Organic matter processing by microbial communities throughout the Atlantic water column as revealed by metaproteomics

Kristin Bergauer[a,1], Antonio Fernandez-Guerra[b,c], Juan A. L. Garcia[a], Richard R. Sprenger[d], Ramunas Stepanauskas[e], Maria G. Pachiadaki[e], Ole N. Jensen[d], and Gerhard J. Herndl[a,f,g]

[a]Department of Limnology and Bio-Oceanography, University of Vienna, A-1090 Vienna, Austria; [b]Microbial Genomics and Bioinformatics Research Group, Max Planck Institute for Marine Microbiology, D-28359 Bremen, Germany; [c]Oxford e-Research Centre, University of Oxford, Oxford OX1 3QG, United Kingdom; [d]Department of Biochemistry and Molecular Biology, VILLUM Center for Bioanalytical Sciences, University of Southern Denmark, DK-5230 Odense M, Denmark; [e]Bigelow Laboratory for Ocean Sciences, East Boothbay, ME 04544; [f]Department of Marine Microbiology and Biogeochemistry, Royal Netherlands Institute for Sea Research, Utrecht University, 1790 AB Den Burg, The Netherlands; and [g]Vienna Metabolomics Center, University of Vienna, A-1090 Vienna, Austria

The phylogenetic composition of the heterotrophic microbial community is depth stratified in the oceanic water column down to abyssopelagic layers. In the layers below the euphotic zone, it has been suggested that heterotrophic microbes rely largely on solubilized particulate organic matter as a carbon and energy source rather than on dissolved organic matter. To decipher whether changes in the phylogenetic composition with depth are reflected in changes in the bacterial and archaeal transporter proteins, we generated an extensive metaproteomic and metagenomic dataset of microbial communities collected from 100- to 5,000-m depth in the Atlantic Ocean. By identifying which compounds of the organic matter pool are absorbed, transported, and incorporated into microbial cells, intriguing insights into organic matter transformation in the deep ocean emerged. On average, solute transporters accounted for 23% of identified protein sequences in the lower euphotic and ~39% in the bathypelagic layer, indicating the central role of heterotrophy in the dark ocean. In the bathypelagic layer, substrate affinities of expressed transporters suggest that, in addition to amino acids, peptides and carbohydrates, carboxylic acids and compatible solutes may be essential substrates for the microbial community. Key players with highest expression of solute transporters were Alphaproteobacteria, Gammaproteobacteria, and Deltaproteobacteria, accounting for 40%, 11%, and 10%, respectively, of relative protein abundances. The in situ expression of solute transporters indicates that the heterotrophic prokaryotic community is geared toward the utilization of similar organic compounds throughout the water column, with yet higher abundances of transporters targeting aromatic compounds in the bathypelagic realm.

transporter proteins | organic matter | deep sea | Atlantic Ocean | metaproteomics

The dark ocean is, together with the deep subsurface, the most extensive and least explored biome on Earth, characterized by high hydrostatic pressure, low temperature, and low metabolic activities (1, 2). The principal carbon and energy source for the dark ocean's biota is primary production in the sunlit surface waters of the ocean and the resulting sedimentation of particulate organic matter (POM) into the ocean's interior, known as the biological carbon pump (3, 4). Once exported into the mesopelagic waters, the vertical flux of POM attenuates with depth due to direct utilization by the biota and/or solubilization via extracellular enzymatic activity by heterotrophic microbes (5, 6). This solubilization process of POM to dissolved organic matter (DOM) and the subsequent uptake of cleavage products by heterotrophic microbes are considered the rate-limiting step in microbial metabolism in the deep sea. Based on microbial activity measurements and dissolved organic carbon profiles throughout the water column, it is estimated that direct DOM utilization accounts for only about 10% of the microbial carbon

demand in the mesopelagic and bathypelagic waters (7, 8). Despite this apparent low contribution of DOM compared with POM in supporting heterotrophic microbial metabolism in the deep ocean, DOM quantity and quality decreases with depth (9). The decrease in the overall nutritional quality of the DOM with depth might reflect the generation of increasingly recalcitrant compounds by heterotrophic microbes (10) in mesopelagic and bathypelagic layers, as suggested by the microbial carbon pump hypothesis (11). Another likely mechanism, formulated in the selective preservation hypothesis, is that labile DOM compounds are preferentially used, leaving behind refractory molecules. Alternatively, it might be simply a consequence of diluting out the large number of DOM molecules, making their utilization inefficient (12). The low concentrations of organics and the extremely high diversity of DOM molecules in the dark ocean (12, 13) constitute a challenge for microbes to maintain a sufficiently high encounter rate with solutes and to minimize the energy requirements associated with compound assimilation (14).

Significance

Circumstantial evidence indicates that especially deep-ocean heterotrophic microbes rely on particulate organic matter sinking through the oceanic water column and being solubilized to dissolved organic matter (DOM) prior to utilization rather than on direct uptake of the vast pool of DOM in the deep ocean. Comparative metaproteomics allowed us to elucidate the vertical distribution and abundance of microbially mediated transport processes and thus the uptake of solutes throughout the oceanic water column. Taken together, our data suggest that, while the phylogenetic composition of the microbial community is depth stratified, the composition and substrate specificities of transporters considered in this study are ubiquitous while their relative abundance changes with depth.

PNAS PNAS PNAS

All of these factors might contribute to the overall rather inefficient microbial utilization of the large stock of deep-sea DOM and the apparent preferential utilization of sedimenting POM in the ocean's interior (15).

In this study, we integrated metaproteomic, metagenomic, and single-cell genomic analyses to elucidate protein expression patterns of microbial communities from the euphotic zone (100 m) to the dark ocean (300–4,050 m) along a large latitudinal range (67°N to 49°S) in the Atlantic Ocean. We focused on the spatial distribution and abundance of expressed transporters as key mechanisms for the uptake of essential nutrients, including the primary active ATP-binding cassette (ABC) transporters, as well as the secondary active tripartite ATP-independent periplasmic transporters (TRAP-Ts), the tripartite tricarboxylate transporters (TTTs), and the TonB-dependent transporters (TBDTs). The prime determinant of selectivity in import systems are the extracytoplasmic substrate-binding proteins (SBPs) (16, 17), which are constituents of primary and secondary active transporters and capture the substrate with high affinity from the cells' surroundings (16, 18). Besides their primary function as substrate translocators, SBPs play an important role in signal transduction (19). All SBP-dependent ABC systems are importers and ubiquitous in Bacteria and Archaea (20). Knowledge on the variability of the transporter proteins and their phylogenetic origin provides insights into the dynamics in nutrient scavenging of the microbial community in response to the organic matter (OM) availability with depth (21–26).

In this study, we addressed the following specific research questions: (*i*) Are changes in the phylogenetic composition of the microbial community tightly or loosely linked to changes in the type and abundance of transporter proteins for OM throughout the water column? (*ii*) Does the distribution of transporter protein indicate major changes in the OM compound classes used as substrate by the microbial communities between surface and bathypelagic waters?

Results and Discussion

The Transporter Repertoire of Open-Ocean Microbes. As a first step toward establishing a proteomic inventory of marine microbial communities, we performed protein mass spectrometry (Fig. S1) at 14 sample sites (Table S1) in the Atlantic Ocean (Fig. S2). Additionally to protein annotations retrieved by homology searches against genomic databases (*Supporting Information* and

Fig. S3), functional classifications were inferred using (*i*) clusters of orthologous groups (COGs) of proteins and (*ii*) the Kyoto Encyclopedia of Genes and Genomes (KEGG) database for higher-order cellular processes (Table 1). A combined set of 1,002 nonredundant transport-related membrane proteins (29% of total proteins identified; Datasets S1 and S2), subsequently referred to as TMPs, was identified. Approximately 71% of the TMPs were identified in at least two depth layers, whereas only 7% (75 proteins) were detected in all metaproteomes (Fig. 1*A*). COG categories, defined for the transporter protein repertoire, resulted in a conserved set of 96 COG families (Dataset S3), subsequently referred to as transporter COGs. The comparison between the major water layers, namely the euphotic, the mesopelagic (300–850 m depth), and upper (1,500–2,000 m depth) and lower (2,750–4,050 m) bathypelagic zones, enabled us to identify 29 COG families (~30%) common to all metaproteomes (i.e., involved in amino acid, inorganic ion, and carbohydrate transport; Fig. 1*B*). Among the 31 COG families present exclusively in the bathypelagic layer, we identified transporter COGs from the categories of "Carbohydrate transport and metabolism" (G; i.e., organic acids, xylose, maltose), "Inorganic ion transport and metabolism" (P), and "Poorly characterized" (R). The high number of unique transporter COGs recovered from deep ocean likely reflects the overall higher number of proteins recovered from these samples. Noteworthy, despite the greater diversity of differential transporter COGs, no additional substrate specificities were predicted from the protein-coding sequences. This observation points at the importance of suitable and "complete" genomic databases and to accurately annotate protein-coding sequences. These results highlight also the potentially broad substrate range encounter by marine microbes, which needs yet to be defined. To account for differences in protein recovery between the samples, we applied semiquantitative analyses of transporter COGs, introducing normalized area abundance factors (NAAFs) (*Supporting Information*). Resulting expression profiles of transporter COGs indicate that functionally similar transport processes are present throughout the water column yet differ in their relative abundances between sampling sites (Fig. 2).

Vertical DOM Uptake Patterns Based on Transporter Proteins. Similar to previous metaproteomic studies (14, 21, 24, 25, 27, 28), ABC transporters (662 proteins, 20% of total identified proteins)

Table 1. Transport-related membrane protein statistics summarizing unique COG and KO classifiers

Layer	Depth, m	Sample ID	No.	%	COG, %	KO, %	Unique TMP, no.	Average TMP, %	Moca, Geotraces, %	Malaspina, %	SAG, %
			TMP								
EUPH	100	MED1_13	55	15	66	82	28	23	19	10	71
		GEO_5	63	25	77	63	16		24	26	49
		MOC_7	132	21	70	85	214		13	11	76
		MOC_6	150	31	72	73	58		26	18	56
MESO	300	MOC_5	87	26	71	59	7	32	23	47	30
	500	GEO_6	59	20	76	76	57		18	17	65
	850	MED1_16	169	51	81	70	21		25	32	43
BATHY	1,475	MED2_24	418	40	80	64	116	39	30	33	37
	1,750	MED2_12	594	38	83	68	68		34	31	35
	2,000	MED2_16	442	39	82	69	36		30	31	39
	2,750	MED2_8	472	40	83	65	2		35	36	29
	2,800	MED2_17	675	38	83	67	104		34	35	31
	3,200	MED2_5	227	44	79	63	10		29	43	28
	4,050	MED1_24	429	35	83	62	162		26	54	20

Detailed information on the reference databases is provided for all depth layers considered in this study. Number (no.) indicates the number of protein sequences identified, and percentages were calculated for the entire metaproteomic data/sample. COG, cluster of orthologous groups; EUPH, lower euphotic layer; KO, KEGG orthology; SAG, single amplified genome; TMP, transport-related membrane protein.
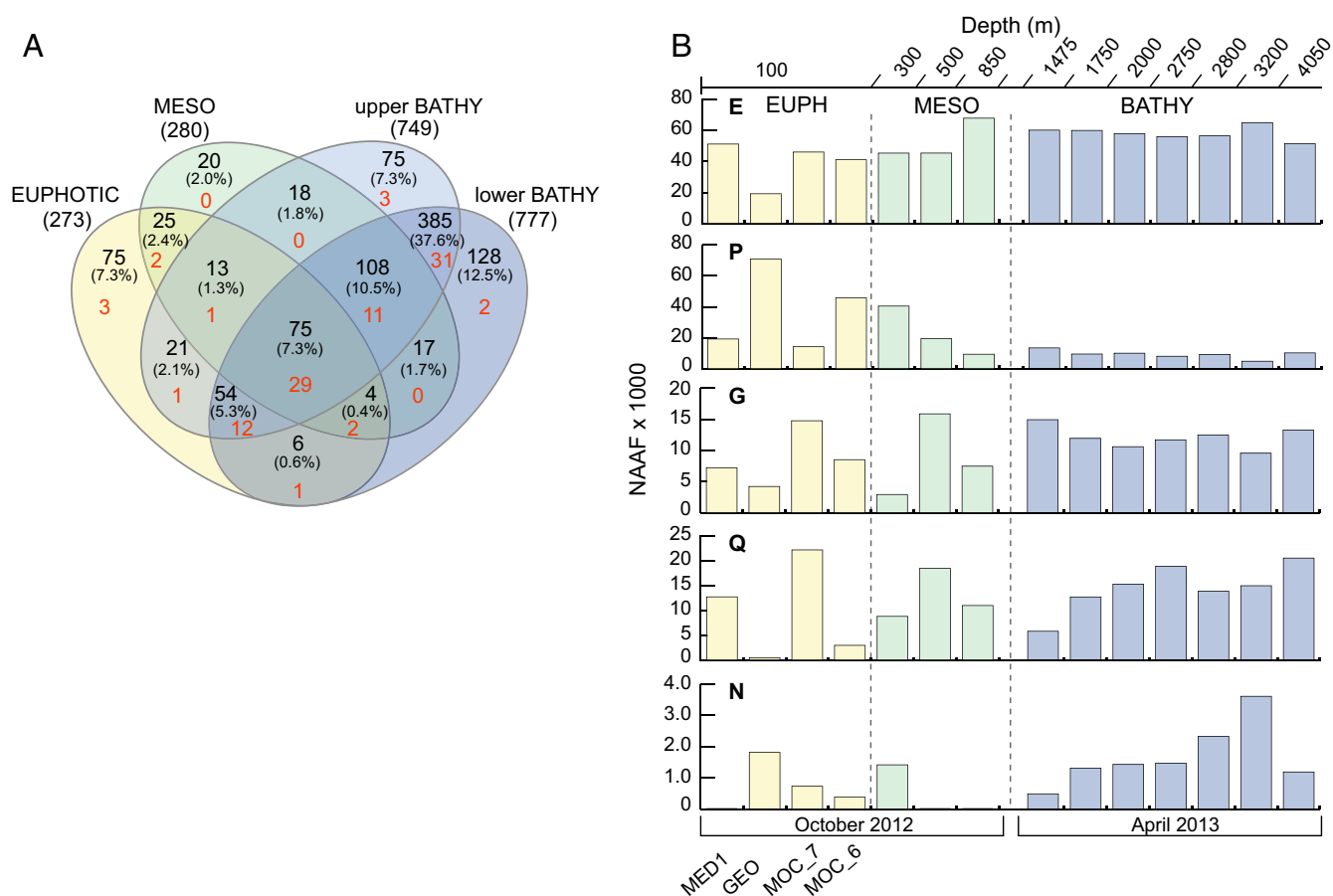
ECOLOGY

Fig. 1. Mascot and SEQUEST-HT search results were combined to create nonredundant lists of protein groups, and shared COGs as well as differences in transporter abundances between the samples are shown. (*A*) Venn diagram illustrating the number of COG families (red) shared between the 14 meta-proteomes, grouped by the distinct water layers: EUPHotic (lower euphotic, 100 m), MESO (mesopelagic, 300–850 m), upper BATHY (bathypelagic, 1,475–1,973 m), and lower BATHY (2,750–4,050 m). Proteins not grouped into COGs are indicated in black letters. (*B*) Distribution and abundance of selected COG functional categories associated with transport functions and cell motility. COG categories: E, amino acid transport and metabolism; P, inorganic ion transport and metabolism; G, carbohydrate transport and metabolism; N, cell motility; Q, secondary metabolites biosynthesis, transport, and catabolism. Relative protein abundances are based on NAAF values.

comprised the most prevalent transporter system in our study. As reported from coastal surface waters off the Antarctic Peninsula (14), the SBPs were the most frequently encountered components of the ABC complex, whereas transmembrane domains were less often identified (4%; Dataset S1). Roughly 85% of ABC transporters were predicted periplasmic SBPs and 11% remained unassigned. TRAP transporters were among the 10 most expressed proteins in the metaproteomic repertoire (86 proteins; 9% of TMP and 2.5% of total identified proteins) and mostly comprised SBPs homologous to the DctP family. Only three SBPs were of the TRAP-associated extracytoplasmic immunity (TAXI) family (29, 30) (Dataset S1). Together with TRAP transporters (31), TTTs are considered as secondary active transporters found in Bacteria and Archaea that employ SBPs to capture their ligands (29, 30, 32, 33). TTTs accounted for the smallest fraction of prokaryotic transport systems (29 proteins; 3% of TMP and 1% of total identified proteins; Fig. 2), which might be explained by their narrow substrate range we know of until now (31, 34, 35), or due to the limited number of reference sequences available in public databases.

The third most abundant transporter system was the outer-membrane TBDT, which plays an important role in high–molecular-weight OM uptake (22). Metaproteomic (21), metagenomic (26), and metatranscriptomic (36) surveys indicate the ocean-wide (coastal to open ocean) presence of TBDTs in surface oceans

and in deep-sea hydrothermal vents. In this study, we provide additional information on the vertical distribution and expression levels of TBDTs (107 proteins) in pelagic waters, accounting for 11% of TMP and 3% of total identified proteins.

The semiquantitative assessment of these transporter systems further confirmed the omnipresence and wealth of ABC transporters in the euphotic zone and the dark ocean (Fig. 2). Relative abundances of expressed ABC transporter proteins were on average six times higher than TRAP-Ts in the euphotic zone and 22 times higher than TBDTs in bathypelagic metaproteomes. TRAP-Ts and TTTs exhibited similar vertical expression patterns, with higher average abundances in the mesopelagic and bathypelagic zones, whereas TBDTs were expressed at higher values in the upper water column.

With the exception of TTT systems, transporter proteins mediating the uptake of both low– and high–molecular-weight OM were detected at all depths. However, relative expression levels of transporter systems varied with depths, suggesting depth-related quantitative changes in substrate uptake patterns while qualitative changes in the transporter systems were not apparent (Fig. S4). Importantly, the fraction of transporters increased from ~23% in the euphotic layers, to 32% in the mesopelagic and 39% in bathypelagic waters (Table 1). The increase in the fraction of transporter proteins from the euphotic to the bathypelagic zone as observed in this study (Table 1) might be interpreted as an
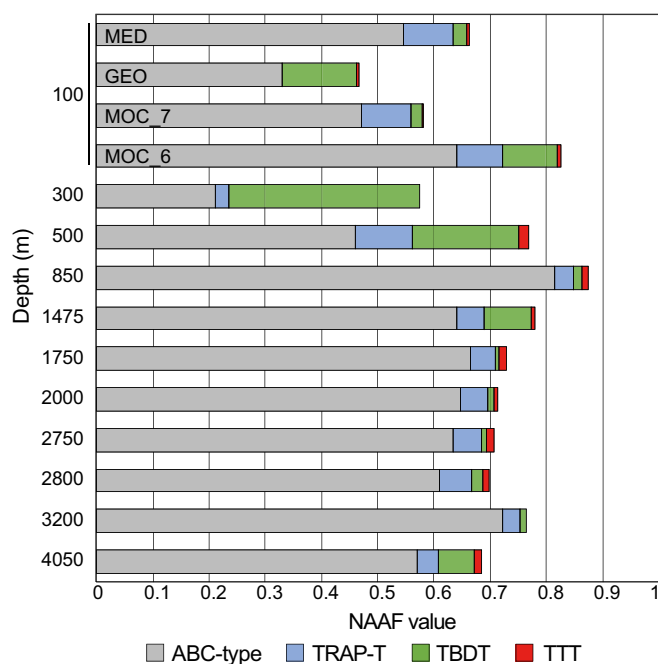
Fig. 2. Vertical distribution and relative abundance (NAAF) of ATP-binding cassette (ABC), tripartite ATP-independent (TRAP-T), tripartite tricarboxylate (TTT), and TonB-dependent (TBDT) transporters. The sample IDs for meta-proteomes from the euphotic zone are indicated; the coordinates and the physical and chemical characteristics of the samples are given in Table S1.

adaptation of the deep-sea microbial community to the changes in quality and quantity of the OM in the water column. Prokaryotic abundances and leucine incorporation rates, used as a proxy for heterotrophic prokaryotic production, typically decrease from the sunlit to the abyssopelagic waters by roughly three orders of magnitude (37). Cell-specific leucine incorporation rates, however, analyzed along the cruise track of Geotraces and Medea (*Supporting Information*) remained fairly constant in bathypelagic and abyssopelagic waters (1,000–5,500 m; Table S2), with average rates of $2.1 \pm 2.2 \times 10^{-5}$ fmol Leu·cell^{-1}·d^{-1}, indicating a fairly constant heterotrophic activity below the mesopelagic zone.

The Substrate-Active Community. The taxonomic diversity derived from expressed transporter proteins, referred to as the substrate-active community, comprised mainly Bacteria (69%) and 2% Archaea (Fig. 3). Roughly 30% of TMPs (294 proteins) remained unclassified due to divergent assignments at the phylum level or insignificant hits (Dataset S1). Expression profiles of the substrate-active community revealed the following contribution of phyla: Proteobacteria (Alphaproteobacteria, 40% of NAAF values; Gammaproteobacteria, 10%; Deltaproteobacteria, 7%), Actinobacteria (5%), Bacteroidetes (5%), Thaumarchaeota (4%), Cyanobacteria (3%), and Euryarchaeota and Planctomycetes (each 1%). A depth-dependent stratification of substrate-responsive phyla, based on the expression of transporter-related proteins, was evident for Cyanobacteria, Actinobacteria, Bacteroidetes, Verrucomicrobia, Firmicutes, and the candidate phylum Marinimicrobia, accounting for higher relative abundances in euphotic and mesopelagic layers than in bathypelagic waters (Fig. 3). Conversely, transporters from Deferribacterales, Planctomycetes, Chloroflexi, and members of the Gammaproteobacteria and Deltaproteobacteria were relatively more abundant in the bathy-pelagic than in water masses above.

Concomitant metagenomes recovered between 1,809 and 65,228 predicted protein-coding genes (Table S3) and 32–438 near–full-length 16S rRNA gene sequences (*Supporting Information* and

Table S4). The recovery of 16S rRNA genes from metagenomic libraries supports the findings on the diversity deduced from 16S rRNA gene PCR amplifications. Vertical distribution patterns based on variations in estimated 16S rRNA gene abundances revealed similar trends in the microbial community structure (*Supporting Information* and Fig. S5). Noteworthy, however, estimated 16S rRNA gene abundances indicated higher relative abundances of Thaumarchaeota throughout the water column than suggested by transporter protein abundances, contributing up to ~20% of the microbial community in bathypelagic waters (Fig. S5). Thus, transporter proteins primarily selected as a proxy for transport and degradation of complex organic molecules, might underestimate the activity and abundance of marine Archaea and the relative contribution of chemolithoautotrophic metabolism versus heterotrophy. The genetic capabilities of Thaumarchaeota for the uptake or translocation of compounds are described in detail elsewhere (38). In our metaproteomes, predicted thau-marchaeal transporters included ammonium channel proteins
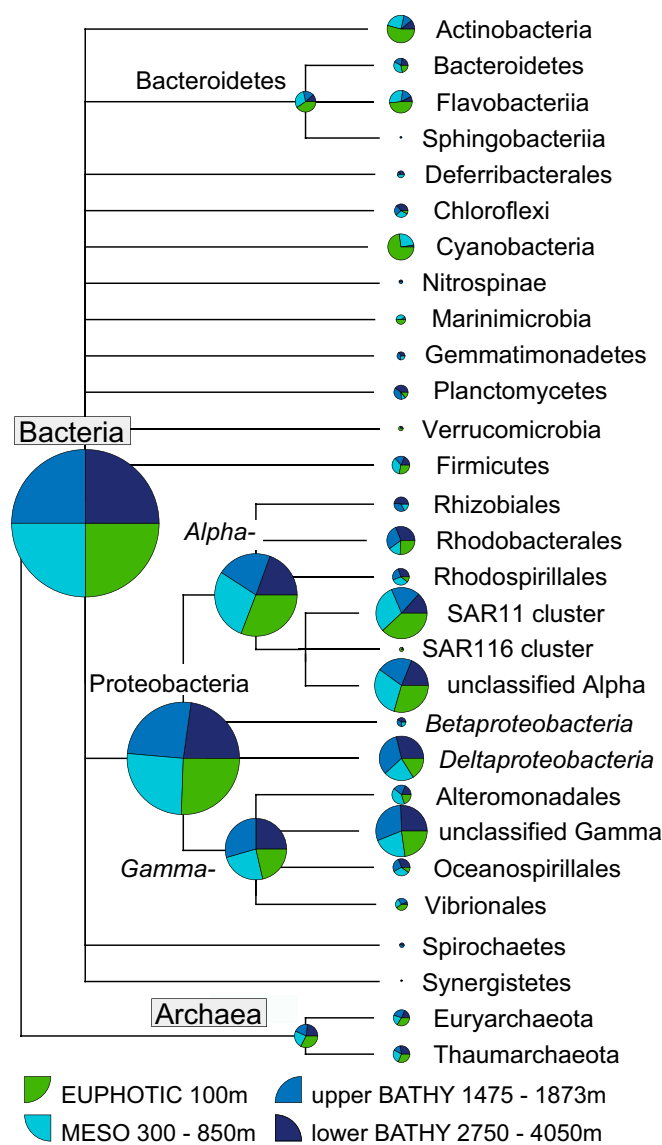

Fig. 3. Taxonomic distribution of *Bacteria* and *Archaea* assigned to transport-related membrane proteins (TMPs). Pie charts represent semi-quantitative abundance estimates based on averaged NAAF values for all TMPs at the designated water layers (see color key).

(Amt), urea, and SSS family transporters revealing, despite the comparatively low expression levels, the potential of the organism to utilize inorganic and organic energy sources (39).

Substrate Uptake Patterns at the Community Level. COG and KEGG Orthology (KO) classifiers were used to resolve the compound specificity of transport proteins (Fig. 4), and for prevalent microbial groups, average NAAFs, referred to as "protein abundance" from here on, were summarized for the euphotic, the mesopelagic, and the upper and lower bathypelagic zones (Fig. 5).

Predicted substrate affinities of expressed transporter proteins suggest that amino acids, including branched-chain amino acids, proline/glycine betaine, and di- and oligopeptides, as well as carbohydrates and carboxylic acids, might represent essential components of the OM pool in the oligotrophic open ocean.

ABC transporters exhibited the highest expression values in every depth layer attributed to their involvement in the uptake of amino acids (245 proteins; Fig. 4). Predicted ligands of the respective SBPs included proteinogenic amino acids such as glutamate, arginine and derivatives, histidine, proline, and glycine or more general branched-chain amino acid (leucine, isoleucine, and valine) and polar- and L-amino acids (Fig. S6 and Dataset S1). Various bacterial phyla and Euryarchaeota Marine Group II (MG-II) expressed amino acid transporters, albeit the proportions of active cells varied among groups (Dataset S1). In agreement with actual measurements of amino acid uptake rates (40), Alphaproteobacteria and Gammaproteobacteria accounted for high relative abundances of amino acid transporters, together

contributing ~30% to the expression values in the lower euphotic, 83% in the mesopelagic, and 85% in the bathypelagic zone (Fig. 5). As previously described based on transcriptome data (41), members of the SAR11 clade invest substantially in the uptake of compatible solutes, for example, proline/glycine betaine (ProU operon), reportedly supporting cell growth (42). In sinking POM, amino acid-like material accounts for the largest component (40–50%) of particulate organic carbon (43). Interestingly, minimal changes in the chemical structure of the bulk POM composition were found throughout the water column (43). Together with amino acids and lipids, carbohydrates constitute a major group of biomolecules in the OM pool and comprise up to 30% (by mass) of sinking particles (43). We identified 109 proteins of ABC carbohydrate-binding transporters: (*i*) monosaccharide and (*ii*) disaccharide and oligosaccharide transporters, which are homologous to the family of oligopeptide and dipeptide transporters. Carbohydrate transporter abundances accounted for similarly high expression values as observed for peptides with only minor variations in relative abundances throughout the water column (Fig. 4). Predicted COG/KO functions indicate a broad range of potential target molecules (Fig. S6), with a remarkably high fraction of transporters involved in the uptake of glycerol-3-phosphate (G3P) (Dataset S4). Expression profiles of the carbohydrate-active community indicated a vertical stratification in the water column, with higher relative abundances of Actinobacteria, Firmicutes, and Gammaproteobacteria in the upper water column, whereas Deltaproteobacteria dominated in underlying waters (Fig. 5). SAR324 clade members showed high expression values of
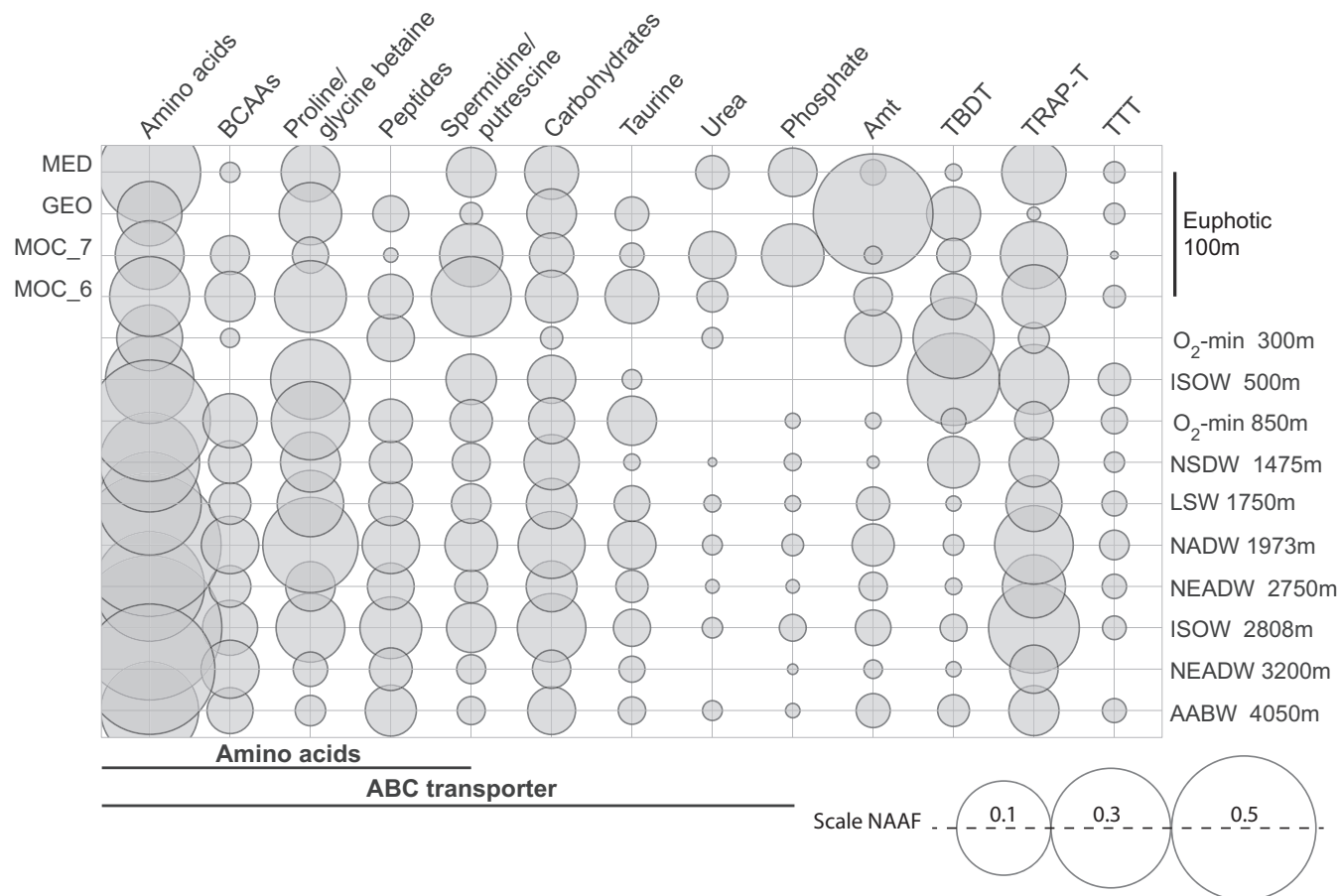

Fig. 4. Vertical expression profiles of selected transporters analyzed in a semiquantitative manner based on NAAF values. Transporter proteins were grouped by the predicted substrate specificity of the SBP. BCAAs, branched-chain amino acids.
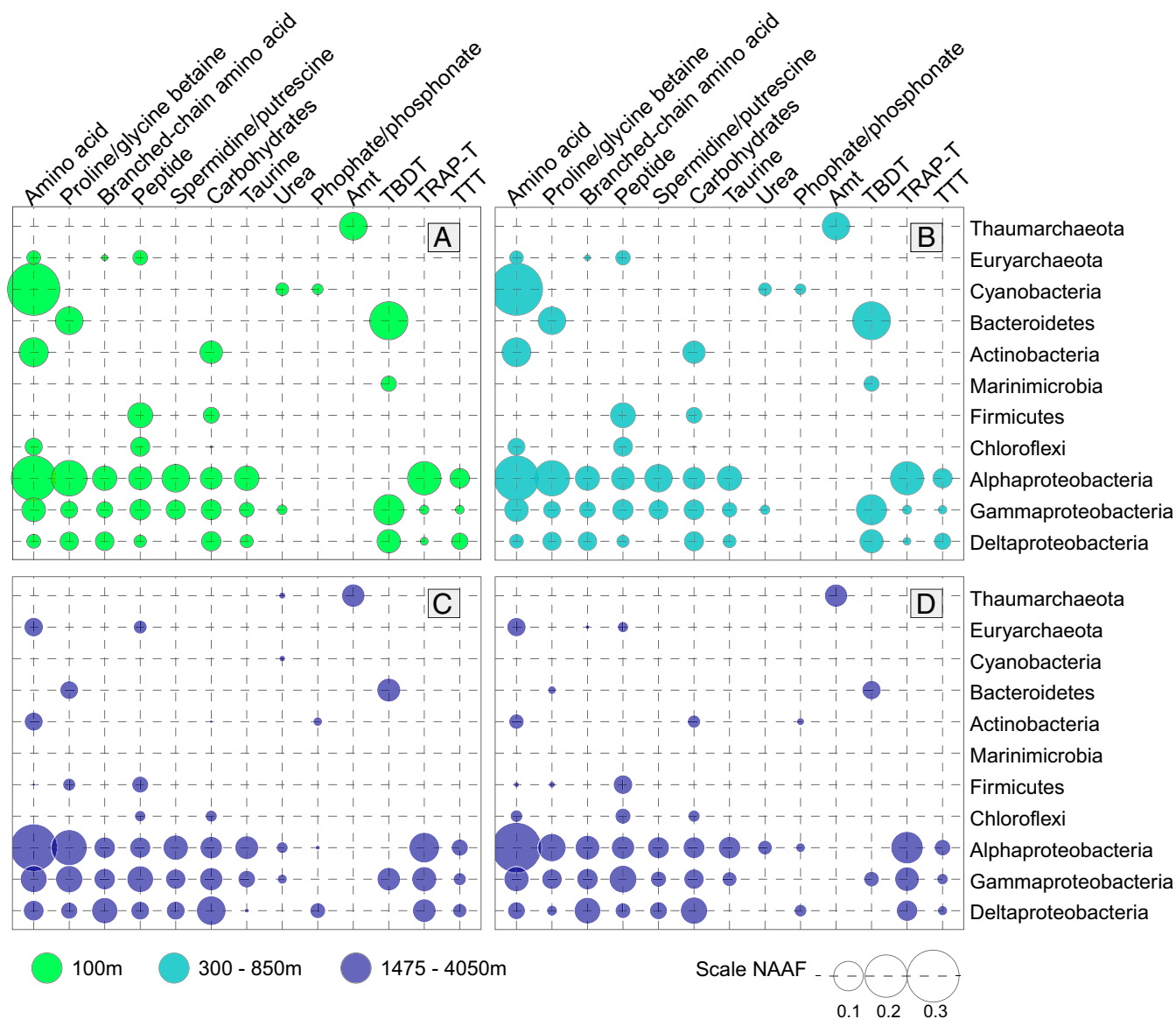
Fig. 5. Vertical expression profiles of transporter proteins of abundant taxa. Expression values were calculated in a semiquantitative manner and average abundances were plotted for selected members of the substrate active community residing in the (*A*) lower euphotic, (*B*) mesopelagic, and (*C*) upper and (*D*) lower bathypelagic water layers.

transporters specific to G3P, which might be utilized as a source of carbon or inorganic phosphate under nutrient-limiting conditions (44, 45). The expression of indicator genes for flagellar-related assembly (COG1344; Dataset S1) may indicate that at least some members of the SAR324 cluster possess a strong gliding or adhesion capability (46), suggesting the tendency to associate with particles (47).

Another omnipresent and abundantly expressed compound class was ABC peptide transporters, mediating the bulk uptake of dipeptides and oligopeptides (Dpp and Opp systems). We identified 115 transporter proteins involved in peptide utilization (Dataset S1). Independent of the sampling depth, relative abundances of dipeptide transporters were twice as high as those for oligopeptides (Fig. S6). While their primary function is the import of peptides as carbon and energy source, their role as messengers in virulence or signaling processes, for example, modulating chemotactic behavior or peptide export for biofilm formation, has to be taken into consideration. Thereby, peptide transporters ex-

emplify that it is inevitable to perform biochemical analyses to verify substrate specificities and/or functions of the predicted transporter components. Various taxa demonstrated protein-processing capabilities, with the MG-II/III Euryarchaeota, Planctomycetes, and Gammaproteobacteria being the prime protein utilizers (Fig. 5). This is in agreement with high transcript levels of extracellular peptidases and carbohydrate-active enzymes (CAZymes) of MG-II and MG-III Euryarchaeota (48), suggesting the metabolic capacity of extracellular protein degradation and utilization of carbohydrates for heterotrophic growth (Table S5).

Another ubiquitous source of labile carbon and nitrogen in the ocean constitute polyamines like putrescine and spermidine. The occurrence of polyamine transporters has been studied in marine bacterial genomes (49), revealing the wide distribution of SBPs among the phyla of Actinobacteria, Chlamydiae/Verrucomicrobia, Cyanobacteria, Firmicutes, and Proteobacteria. Surveys of marine metagenomes revealed that up to 32% of surface ocean

bacterioplankton potentially encode homologs implicated in the transport or degradation of polyamines, suggesting the important role of polyamines in carbon and nitrogen cycling (50). In our study, the active polyamine-transforming community comprised Alphaproteobacteria (25 proteins; including Rhodobacterales, Rhizobiales, and Pelagibacterales), Gammaproteobacteria (14 proteins), and Deltaproteobacteria (7 proteins), with SAR11 dominating in the euphotic zone (Fig. 5). Overall, relative abundances of polyamine transporter systems (PotD, PotF; 56 proteins) suggest a higher expression in the euphotic zones than underlying water masses (Fig. 4).

ABC transporters expressed at considerably lower abundances had predicted substrate specificities to urea, phosphate/phosphonate, or taurine (2-aminoethane sulfonate), sources of organic carbon, nitrogen, and/or sulfur moieties (Fig. 4). Taurine, in particular, constitutes a valuable food source for heterotrophic microbes due to its carbon, nitrogen, and sulfur moieties and is found at nanomolar concentrations in the marine environment (51). Key enzymes required to utilize these nutritional elements (52, 53) include the taurine-pyruvate aminotransferase (Tpa), alanine dehydrogenase (Ald), and sulfoacetaldehyde acetyltransferase (Xsc), facilitating the utilization of taurine as C source (52, 53). Thirteen proteins encoding the SBP (TauA) of the taurine transport system (TauABC) were detected in our metaproteomes at higher average abundances in the euphotic zone than in deeper layers (Fig. 4; NAAFs ranging between 0.003 and 0.06). Proteins encoding the Xsc (Dataset S1) were recovered at 500 m and 1,475- to 2,808-m depth, suggesting that taurine might serve as a favorable C source also in mesopelagic and bathypelagic waters. Function predictions of several proteobacterial single amplified genomes revealed genes or homologs for the TauD and/or Tpa, suggesting the genetic potential of C and S utilization. Our metagenomes clearly indicated an overall increase of TauD gene abundances below the euphotic zone, pointing at the importance of taurine as S source in the dark ocean (Dataset S5). Within the Alphaproteobacteria, SAR11 cluster members appeared as the prime utilizers of taurine in bathypelagic waters (Fig. 5). This is in agreement with a previous study, experimentally demonstrating the growth of SAR11 on taurine (54) and the ubiquity of SAR11 taurine transporters in metaproteomic and metatranscriptomic surveys in coastal surface waters (14, 28, 41).

Additionally to ABC transporters, diverse members of the microbial community expressed TRAP transporters at relatively high abundances, particularly in bathypelagic layers (Fig. 5). TRAP transporters are ATP independent, and thus less energy consuming than ABC transporters and, consequently, more advantageous under oligotrophic, deep-sea conditions. Known substrates recognized by the SBPs include a wide variety of compounds, all characterized by the presence of a carboxylate group, that is, organic acids (amino acids and acid sugars) and other carboxylate-containing small metabolites (29, 30, 32, 33, 55). Besides nutritional benefits, TRAP transporters support the thermostabilization of microbial cells by translocating osmoprotectants such as ectoine and 5-hydroxyectone (56). Predicted substrate affinities of TRAP SBPs identified in this study included mannitol/chloroaromatic compounds and aromatic acids (35) (COG4663, FcbT1; 45 proteins) and C4-dicarboxylates such as malate, fumarate, and succinate used as carbon and energy sources (COG1638, DctP; 36 proteins). Collectively, the phylogenetically diverse expression of these transporters and their ubiquity may indicate their ecological importance in dark ocean substrate turnover.

High proportional abundances of outer membrane TBDTs (107 proteins) were observed in euphotic and mesopelagic layers (NAAF euphotic, 0.0747 ± 0.047; meso, 0.1778 ± 0.117; bathy, 0.03 ± 0.03), similar to environmental metatranscriptomic (57) and metaproteomic (21) studies. TonB transporters were originally identified in the context of iron transport but are now reported to be involved in the uptake of a variety of compounds (58).

TonB-dependent transporter proteins encoded by cobalamine or vitamin B12 receptors (COG4206, BtuB), Ton box of ferric citrate (COG4772, FecA), outer membrane Fe transporters (COG1629, CirA), and outer membrane receptors for ferrienterochelin and colicins (COG4771, FepA). Also, we identified eight proteins encoding the SusC/RagA clade of TonB-linked outer membrane proteins, presumably utilizing large protein fragments (i.e., RagA) or carbohydrates (i.e., SusC) as organic compounds (Dataset S1). TonB-related proteins were also identified as among the most abundant transcripts assigned to DOM-responsive Idiomarinaceae and Alteromonadacea in a DOM-enriched marine microcosm (22). Both FepA and FecA facilitate the uptake of iron complexed to OM in diverse marine bacteria (59) in addition to other transport capacities. Fe is a critical trace element for bacterioplankton and impacts carbon and nitrogen fixation over broad regions of the ocean (60). However, to our knowledge, Fe(II) oxidation has not been observed in the open water column. In agreement with previous reports (26, 61), phylogenetic analyses of TBDTs strongly indicated that Gammaproteobacteria (34 proteins) and Bacteroidetes (including Cytophaga, Flavobacteria, and Sphingobacteria; 19 proteins) degrade polymeric matter throughout the water column. The taxonomic classification and relative abundance of TBDTs throughout the water column reveal a clear stratification of Flavobacteria, Alteromonadales, SAR86 cluster bacteria, and Marinimicrobia, dominating in the upper ocean. Conversely, TBDTs of Deferribacteres and Gemmatimonadetes were expressed at higher abundances in the dark ocean (Dataset S1). Taken together, these observations further highlight the significance of transporter processes throughout the water column and likely illustrate depth-related partitioning of marine bacteria according to environmental conditions.

Conclusion

Understanding the biogeochemistry of the OM pool remains a central challenge in microbial ecology (23, 62, 63), and thus, the analyses of expressed transporters in the open water column provide fundamental insights on this important topic. In concert, our data suggest that despite the commonly observed decreasing concentrations of OM with depth, transport proteins are expressed at high levels accounting for up to 39% of assigned COGs in bathypelagic zones (Table 1). Our study revealed expression patterns of prevalent transporter systems targeting various substrates, with no indications of major changes in the substrate affinity of transporter systems. The relative quantity of individual transporter systems, however, changed with depth such as TRAP transporters being particularly abundant among the bathypelagic microbial community (Fig. 4). Thus, while the phylogenetic composition of the microbial community is depth- and water mass-specific (Fig. S5), the composition of transporter proteins, indicative of the substrate quality, changes only marginally with depth. Using transporter components as proxy for OM translocation and utilization, we hypothesize that low substrate concentrations might be compensated by either modifying the abundance of the corresponding transporters by the individual microbes or by a shift in the microbial community. Despite multiple challenges imposed by "omics" tools, for example, peptide recovery and annotation accuracy (64, 65), or misleading annotations for protein families with diverse functions, transport proteins currently give the most subtle clue to assess the organic and inorganic compounds actually being used by microbes in the deep ocean. Overall, the distribution of expressed transporter proteins from the subsurface to the bathypelagic waters lends support to the major role of POM solubilization as a carbon and energy source for deep-ocean microbes rather than direct DOM utilization.

Materials and Methods

Sampling and Metadata Collection. Sampling was performed during the research cruises Geotraces III (March 2010), Moca (October 2010), and Medea I and II (October 2011, June/July 2012), spanning a latitudinal range in the Atlantic Ocean from 67°N to 49°S (Fig. S2). To investigate the vertical structure and activity of microbial communities, water samples (100–600 L) for metagenomic (100–5,002 m) and metaproteomic (100–4,050 m) analyses were collected from eight distinct water masses and 14 depth layers (Table S1).

Metagenomics. Preparation of genomic DNA was performed using a standard phenol extraction protocol (*Supporting Information*). Contigs were assembled using paired-end Illumina HiSeq reads, and the assembly and prediction of ORFs were performed using the software tools MetaVelvet and Prokka (Fig. S1). 16S rRNA gene sequences were assembled using EMIRGE (66) and identified by BLASTn searching against the Silva SSU 128 database (Fig. S5).

Metaproteomics. Protein extraction and spin filter-aided in-solution digestion (SF-ISD) was performed according to León et al. (67) with modifications to optimize protein recovery (*Supporting Information*). Equal peptide aliquots generated from each sample were analyzed by ultrahigh-pressure nanoLC coupled to an LTQ-Orbitrap Velos Mass Spectrometer (Thermo Scientific). Peptide separation was performed with a nanoAcquity UPLC system (Waters) fitted with a 2-cm, 180-μm ID Symmetry C18 trap column (Waters) and a 25-cm, 75-μm ID, BEH C18 (1.7-μm particles) analytical column (Waters). Peptides were trapped for 2 min at 10 μL/min with 0.1% TFA and separated at 350 nL/min using a gradient of 5–35% acetonitrile with 0.1% formic acid for 75 min.

Detailed information on the experimental procedures, functional assignments, and quantification methods can be found in *SI Materials and Methods*.

1. Nagata T, et al. (2010) Emerging concepts on microbial processes in the bathypelagic ocean—ecology, biogeochemistry, and genomics. *Deep Sea Res Part II Top Stud Oceanogr* 57:1519–1536.
2. Aristegui J, Gasol JM, Duarte CM, Herndl GJ (2009) Microbial oceanography of the dark ocean's pelagic realm. *Limnol Oceanogr* 54:1501–1529.
3. Buesseler KO, Lampitt RS (2008) Introduction to "Understanding the ocean's biological pump: Results from VERTIGO"—preface. *Deep Sea Res Part II Top Stud Oceanogr* 55:1519–1521.
4. Ducklow HW, Steinberg DK, Buesseler KO (2001) Upper ocean carbon export and the biological pump. *Oceanography* 14:50–58.
5. Delong EF, Franks DG, Alldredge AL (1993) Phylogenetic diversity of aggregate-attached vs free-living marine bacterial assemblages. *Limnol Oceanogr* 38:924–934.
6. Herndl GJ, Reinthaler T (2013) Microbial control of the dark end of the biological pump. *Nat Geosci* 6:718–724.
7. Arístegui J, et al. (2002) Dissolved organic carbon support of respiration in the dark ocean. *Science* 298:1967.
8. Arístegui J, Del Giorgio PA, Williams PJlB (2005) Respiration in the mesopelagic and bathypelagic zones of the ocean. *Respiration in Aquatic Ecosystems* (Oxford Univ Press, Oxford), pp 181–205.
9. Baltar F, Aristegui J, Gasol JM, Sintes E, Herndl GJ (2009) Evidence of prokaryotic metabolism on suspended particulate organic matter in the dark waters of the subtropical North Atlantic. *Limnol Oceanogr* 54:182–193.
10. Brophy JE, Carlson DJ (1989) Production of biologically refractory dissolved organic-carbon by natural seawater microbial-populations. *Deep-Sea Res* 36:497–507.
11. Jiao N, et al. (2010) Microbial production of recalcitrant dissolved organic matter: Long-term carbon storage in the global ocean. *Nat Rev Microbiol* 8:593–599.
12. Arrieta JM, et al. (2015) Ocean chemistry. Dilution limits dissolved organic carbon utilization in the deep ocean. *Science* 348:331–333.
13. Zark M, Christoffers J, Dittmar T (2017) Molecular properties of deep-sea dissolved organic matter are predictable by the central limit theorem: Evidence from tandem FT-ICR-MS. *Mar Chem* 191:9–15.
14. Williams TJ, et al. (2012) A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *ISME J* 6:1883–1900.
15. Hansell DA, Carlson CA, Repeta DJ, Schlitzer R (2009) Dissolved organic matter in the ocean: A controversy stimulates new insights. *Oceanography* 22:202–211.
16. Berntsson RP, Smits SH, Schmitt L, Slotboom DJ, Poolman B (2010) A structural classification of substrate-binding proteins. *FEBS Lett* 584:2606–2617.
17. Davidson AL, Dassa E, Orelle C, Chen J (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* 72:317–364.
18. Lewinson O, Livnat-Levanon N (2017) Mechanism of action of ABC importers: Conservation, divergence, and physiological adaptations. *J Mol Biol* 429:606–619.
19. Piepenbreier H, Fritz G, Gebhard S (2017) Transporters as information processors in bacterial signalling pathways. *Mol Microbiol* 104:1–15.
20. Saier MH, Jr (1998) Molecular phylogeny as a basis for the classification of transport proteins from bacteria, archaea and eukarya. *Adv Microb Physiol* 40:81–136.
21. Morris RM, et al. (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* 4:673–685.
22. McCarren J, et al. (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc Natl Acad Sci USA* 107:16420–16427.
23. Poretsky RS, Sun S, Mou X, Moran MA (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environ Microbiol* 12:616–627.
24. Sowell SM, et al. (2011) Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. *ISME J* 5:856–865.
25. Sowell SM, et al. (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* 3:93–105.
26. Tang K, Jiao N, Liu K, Zhang Y, Li S (2012) Distribution and functions of TonB-dependent transporters in marine bacteria and environments: Implications for dissolved organic matter utilization. *PLoS One* 7:e41204.
27. Colatriano D, et al. (2015) Metaproteomics of aquatic microbial communities in a deep and stratified estuary. *Proteomics* 15:3566–3579.
28. Georges AA, El-Swais H, Craig SE, Li WKW, Walsh DA (2014) Metaproteomic analysis of a winter to spring succession in coastal northwest Atlantic Ocean microbial plankton. *ISME J* 8:1301–1313.
29. Mulligan C, Fischer M, Thomas GH (2011) Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiol Rev* 35:68–86.
30. Kelly DJ, Thomas GH (2001) The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *FEMS Microbiol Rev* 25:405–424.
31. Winnen B, Hvorup RN, Saier MH, Jr (2003) The tripartite tricarboxylate transporter (TTT) family. *Res Microbiol* 154:457–465.
32. Fischer M, Zhang QY, Hubbard RE, Thomas GH (2010) Caught in a TRAP: Substrate-binding proteins in secondary transport. *Trends Microbiol* 18:471–478.
33. Rabus R, Jack DL, Kelly DJ, Saier MH, Jr (1999) TRAP transporters: An ancient family of extracytoplasmic solute-receptor-dependent secondary active transporters. *Microbiology* 145:3431–3445.
34. Antoine R, et al. (2005) The periplasmic binding protein of a tripartite tricarboxylate transporter is involved in signal transduction. *J Mol Biol* 351:799–809.
35. Hosaka M, et al. (2013) Novel tripartite aromatic acid transporter essential for terephthalate uptake in *Comamonas* sp. strain E6. *Appl Environ Microbiol* 79:6148–6155.
36. Li M, et al. (2014) Microbial iron uptake as a mechanism for dispersing iron from deep-sea hydrothermal vents. *Nat Commun* 5:3192.
37. De Corte D, Sintes E, Yokokawa T, Reinthaler T, Herndl GJ (2012) Links between viruses and prokaryotes throughout the water column along a North Atlantic latitudinal transect. *ISME J* 6:1566–1577.
38. Offre P, Kerou M, Spang A, Schleper C (2014) Variability of the transporter gene complement in ammonia-oxidizing archaea. *Trends Microbiol* 22:665–675.
39. Kerou M, et al. (2016) Proteomics and comparative genomics of *Nitrososphaera viennensis* reveal the core genome and adaptations of archaeal ammonia oxidizers. *Proc Natl Acad Sci USA* 113:E7937–E7946.
40. Alonso C, Pernthaler J (2006) Concentration-dependent patterns of leucine incorporation by coastal picoplankton. *Appl Environ Microbiol* 72:2141–2147.
41. Gifford SM, Sharma S, Booth M, Moran MA (2013) Expression patterns reveal niche diversification in a marine microbial assemblage. *ISME J* 7:281–298.
42. Tripp HJ, et al. (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* 452:741–744.
43. Hedges JI, et al. (2001) Evidence for non-selective preservation of organic matter in sinking marine particles. *Nature* 409:801–804.
44. Vetting MW, et al. (2015) Experimental strategies for functional annotation and metabolism discovery: Targeted screening of solute binding proteins and unbiased panning of metabolomes. *Biochemistry* 54:909–931.
45. Lackner G, Peters EE, Helfrich EJ, Piel J (2017) Insights into the lifestyle of uncultured bacterial natural product factories associated with marine sponges. *Proc Natl Acad Sci USA* 114:E347–E356.
46. Cao H, et al. (2016) Delta-proteobacterial SAR324 group in hydrothermal plumes on the South Mid-Atlantic Ridge. *Sci Rep* 6:22842.
47. Swan BK, et al. (2011) Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science* 333:1296–1300.
48. Li M, et al. (2015) Genomic and transcriptomic evidence for scavenging of diverse organic compounds by widespread deep-sea archaea. *Nat Commun* 6:8933.

Bergauer et al.

49. Mou XZ, Sun SL, Rayapati P, Moran MA (2010) Genes for transport and metabolism of spermidine in *Ruegeria pomeroyi* DSS-3 and other marine bacteria. *Aquat Microb Ecol* 58:311–321, and erratum (2010) 59:102.

50. Howard EC, Sun S, Biers EJ, Moran MA (2008) Abundant and diverse bacteria involved in DMSP degradation in marine surface waters. *Environ Microbiol* 10:2397–2410.

51. Clifford EL, et al. (2017) Crustacean zooplankton release copious amounts of dissolved organic matter as taurine in the ocean. *Limnol Oceanogr* 62:2745–2758.

52. Chien CC, Leadbetter ER, Godchaux W (1999) *Rhodococcus* spp. utilize taurine (2-aminoethanesulfonate) as sole source of carbon, energy, nitrogen and sulfur for aerobic respiratory growth. *FEMS Microbiol Lett* 176:333–337.

53. Cook AM, Denger K, Smits THM (2006) Dissimilation of C3-sulfonates. *Arch Microbiol* 185:83–90.

54. 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.

55. Mulligan C, Kelly DJ, Thomas GH (2007) Tripartite ATP-independent periplasmic transporters: Application of a relational database for genome-wide analysis of transporter gene frequency and organization. *J Mol Microbiol Biotechnol* 12:218–226.

56. Kuhlmann SI, Terwisscha van Scheltinga AC, Bienert R, Kunte HJ, Ziegler C (2008) 1.55 A structure of the ectoine binding protein TeaA of the osmoregulated TRAP-transporter TeaABC from *Halomonas elongata*. *Biochemistry* 47:9475–9485.

57. Ottesen EA, et al. (2011) Metatranscriptomic analysis of autonomously collected and preserved marine bacterioplankton. *ISME J* 5:1881–1895.

58. Schauer K, Rodionov DA, de Reuse H (2008) New substrates for TonB-dependent transport: Do we only see the "tip of the iceberg"? *Trends Biochem Sci* 33:330–338.

59. Chakraborty R, Storey E, van der Helm D (2007) Molecular mechanism of ferricsiderophore passage through the outer membrane receptor proteins of *Escherichia coli*. *Biometals* 20:263–274.

60. Martin JH, Gordon RM, Fitzwater SE (1990) Iron in Antarctic waters. *Nature* 345:156–158.

61. Dupont CL, et al. (2012) Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J* 6:1186–1199.

62. Benner R (2002) *Biogeochemistry of Marine Dissolved Organic Matter*, eds Hansell DA, Carlson CA (Academic, London), pp 59–90.

63. Kirchman DL, Suzuki Y, Garside C, Ducklow HW (1991) High turnover rates of dissolved organic-carbon during a spring phytoplankton bloom. *Nature* 352:612–614.

64. May DH, et al. (2016) An alignment-free "metapeptide" strategy for metaproteomic characterization of microbiome samples using shotgun metagenomic sequencing. *J Proteome Res* 15:2697–2705.

65. Timmins-Schiffman E, et al. (2017) Critical decisions in metaproteomics: Achieving high confidence protein annotations in a sea of unknowns. *ISME J* 11:309–314.

66. Miller CS, et al. (2013) Short-read assembly of full-length 16S amplicons reveals bacterial diversity in subsurface sediments. *PLoS One* 8:e56018.

67. León IR, Schwämmle V, Jensen ON, Sprenger RR (2013) Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol Cell Proteomics* 12:2992–3005.