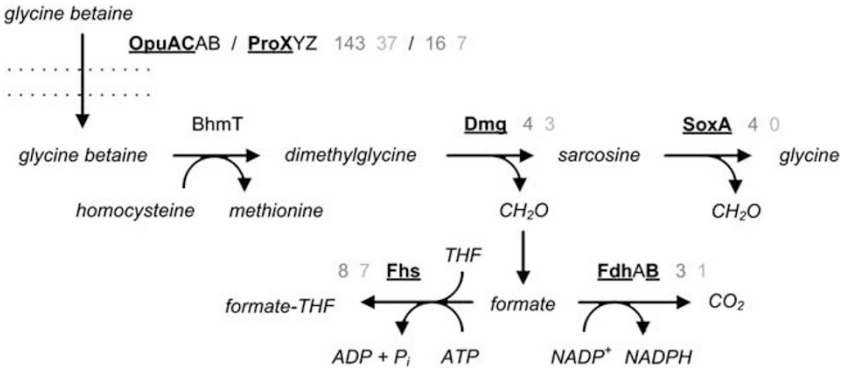

Figure 5 Putative pathway for the degradation of glycine betaine in SAR11. Genes for all proteins were found in SAR11 genomes; bolded and underlined proteins were detected in the metaproteome; and spectral counts per protein are shown for winter (first number after the protein) and summer (second number after the protein). BhmT, betaine-homocysteine methyltransferase; Dmg, dimethylglycine dehydrogenase; FhdAB, formate dehydrogenase; Fhs, formate-tetrahydrofolate ligase; OpuACAB, glycine betaine ABC transporter; ProXYZ, proline/glycine betaine ABC transporter; SoxA, sarcosine oxidase (monomeric); THF, tetrahydrofolate.

Other proteins related to C1 metabolism include two methanol dehydrogenase large subunit proteins from summer and winter that belong to the OM43 clade of Betaproteobacteria, with best matches to the obligately methylotrophic strain HTCC2181 (Giovannoni *et al.*, 2008). Consistent with the view that methylotrophic members of the OM43 clade are important components of coastal marine ecosystems (Giovannoni *et al.*, 2008; Sowell *et al.*, 2011), our data imply that C1 compounds such as methanol and formaldehyde are able to be metabolized. Possible sources of methanol include deposition from the atmosphere, a byproduct of marine phytoplankton growth (Heikes *et al.*, 2002) or photochemical degradation of surface DOC (Dixon *et al.*, 2011).

Methionine and DMSP degradation

The ABC transporter for glycine betaine (OpuA) has also been shown to import DMSP, an algal osmolyte and reduced sulfur source (Holtmann and Bremer, 2004). The general amino-acid transporter (Yhd) has been proposed to import methionine in SAR11 (Sowell *et al.*, 2009). In the absence of genes for the reductive assimilation of sulfur, SAR11 requires exogenous sources of reduced sulfur, such as methionine or DMSP (Tripp *et al.*, 2008). While SAR11 proteins corresponding to the postulated methionine degradation pathway (Tripp *et al.*, 2008) were not detected, two proteins in the pathway (adenosylhomocysteinase and *O*-acetyl-homoserine [thiol]-lyase) that matched other bacterioplankton were detected (Table 1 and Supplementary Table S1). The former matched to a range of bacteria, and MGI, suggesting that the biosynthesis of cysteine from methionine is a relatively common pathway in the Antarctic community that was sampled.

A novel pathway for the assimilation of DMSP was recently demonstrated for the *Roseobacter* clade bacterium *Ruegeria pomeroyi* (Reisch *et al.*, 2011). Following initial demethylation of DMSP, the

intermediate methylmercaptopropionate (MMPA) is further catabolized to methanethiol, which can be readily assimilated as a reduced sulfur source (Reisch *et al.*, 2011). SAR11 was also shown to possess enzyme activities for the first two steps of MMPA catabolization (Reisch *et al.*, 2011). The first step is mediated by DMSP demethylase (DmdA), a homolog of the glycine cleavage T-protein (GcvT). GcvT homologs were detected in the metaproteome, and we infer two (one from SAR11 and one from OMG) to be DmdA (Table 1 and Supplementary Table S1). SAR11 proteins for the subsequent MMPA catabolization steps were not detected in the metaproteome, although two proteins that correspond to the enzyme that catalyzes the second step of MMPA catabolization, methylmercaptopropionyl-CoA ligase (DmdC), were detected for Gammaproteobacteria (including OMG) (Table 1 and Supplementary Table S1). A SAR11 acetaldehyde dehydrogenase was detected for the oxidation of acetaldehyde (another product of MMPA catabolization) to acetate.

GSO–EOSA-1 bacteria (ARCTIC96BD-19/SUP05)

Many proteins in the summer and winter metaproteome had the highest identity to proteins from members of the SUP05 cluster of Gammaproteobacteria, including the clam symbiont *Ca.* Ruthia magnifica (Newton *et al.*, 2007), and the free-living SUP05 Walsh *et al.*, 2009). The SUP05 and ARCTIC96BD-19 clusters together make up the GSO–EOSA-1 complex, and pelagic members of this group are inferred to be major contributors to global chemolithoautotrophic carbon fixation via sulfur oxidation (Walsh *et al.*, 2009; Swan *et al.*, 2011). The SUP05 cluster has been reported in high abundance in some oxygen minimum zones (Walsh *et al.*, 2009), and the closely related ARCTIC98BD-19 cluster have been recovered from diverse oceanic habitats and were recently reported to be widespread in the mesopelagic Atlantic and Pacific oceans (Swan *et al.*, 2011). Overall, our metaproteomic data suggest that GSO–EOSA-1 bacteria are metabolically active in the aerobic surface waters of the Antarctic coast, with a higher representation in winter. As well as components of the transcription and translation machinery, these included proteins pertaining to ammonia assimilation, energy generation (subunits of cytochrome c oxidase (Cox) and ATP synthase) and biosynthesis (for example, amino acids and nucleotides) (Table 1 and Supplementary Table S1).

SUP05 has the potential for chemolithoautotrophy using reduced sulfur compounds coupled to oxygen or nitrate reduction and carbon fixation using form II RuBisCO (Walsh *et al.*, 2009), while the genome sequence of a member of the ARCTIC96BD-19 cluster indicates the capacity for carbon fixation using form I RuBisCO and sulfur oxidation (AprA), and the potential for mixotropic growth (Swan *et al.*, 2011). Sequences from the winter metagenomic library matched genes in *Ca.* R. magnifica required for APS reductase, sulfur oxidation (Sox) and reverse dissimilatory sulfate reduction pathway (Grzymski *et al.*, 2012), with the latter possibly used for the oxidation of stored sulfur (S^0). The winter metagenome also contained closely related sequences to ARCTIC96BD-19 cluster related form I RuBisCO (Swan *et al.*, 2011). The detection of CoxII in the winter metaproteome indicates an aerobic metabolism by GSO–EOSA-1 bacteria, whereas the detection of AprA (both seasons; Supplementary Table S1) could indicate sulfur oxidation at the surface. Studies examining the metabolic activity of GSO–EOSA-1 bacteria in aerobic waters are needed in order to determine the ecological role that these bacterioplankton have in Southern Ocean waters.

PR

PRs are photoactive cell membrane proteins in bacteria (Béjà *et al.*, 2000, 2001; DeLong and Béjà, 2010). Three PRs were detected in our metaproteome (both seasons), two of which had best matches to SAR11 and the third to the OMG group (Table 1 and Supplementary Table S1). Two of the PRs are inferred to be green-light-absorbing (one SAR11 and one OMG) and the other blue-light-absorbing (SAR11) (Man *et al.*, 2003) (Table 1 and Supplementary Table S1). Laboratory studies of *Ca.* P. ubique strain HTCC1062 (which has the 'green' PR) indicate that PR serves as a light-dependent proton pump, and is important under illuminated, carbon-limiting conditions when PR-mediated ATP generation can substitute for respiration (Giovannoni *et al.*, 2005; Steindler *et al.*, 2011). Although 'green' PR is optimal for capturing light at the surface and 'blue' for light deep in the photic zone, both variants are found in shallow water (Fuhrman *et al.*, 2008). Constitutive expression of PR for light harvesting in *Ca.* P. ubique may facilitate the ability to immediately respond to cellular energy deficits caused by carbon starvation (Steindler *et al.*, 2011).

Chaperones and quality control

Protein chaperones were highly abundant in the metaproteome, including DnaK, HtpG, peptidyl–prolyl *cis/trans* isomerases, trigger factor, and bacterial GroEL/GroES and the equivalent archaeal chaperonin Cpn60. Proteins of the FtsH–HflKC membrane complex involved in proteolytic turnover of membrane and soluble proteins were also detected. Chaperones and proteases were abundant in both seasons, and the high abundance of these proteins has been linked to microbial survival in surface ocean water as a response to exposure to environmental stresses (Sowell *et al.*, 2009, 2011), as well as oxidation of endogenous biomass via

proteolysis under carbon-limiting conditions (Steindler *et al.*, 2011). Additionally, proteasome, exosome and degradosome (particularly polyribonucleotide nucleotidyltransferase) components were detected. Consistent with the detection of MGI proteins, the archaeal proteasome and exosome proteins were found exclusively in winter. Proteomic studies of the Antarctic archaeon *M. burtonii* have linked the abundance of proteasome and exosome proteins to cold adaptation (Williams *et al.*, 2010). One possibility is that the inhibitory effects of low temperature on translation, including correct folding of the nascent polypeptide, leads to an increase in the number of irretrievably misfolded proteins. Cold temperatures also lead to an increase in adventitious secondary structures formed by RNA transcripts, which can be resolved by DEAD box RNA helicases, the latter were found in the metaproteome (two in winter and one in summer) (Supplementary Table S1). Increased exosome and degradosome levels are consistent with the need to degrade irreparably damaged or tangled RNA.

Conclusions

Our metaproteomic analysis provided insight into functionally important transport, metabolism and energy generation processes in surface water communities in the coastal WAP. Bacterial ABC, TRAP and TBDR transporters have a critical role in nutrient scavenging. We have obtained evidence that glycine betaine is being metabolized to glycine in Antarctic coastal waters to redress conditional glycine auxotrophy in SAR11 (Tripp *et al.*, 2009). We also posit that the purpose of APS reductase *in lieu* of sulfate reduction by SAR11 (Tripp *et al.*, 2008) is to detoxify sulfite that is generated endogenously (for example, by taurine desulfonation) without providing any benefit in energy conservation. Thus, the prevalence of *aprA* in functional gene studies of environmental samples may be partly explained by the utility of a reverse-acting APS reductase in organic sulfur compound degradation, including by SAR11 (Meyer and Kuever, 2007).

MGI have been proposed to have an important role in oceanic nitrification, based on the detection of these organisms in high abundance in coastal Antarctic surface waters (DeLong *et al.*, 1994; Murray and Grzymski, 2007) and studies from the North Sea, which demonstrated that MGI abundance correlates with ammonia oxidation to nitrite (Wuchter *et al.*, 2006). Components of the 3HP/4HB cycle and the oxidation of ammonia to nitrite were detected in the winter metaproteome indicating that this pathway is active and the MGI are causing nitrification in coastal WAP waters. Ammonia-oxidizing and nitrite-oxidizing bacteria were also represented in the data, indicating that nitrification in these waters involved multiple species. As this nitrification process fuels the fixation of inorganic carbon, the metaproteome corroborates metagenomic and SSU ribosomal RNA interpretations that the chemolithoautotrophic archaea and bacteria constitute a major CO_2 sink during the cold, dark winter months (Grzymski *et al.*, 2012). Overall, these data emphasize the role that specific types of Southern Ocean microorganisms have, year round, in primary production and remineralization in a coastal region of West Antarctica.

Acknowledgements

We extend our thanks to our field team M Erickson, K Myers, V Peng and J-F Ghiglione for their expert assistance sampling in Antarctica and to the Palmer Station personnel in the US Antarctic Program for support, especially during the winter field season. The work was supported by the NSF grants, ANT 0632389 to AEM and JJG, and ANT 0632278 and ANT 0217282 to HD. The work of the Australian contingent was supported by the Australian Research Council. Mass spectrometric results were obtained at the Bioanalytical Mass Spectrometry Facility within the Analytical Centre of the University of New South Wales. This work was undertaken using infrastructure provided by NSW Government co-investment in the National Collaborative Research Infrastructure Scheme. Subsidized access to this facility is gratefully acknowledged.

References

Abell GCJ, Bowman JP. (2005). Ecological and biogeographic relationships of class Flavobacteria in the Southern Ocean. *FEMS Microbiol Ecol* **51**: 265–277.

Allen M, Lauro FM, Williams TJ, Burg D, Siddiqui KS, De Francisci D *et al.* (2009). The genome sequence of the psychrophilic archaeon, *Methanococcoides burtonii*: the role of genome evolution in cold-adaptation *ISME J* **3**: 1012–1035.

Australian Commonwealth Government *Australian Antarctic Science Strategic Plan 2011–12 to 2020–21 (2011). Australian Commonwealth Government: ISBN 978 1 876934 16 6.*

Béjà O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP *et al.* (2000). Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **15**: 1902–1906.

Béjà O, Spudich EN, Spudich JL, Leclerc M, DeLong EF. (2001). Proteorhodopsin phototrophy in the ocean. *Nature* **411**: 786–789.

Berg IA, Kockelkorn D, Buckel W, Fuchs G. (2007). A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* **318**: 1782–1786.

Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N *et al.* (2007). Plant carbohydrate scavenging through TonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria *PLoS One* **2**: 224.

Buchan A, González JM, Moran MA. (2005). Overview of the marine *Roseobacter* lineage. *Appl Environ Microbiol* **71**: 5665–5677.

Burg DW, Lauro FM, Williams TJ, Raftery MJ, Guilhaus M, Cavicchioli R. (2010). Analyzing the hydrophobic proteome of the antarctic archaeon *Methanococcoides burtonii* using differential solubility fractionation *J Proteome Res* **9**: 664–676.

Canfield DE, Stewart FJ, Thamdrup B, De Brabandere L, Dalsgaard T, Delong EF *et al.* (2010). A cryptic sulfur cycle in oxygen-minimum-zone waters off the Chilean coast. *Science* **330**: 1375–1378.

Cho JC, Giovannoni SJ. (2004). Cultivation and growth characteristics of a diverse group of oligotrophic marine *Gammaproteobacteria*. *Appl Environ Microbiol* **70**: 432–440.

Church MJ, DeLong EF, Ducklow HW, Karner MB, Preston CM, Karl DM. (2003). Abundance and distribution of planktonic Archaea and Bacteria in the waters west of the Antarctic Peninsula. *Limnol Oceanogr* **48**: 1893–1902.

Cottrell MT, Kirchman DL. (2000). Natural assemblages of marine Proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular weight dissolved organic matter. *Appl Environ Microbiol* **66**: 1692–1697.

Danson MJ, Lamble HJ, Hough DW. (2007). Central metabolism. In: Cavicchioli R (ed): *Archaea: Molecular and Cellular Biology*. ASM Press: Washington DC, pp 260–287.

DeLong EF, Béjà O. (2010). The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. *PLoS Biol* **8**: e1000359.

DeLong EF, Wu KY, Prezelin BB, Jovine RVM. (1994). High abundance of Archaea in Antarctic marine picoplankton. *Nature* **371**: 695–697.

DeMaere MZ, Lauro FM, Thomas T, Yau S, Cavicchioli R. (2011). Simple high-throughput annotation pipeline (SHAP). *Bioinformatics* **27**: 2431–2432.

Dixon JL, Beale R, Nightingale PD. (2011). Rapid biological oxidation of methanol in the tropical Atlantic: significance as a microbial carbon source. *Biogeosci Discuss* **8**: 3899–3921.

Farmer CT, Hansell DH. (2007). Determination of dissolved organic carbon and total dissolved nitrogen in sea water. In: Dickson AG, Sabina CL, Christian JR (eds): *Guide to Best Practices for Ocean CO_2 Measurements*. PICES Special Publication: Sydney, British Columbia, Canada, p 3191.

Forward JA, Behrendt MC, Wyborn NR, Cross R, Kelly DJ. (1997). TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *J Bacteriol* **179**: 5482–5493.

Fuhrman JA, Schwalbach MS, Stingl U. (2008). Proteorhodopsins: an array of physiological roles? *Nat Rev Microbiol* **6**: 488–494.

Ghiglione JF, Murray AE. (2012). Pronounced summer to winter differences and higher wintertime richness in coastal sub-Antarctic and Antarctic marine bacterioplankton. *Environ Microbiol* **14**: 617–629.

Gilbert HJ. (2008). Sus out sugars in. *Structure* **16**: 987–989.

Giovannoni SJ, Bibbs L, Cho JC, Stapels MD, Desiderio R, Vergin KL *et al.* (2005). Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* **438**: 82–85.

Giovannoni SJ, Hayakawa DH, Tripp HJ, Stingl U, Givan SA, Cho JC *et al.* (2008). The small genome of an abundant coastal ocean methylotroph. *Environ Microbiol* **10**: 1771–1782.

Goltsman DS, Denef VJ, Singer SW, VerBerkmoes NC, Lefsrud M *et al.* (2009). Community genomic and proteomic analyses of chemoautotrophic iron-oxidizing ''*Leptospirillum rubarum*'' (Group II) and ''*Leptospirillum ferrodiazotrophum*'' (Group III) bacteria in acid mine drainage biofilms. *Appl Environ Microbiol* **75**: 4599–4615.

Gómez-Consarnau L, Gonzalez JM, Coll-Llado M, Gourdon P, Pascher T, Neutze R *et al.* (2007). Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature* **445**: 210–213.

González JM, Fernández-Gómez B, Fernàndez-Guerra A, Gómez-Consarnau L, Sánchez O, Coll-Lladó M *et al.* (2008). Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). *Proc Natl Acad Sci USA* **105**: 8724–8729.

Grossart H-P, Levold F, Allgaier M, Simon M, Brinkhoff T. (2005). Marine diatom species harbour distinct bacterial communities. *Environ Microbiol* **7**: 860–873.

Grzymski JJ, Riesenfeld CS, Williams TJ, Dussaq AM, Ducklow H, Erickson M *et al.* (2012). A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *ISME J*; e-pub ahead of print; doi:10.1038/ismej.2012.31.

Hallam SJ, Konstantinidis KT, Putnam N, Schleper C, Watanabe Y, Sugahara J *et al.* (2006a). Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc Natl Acad Sci USA* **103**: 18296–18301.

Hallam SJ, Mincer TJ, Schleper C, Preston CM, Roberts K, Richardson PM *et al.* (2006b). Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine Crenarchaeota. *PLoS Biol* **4**: e95.

Hayashi NR, Arai H, Kodama T, Igarashi Y. (1997). The novel genes, *cbbQ* and *cbbO*, located downstream from the RubisCO genes of *Pseudomonas hydrogenothermophila*, affect the conformational states and activity of RubisCO. *Biochem Biophys Res Commun* **241**: 565–569.

Heikes BG, Chang W, Pilson MEQ, Swift E, Singh HB, Guenther A *et al.* (2002). Atmospheric methanol budget and ocean implication. *Global Biogeochem Cycles* **16**: 1133.

Holtmann G, Bremer E. (2004). Thermoprotection of *Bacillus subtilis* by exogenously provided glycine betaine and structurally related compatible solutes: involvement of Opu transporters. *J Bacteriol* **186**: 1683–1693.

Kappler U. (2011). Bacterial sulfite-oxidizing enzymes. *Biochim Biophys Acta* **1807**: 1–10.

Karner M, DeLong EF, Karl DM. (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507–509.

Kirchman D, K'Nees E, Hodson R. (1985). Leucine incorporation and its potential as a measure of protein synthesis by bacteria in natural aquatic systems. *Appl Environ Microbiol* **49**: 599–607.

Kockelkorn D, Fuchs G. (2009). Malonic semialdehyde reductase, succinic semialdehyde reductase, and succinyl-coenzyme A reductase from *Metallosphaera sedula*: enzymes of the autotrophic 3-hydroxypropionate/

4-hydroxybutyrate cycle in Sulfolobales. *J Bacteriol* **191**: 6352–6362.

Koebnik R. (2005). TonB-dependent trans-envelope signalling: the exception or the rule? *Trends Microbiol* **13**: 343–347.

Könneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon *Nature* **437**: 543–546.

Lauro FM, DeMaere MZ, Yau S, Brown M, Ng C, Wilkins D *et al.* (2011). An integrative study of a meromictic lake ecosystem in Antarctica. *ISME J* **5**: 879–895.

Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, Rice S *et al.* (2009). The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci USA* **106**: 15527–15533.

Lücker S, Wagner M, Maixner F, Pelletier E, Koch H, Vacherie B *et al.* (2010). A *Nitrospira* metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *Proc Natl Acad Sci USA* **107**: 13479–13484.

Man D, Wang W, Sabehi G, Aravind L, Post AF, Massana R *et al.* (2003). Diversification and spectral tuning in marine proteorhodopsins. *EMBO J* **22**: 1725–1731.

Markowitz VM, Chen IM, Chu K, Szeto E, Palaniappan K, Grechkin Y *et al.* (2012). IMG/M: the integrated metagenome data management and comparative analysis system. *Nucleic Acids Res* **40**: (Database issue) D123–D129.

Martens-Habbena W, Berube PM, Urakawa H, de la Torre JR, Stahl DA. (2009). Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* **461**: 976–979.

Massana R, Pedros-Alio C, Casamayor EO, Gasol JM. (2001). Changes in marine bacterioplankton phylogenetic composition during incubations designed to measure biogeochemically significant parameters. *Limnol Oceanogr* **46**: 1181–1188.

Meyer B, Kuever J. (2007). Molecular analysis of the diversity of sulfate-reducing and sulfur-oxidizing prokaryotes in the environment, using *aprA* as functional marker gene. *Appl Environ Microbiol* **73**: 7664–7679.

Moran MA, Belas R, Schell MA, González JM, Sun F, Sun S *et al.* (2007). Ecological genomics of marine Roseobacters. *Appl Environ Microbiol* **73**: 4559–4569.

Moran MA, Zepp RG. (1997). Role of photoreactions in the formation of biologically labile compounds from dissolved organic matter. *Limnol Oceanogr* **42**: 1307–1316.

Morris RM, Nunn BL, Frazar C, Goodlett DR, Ting YS, Rocap G. (2010). Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* **4**: 673–685.

Murray AE, Grzymski JJ. (2007). Diversity and genomics of Antarctic marine micro-organisms. *Philos Trans R Soc Lond B Biol Sci* **362**: 2259–2271.

Murray AE, Preston CM, Massana R, Taylor LT, Blakis A, Wu K *et al.* (1998). Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters off Anvers Island, Antarctica. *Appl Environ Microbiol* **64**: 2585–2595.

Nagano K, Murakami Y, Nishikawa K, Sakakibara J, Shimozato K, Yoshimura F. (2007). Characterization of RagA and RagB in Porphyromonas gingivalis: study using gene deletion mutants. *J Med Microbiol* **56**: 1536–1548.

Newton IL, Woyke T, Auchtung TA, Dilly GF, Dutton RJ, Fisher MC *et al.* (2007). The *Calyptogena magnifica* chemoautotrophic symbiont genome. *Science* **315**: 998–1000.

Ng C, DeMaere MZ, Williams TJ, Lauro FM, Raftery M, Gibson JA *et al.* (2010). Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica. *ISME J* **4**: 1002–1019.

Piquet AM, Bolhuis H, Meredith MP, Buma AG. (2011). Shifts in coastal Antarctic marine microbial communities during and after melt water-related surface stratification. *FEMS Microbiol Ecol* **76**: 413–427.

Ramos-Vera WH, Weiss M, Strittmatter E, Kockelkorn D, Fuchs G. (2011). Identification of missing genes and enzymes for autotrophic carbon fixation in crenarchaeota. *J Bacteriol* **193**: 1201–1211.

Reeves AR, Wang G-R, Salyers AA. (1997). Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*. *J Bacteriol* **179**: 643–649.

Reisch CR, Stoudemayer MJ, Varaljay VA, Amster IJ, Moran MA, Whitman WB. (2011). Novel pathway for assimilation of dimethylsulphoniopropionate widespread in marine bacteria. *Nature* **473**: 208–211.

Rodriguez-Brito B, Rohwer F, Edwards RA. (2006). An application of statistics to comparative metagenomics. *BMC Bioinformatics* **7**: 162.

Sharp JH, Benner R, Carlson CA, Dow R, Fitzwater SE. (1993). Re-evaluation of high temperature combustion and chemical oxidation measurements of dissolved organic carbon in seawater. *Limnol Oceanogr* **38**: 1774–1782.

Schattenhofer M, Fuchs BM, Amann R, Zubkov MV, Tarran GA, Pernthaler J. (2009). Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ Microbiol* **11**: 2078–2093.

Schwalbach MS, Tripp HJ, Steindler L, Smith DP, Giovannoni SJ. (2010). The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. *Environ Microbiol* **12**: 490–500.

Smith DP, Kitner JB, Norbeck AD, Clauss TR, Lipton MS, Schwalbach MS *et al.* (2010). Transcriptional and translational regulatory responses to iron limitation in the globally distributed marine bacterium *Candidatus* Pelagibacter ubique. *PLoS One* **5**: e10487.

Sowell SM, Abraham PE, Shah M, Verberkmoes NC, Smith DP, Barofsky DF *et al.* (2011). Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. *ISME J* **5**: 856–865.

Sowell SM, Norbeck AD, Lipton MS, Nicora CD, Callister SJ, Smith RD *et al.* (2008). Proteomic analysis of stationary phase in the marine bacterium ''*Candidatus* Pelagibacter ubique''. *Appl Environ Microbiol* **74**: 4091–4100.

Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, Barofsky DF *et al.* (2009). Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* **3**: 93–105.

Steindler L, Schwalbach MS, Smith DP, Chan F, Giovannoni SJ. (2011). Energy starved *Candidatus* Pelagibacter ubique substitutes light-mediated ATP production for endogenous carbon respiration. *PLoS One* **6**: e19725.

Stewart FJ, Ulloa O, Delong EF. (2012). Microbial metatranscriptomics in a permanent marine oxygen minimum zone. *Environ Microbiol* **14**: 23–40.

Stocker R, Seymour JR, Samadani A, Hunt DE, Polz MF. (2008). Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches. *Proc Natl Acad Sci USA* **105**: 4209–4214.

Strous M, Pelletier E, Mangenot S, Rattei T, Lehner A, Taylor MW *et al.* (2006). Deciphering the evolution and metabolism of an anammox bacterium from a community genome. *Nature* **440**: 790–794.

Sullivan MB, Coleman ML, Weigele P, Rohwer F, Chisholm SW. (2005). Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biol* **3**: e144.

Sun J, Steindler L, Thrash JC, Halsey KH, Smith DP, Carter AE *et al.* (2011). One carbon metabolism in SAR11 pelagic marine bacteria. *PLoS One* **6**: e23973.

Swan BK, Martinez-Garcia M, Preston CM, Sczyrba A, Woyke T, Lamy D *et al.* (2011). Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science* **333**: 1296–1300.

Thomas T, Egan S, Burg D, Ng C, Ting L, Cavicchioli R. (2007). The integration of genomics and proteomics into marine, microbial ecology. Theme series on ''Genomics, proteomics and metabolomics in marine ecology''. *Mar Ecol Prog Ser* **332**: 291–299.

Ting L, Williams TJ, Cowley MJ, Lauro FM, Guilhaus M, Raftery MJ *et al.* (2010). Cold adaptation in the marine bacterium, *Sphingopyxis alaskensis*, assessed using quantitative proteomics. *Environ Microbiol* **12**: 2658–2676.

Tourna M, Stieglmeier M, Spang A, Könneke M, Schintlmeister A, Urich T *et al.* (2011). *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc Natl Acad Sci USA* **108**: 8420–8425.

Tripp HJ, Kitner JB, Schwalbach MS, Dacey JW, Wilhelm LJ, Giovannoni SJ. (2008). SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* **452**: 741–744.

Tripp HJ, Schwalbach MS, Meyer MM, Kitner JB, Breaker RR, Giovannoni SJ. (2009). Unique glycine-activated riboswitch linked to glycine-serine auxotrophy in SAR11. *Environ Microbiol* **11**: 230–238.

Turner J, Colwell SR, Marshall GJ, Lachlan-Cope TA, Carleton AM, Jones PD *et al.* (2005). Antarctic climate change during the last 50 years. *Int J Climatol* **25**: 279–294.

Walker CB, de la Torre JR, Klotz MG, Urakawa H, Pinel N, Arp DJ *et al.* (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA* **107**: 8818–8823.

Walsh DA, Zaikova E, Howes CG, Song YC, Wright JJ, Tringe SG *et al.* (2009). Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. *Science* **326**: 578–582.

Williams TJ, Burg DW, Raftery MJ, Poljak A, Guilhaus M, Pilak O *et al.* (2010). Global proteomic analysis of the insoluble, soluble, and supernatant fractions of the psychrophilic archaeon *Methanococcoides burtonii* Part I: the effect of growth temperature. *J Proteome Res* **9**: 640–652.

Williams TJ, Ertan H, Ting L, Cavicchioli R. (2009). Carbon and nitrogen substrate utilization in the marine bacterium *Sphingopyxis alaskensis* strain RB22. *ISME J* **3**: 1036–1052.

Winnen B, Hvorup RN, Saier MH Jr. (2003). The tripartite tricarboxylate transporter (TTT) family. *Res Microbiol* **154**: 457–465.

Wuchter C, Abbas B, Coolen MJ, Herfort L, van Bleijswijk J, Timmers P *et al.* (2006). Archaeal nitrification in the ocean. *Proc Natl Acad Sci USA* **103**: 12317–12322.

Yau S, Lauro FM, DeMaere MZ, Brown MV, Thomas T, Raftery MJ *et al.* (2011). Virophage control of Antarctic algal host–virus dynamics. *Proc Natl Acad Sci USA* **108**: 6163–6168.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)