Limnol. Oceanogr., 56(5), 2011, 1641–1652

doi:10.4319/lo.2011.56.5.1641

Metaproteomic characterization of dissolved organic matter in the water column of the South China Sea

Da-Zhi Wang,* Hong-Po Dong, Zhang-Xian Xie, Min-Han Dai, and Hua-Sheng Hong

State Key Laboratory of Marine Environmental Science/Environmental Science Research Center, Xiamen University, Xiamen, China

Abstract

We characterized dissolved organic matter (DOM; < 0.7 μm in size) collected from the surface (10-m and 75-m) and bathypelagic (3000-m) layers in the South China Sea using the shotgun proteomic approach. A total of 182 proteins matched by 286 unique peptides were identified from three DOM samples. The number of proteins in the large DOM (LDOM; 0.2–0.7-μm fraction) was significantly greater than that in the small DOM (SDOM; 5-kD–0.2-μm fraction). There was no remarkable difference in the number of proteins between the surface and bathypelagic SDOMs. The sources of dissolved proteins were diverse in surface DOM, including various bacterial and phytoplankton groups as well as Oomycetes, whereas the Archaea, Proteobacteria, and some phytoplankton groups were the major contributors to bathypelagic DOM. Proteins involved in cytoskeleton, energy production and conversion, posttranslational modification, protein turnover, and chaperones presented high abundance in surface LDOM, whereas proteins involved in translation, ribosomal structure, and biogenesis were more abundant in bathypelagic LDOM. Proteins involved in transport and metabolism, cell wall or membrane or envelope biogenesis, and photosynthesis were abundant in the 75-m LDOM. A urea ABC transporter assigned to amino acid transport and metabolism was the most abundant protein in the 10-m SDOM, whereas methylenetetrahydromethanopterin reductase involved in energy production and conversion dominated the protein profiles in the 75- and 3000-m SDOMs. The dissolved proteins in the water column are diverse and dynamic, with each layer characterized by unique proteins, and only a very minor amount of proteins from the surface are protected and transferred to the deep sea.

Dissolved organic matter (DOM) in seawater is one of the largest reservoirs of organic matter on earth (Siegenthaler and Sarmiento 1993), and plays important roles in the global carbon cycle (Ducklow 2002). The amount of carbon in DOM is approximately equivalent to that present in atmospheric CO_2, and even slight changes in the processes regulating the DOM pool can have consequences for the global carbon cycle and climate change. Moreover, the huge amount of DOM in the ocean plays various roles in ecological processes, depending on its value as a food source or as a stocking pool (Hansell and Carlson 2001). It is estimated that 20–40% of photosynthetically produced organic matter is utilized by bacteria as biologically available DOM and enters into biogeochemical recycling through the microbial loop (Carlson 2002), whereas much of the DOM is refractory and accumulates in seawater as a carbon pool (Jiao et al. 2010). However, a deeper understanding of the role of marine DOM in the global carbon cycle is necessary, particularly with respect to its origin, chemical characteristics, and biogeochemical cycling.

In the past few decades, our knowledge on DOM has been rapidly increasing. The bulk chemical composition and structure of DOM has been examined by a variety of approaches. A series of compounds such as peptidoglycan, lipopolysaccharide, D-enantiomers of amino acids, muramic acid, N-acetyl amino sugars, and fatty acid components have been successively identified from DOM (Dittmar et al. 2001; Benner and Kaiser 2003; Zou et al. 2004), and some of them have been used as potential biomarkers of DOM. However, DOM is still the least understood organic reservoir in terms of its source, its chemical nature, and the mechanism protecting it from biodegradation. To date, less than 30% of DOM has been characterized at the molecular level.

Proteins are the essential constituents of all organisms, making up more than 50% of organic matter as well as containing 85% of the organic nitrogen found in marine organisms (Tanoue 1995). The proteins in living organisms are eventually transferred to the dissolved pool via various biogeochemical processes. However, at present, both the chemical forms and the sources of these proteins and their degradation products (peptides) in DOM are unclear. Molecular-level analyses of dissolved proteins and peptides in DOM may provide valuable information concerning the dynamics of the bulk DOM pool in the ocean. Although dissolved proteins have been universally detected in oceanic and coastal waters (Yamada and Tanoue 2006, 2009), only a few dissolved proteins have been identified so far, such as porin P protein, OmpA-like protein, alkaline phosphatase, long-chain fatty acyl CoA synthetase, anthranilate synthase, ribulose bisphosphate carboxylase, and luminal binding protein (Tanoue et al. 1995; Powell et al. 2005; Yamada and Tanoue 2006). Obviously, it is insufficient to attempt to describe the features of bulk proteins in DOM and to reveal the biological origin and chemical nature of DOM solely based on a few characterized proteins.

Proteomics is one of the fastest developing research areas and contributes substantially to our understanding of organisms at the molecular level (Aebersold and Mann 2003). Recently considerable efforts have been devoted to extend proteomic analysis beyond the boundary of

* Corresponding author: dzwang@xmu.edu.cn

organisms. Tandem mass spectrometry (MS^2)-based proteomics is a new approach that enables the direct analysis of proteins expressed by mixed microbial assemblages, and it has been applied successfully to study various environmental samples such as acid-mine drainage, wastewater sludge, decomposing plant material, soil leachates, and surface-water samples (Ram et al. 2005; Lo et al. 2007; Wilmes et al. 2008), as well as the oceans (Kan et al. 2005; Sowell et al. 2009; Morris et al. 2010). More recently, by making use of the shotgun proteomic approach, thousands of proteins have been successfully identified from marine particulate organic matter (POM), and their biological source and subcellular location as well as the protection mechanism of these particulate proteins in different water layers have been examined (Dong et al. 2010). In this study, we applied the mass spectrometry–based metaproteomic approach to characterize dissolved proteins in DOM samples collected from different water layers in the South China Sea (SCS), in order to trace the biological origin and cellular location of these dissolved proteins, and to discuss the features of dissolved proteins in different water layers as well as their protection mechanism. To our knowledge, this study is the first attempt to characterize dissolved proteins of DOM at a metaproteomic level, and to compare the features of dissolved proteins from the sea surface to the bathypelagic zone. We found that each water layer displayed distinctive characteristics in terms of the protein sources, their cellular location, and their functional classification, and we revealed the discrepancy in terms of biological source and protection mechanism of the dissolved proteins among different water layers.

Methods

Sampling and preconcentration of dissolved proteins—Samples from different water depths were collected at Sta. SEATS (18.0°N, 115.6°E; water depth, 3850 m) during a summer cruise to the northern SCS from 17 July 2009 to 17 August 2009 (Fig. 1). Seawater samples were collected with a rosette sampler fitted with 12-liter Niskin Go-flo bottles (General Oceanics). The depths and seawater volumes were 10 m, 150 liters; 75 m, 210 liters; and 3000 m, 150 liters, respectively. Each sample was filtered through a GF/F glass fiber filter (with a nominal pore size of 0.7 μm, Whatman) immediately after sampling. Sodium dodecyl sulfate (SDS; 0.01% w : v, final concentration) was added to the filtrates to help to solubilize the proteins. The filtrates were further filtered through a 0.2-μm filter (Durapore membrane filter, Millipore) and sodium azide (NaN_3; 5 mmol L^{-1}, final concentration) was added to the filtrates to prevent bacterial growth. We defined in this study the fraction of 0.2–0.7 μm as a large-size fraction of DOM (LDOM), and the fraction of 5 kD–0.2 μm as a small-size fraction of DOM (SDOM), fractionated as described below.

All the filtrates < 0.2 μm were preconcentrated using the Pelicon 2 cross-flow ultrafiltration system (Millipore; 1 m^2 of filter area) with a filter having a nominal molecular weight cutoff of 1 kDa (low–protein-binding regenerated cellulose) onboard ship at room temperature. The system was precleaned with 0.1 mol L^{-1} NaOH and washed with deionized water according to the manufacturer's instruction. Before use of each new cross-flow ultrafiltration membrane and tubing, a 3.5% solution of NaCl (20 liters) was concentrated in the same manner as the seawater sample to serve as a control. The preconcentrated samples (approximately 600–700 mL) were kept frozen (−80°C) until further concentration using a 5-kD ultrafiltration cell, which is described below.

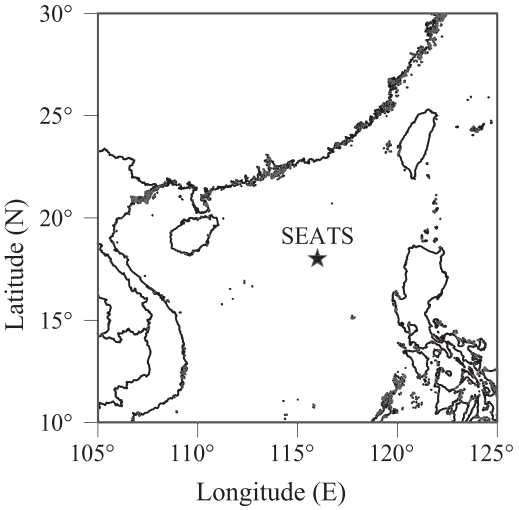

Fig. 1. Map of the SCS showing SEATS, the sampling location of DOM.

Fluorescein isothiocyanate–conjugated lactalbumin (14.5 kD) was used to evaluate the retention characteristics of the 1-kD cross-flow ultrafiltration membrane in the system. The applicability of the fluorescein standards for evaluating relative membrane retention and sorptive losses in seawater has been described previously (Dai et al. 1998). Seawater that had been run through a 0.2-μm prefilter was spiked with the fluorescein-conjugated standard. The spiked solution was then run through the Pelicon 2 cross-flow ultrafiltration system. The operating condition used in this procedure was the same as for the above DOM samples. The fluorescence was measured using a Hitachi F-1000 with the excitation wavelength set at 493 nm and a 520 nm cutoff emission filter. Source seawater without any spikes was used as background substitution. The results showed that the membrane owned more than 87% retention coefficient (RC% = (1 − [permeate]/[retentate]) × 100) and a 55–64% loss for the fluorescein-tagged lactalbumin protein.

Preparation of the dissolved proteins of SDOM—The preconcentrated samples from the cross-flow ultrafiltration were further concentrated and desalted with a 50-mL stirred ultrafiltration cell (Millipore) with a filter having a nominal molecular mass cutoff of 5 kDa (filter area, 13.4 cm^2; low–protein-binding regenerated cellulose, Millipore). When the sample was concentrated to approximately 25 mL, the same volume of the desalting buffer (35 mmol L^{-1} NH_4HCO_3, 0.01% SDS [w : v]) was added, and the mixture was concentrated continuously. This step was repeated three times to desalt the retentate. After retrieval

of the desalted retentate, the cell was washed using desalting buffer, and the final volume of the combined solutions was less than 30 mL. Then the combined solution was precipitated with ice-cold 20% trichloroacetic acid (TCA) in acetone for at least 12 h at −20°C. The mixture was centrifuged at 20,000 × *g* for 30 min at 4°C, and then the pellets were rinsed twice with ice-cold acetone and air dried. Finally the powder was solubilized in rehydration buffer containing 7 mol L^{-1} urea, 2 mol L^{-1} thiourea, and 3-((3-cholamidopropyl) dimethylammonium)-1-propanesulfonate (CHAPS; 4% w : v), and stored in a freezer (−80°C) until further SDS–polyacrylamide gel electrophoresis (PAGE) analysis.

Preparation of the dissolved proteins of LDOM—Protein extraction followed the method described previously (Dong et al. 2009). Briefly, the 0.2-μm filter (Durapore membrane filter) was cut into chips. The chips were suspended in lysis buffer consisting of urea, thiourea, CHAPS, Triton-X100, carrier ampholytes, dithiothreitol, and protease inhibitor cocktail, shaken in an ice-cold water bath using an ultrasonic shaker, and then lysed with a sonicator using a microprobe. The solution was incubated at 25°C for 1 h and then centrifuged at 20,000 × *g* for 30 min at 10°C. The supernatant was precipitated with ice-cold 20% TCA in acetone. The mixture was centrifuged, and the resultant pellets were rinsed twice with ice-cold acetone and air dried. Finally, the powder was dissolved in the above-mentioned rehydration buffer, and stored in a freezer (−80°C) until further SDS-PAGE analysis.

One-dimensional SDS-PAGE and enzymatic digestion of dissolved proteins—A rehydrated dissolved protein sample (40–60 μL), equivalent to ca. 150 liters of the original seawater, was applied to each lane of a 5–12% Bis–Tris gel (13 cm × 13 cm). Electrophoresis was performed at a constant voltage of 180 V in the electrode buffer solution (25 mmol L^{-1} Tris, 192 mmol L^{-1} glycine, and 0.1% SDS) on a Hoefer SE 600 apparatus (Amersham). Low-molecular-weight standards were used for reference. After electrophoresis, the proteins on the gel were visualized with colloidal Coomassie Brilliant Blue G-250 (CBB; Bio-Rad). The gel was cut into four equal pieces to isolate molecular weight (kDa) fractions, followed by reduction, alkylation, and in-gel digestion with trypsin as described previously (Wilm et al. 1996).

Mass spectrometry analysis—A nano-liquid chromatography (LC) MS^2 experiment was performed on a high-performance LC system composed of two LC-20AD nano-flow LC pumps, an SIL-20AC auto-sampler, and an LC-20AB micro-flow LC pump (Shimadzu) connected to a linear trap quadrupole–Orbitrap mass spectrometer (Thermo Fisher). The peptides extracted from the gel bands were resuspended with 5% acetonitrile in 0.1% formic acid and loaded onto a Captrap column (0.5 × 2 mm, Michrom Bioresources) for 5 min at a flow rate of 20 μL min^{-1}. The sample was subsequently separated by a C18 reverse-phase column (0.075 × 150 mm, packed with 3-μm Magic C18-AQ particles, Michrom Bioresources) at a flow rate of 500 nL min^{-1}. The mobile phases were 5% acetonitrile with 0.1% formic acid (phase A and the loading phase) and 95% acetonitrile with 0.1% formic acid (phase B). To achieve proper separation, a 60-min linear gradient from 5% to 45% phase B was employed. The separated sample was introduced into the mass spectrometer via an Advance 30-μm silica tip (Michrom Bioresources). The spray voltage was set at 1.8 kV and the heated capillary at 180°C. The mass spectrometer was operated in data-dependent mode and each cycle of duty consisted of one full-MS survey scan at the mass range 400 ~ 2000 Da with resolution power of 60,000 using the Orbitrap section, followed by MS^2 experiments for the 10 strongest peaks using the LTQ section. The automatic gain control expectation during full MS and MS^2 were 500,000 and 10,000, respectively. Peptides were fragmented in the LTQ section using collision-induced dissociation with helium and the normalized collision energy value was set at 35%. Previously fragmented peptides were excluded for 60 s.

Protein identification—A combined database was used for this study, which was downloaded on 14 January 2010. The database contained protein datasets from potential groups of phytoplankton and bacterioplankton according to the in situ planktonic community data at SEATS downloaded from the National Center for Biotechnology Information (NCBI) website, and the Moore Foundation marine microbial peptide dataset available through the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis (CAMERA) (http://web4.camera.calit2.net/files/) (Table 1). The NCBI database contained 1,616,403 protein entries, and the Moore Foundation marine microbial peptide dataset contained 293,532 sequences from over 155 diverse, ecologically relevant marine microorganisms. It should be noted that the protein database in this study did not include the viral sequences because of the lack of specific marine viral protein database as well as the limitation of the present MS^2 technology.

Tandem mass spectra were extracted using BioWorks version 3.3.1 sp1 (Thermo Fisher). All MS^2 spectra were searched against the above-mentioned combined database using the SEQUEST algorithm (version 28, Thermo Fisher). Data searches were completed, with the following parameters: enzyme type trypsin; parent ion tolerance, 50 ppm; fragment ion mass tolerance, 1.00 Da. Modifications of cysteine residues by 57.021 Da (resulting from the iodoacetamide modification) and methionine by 15.999 Da (oxidation) were allowed. The Trans-Proteomic Pipeline (version 4.1.1, Institute for Systems Biology) was used to validate MS^2-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm (Keller et al. 2002). Protein identifications were accepted if they could be established at greater than 99.0% probability. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al. 2003). Proteins that contained similar peptides and could not be differentiated based on MS^2 analysis alone were grouped to satisfy the principles of

Table 1. The combined database created in this study. The Moore Foundation marine microbial peptide dataset was downloaded from the CAMERA website, and the other datasets were downloaded from the NCBI website.

Protein database	No. of protein entries	Protein database	No. of protein entries
Pseudomonas	279,746	Archaea	441,830
Flavobacteria	143,813	Cyanobacteria	463,511
Erythrobacter	18,609	Cryptophyta	3494
Photobacterium	51,459	Chlorophyta	96,860
Coxiella	30,560	Haptophyceae	1006
Environmental samples of Alphaproteobacteria	350	Dinophyceae	2884
Environmental samples of Betaproteobacteria	1638	Stramenopiles	79,944
Environmental samples of Gammaproteobacteria	699	Moore Foundation marine microbial peptide dataset	293,532

parsimony. Owing to using this high-stringency cutoff and statistical algorithm, those proteins identified by one peptide were also accepted. In addition, in view of the capacity and complexity of the search database, it was not practical to calculate false discovery rates.

Bioinformatics analysis—Relative protein abundance was assessed at the peptide level using a semiquantitative method: peptide spectral counting (Ryu et al. 2008). Peptide spectral counts were determined by the number of times a peptide that correlates with a given protein was selected for collision-induced dissociation, including all repeated selections of the same peptide.

Redundant proteins in a group were removed according to the strategy of He et al. (2005), so that only one protein remained in a group. The method of tracing for biological origin of the dissolved proteins was as previously described (Dong et al. 2010). Briefly, the bacterial proteins were classified into six subgroups: Proteobacteria, Bacteroidetes, Planctomycetes, Firmicutes, Actinobacteria, and Cyanobacteria; and the proteins from eukaryotic algae were classified into seven subgroups: Prasinophyta, Dinophyta, Cryptophyta, Haptophyta, Chrysophyta, Bacillariophyta, and Chlorophyta. Other proteins, not belonging to the above groups, were grouped into Archaea, Oomycetes, and Stramenopiles. The functional categories of the proteins identified were performed using Clusters of Orthologous Groups of proteins (COGs) (http://www.ncbi.nlm.nih.gov/COG/) (Tatusov et al. 2000). Protein subcellular location was predicted using the Proteome Analyst–Subcell Specialization Server 2.5 (http://webdocs.cs.ualberta.ca/~bioinfo/PA/Sub/) (Lu et al. 2004).

Results

Separation and identification of dissolved proteins in DOM—An electrophotogram of the dissolved proteins from six samples separated using one-dimensional (1-D) SDS-PAGE is shown in Fig. 2. Lanes of all samples visualized using CBB stain showed a smeared staining pattern throughout the molecular mass range, and some faint bands were recognized on background staining. Generally, the staining intensity of the dissolved proteins of LDOM was higher than that of SDOM (Fig. 2). There were two distinct bands in each sample between 45 and 66.2 kDa. For each DOM sample, four equal slices were cut from the SDS-PAGE gel and each slice was analyzed using an LTQ-Orbitrap mass spectrometer. Peak list files obtained from the four fractions were processed separately and the peptide sequences were identified. Using the combined database, a total of 182 proteins matched by 286 unique peptides from 1263 spectra were identified from three DOM samples, in which the maximum number of spectra and proteins (76 proteins and 476 spectra) was observed in the DOM sample at the 75-m layer, whereas the minimum number (50 proteins and 331 spectra) was observed in the DOM sample at the 3000-m layer (Fig. 3). The number of spectra from the LDOM was significantly greater than that of spectra from the SDOM (approximately 3.4 times). However, the protein number of SDOM presented no significant difference between the surface and bathypelagic DOM samples. The NCBI accession number, protein name, protein probability, sequence coverage (%), number of unique peptides and total spectra used in the identification, theoretical molecular weight (MW) and isoelectric point, and the subcellular location of proteins identified are listed in the Web Appendix (www.aslo.org/lo/toc/vol_56/issue_5/1641a.html).

Biological origin of dissolved proteins in DOM—The number of spectra that matched phylogenetic groups contributing to dissolved proteins in the two DOM fractions from three water layers is shown in Table 2. Spectra of proteins derived from Proteobacteria were detected in each DOM fraction from the three water layers, with the highest number in the 75-m layer and the lowest in the 10-m layer (Table 2), and presented higher abundance in the LDOM, whereas the highest number of spectra from cyanobacterial proteins was observed in the 75-m layer and the lowest in the 3000-m layer. No spectra from archaeal protein were detected in the 10-m layer,

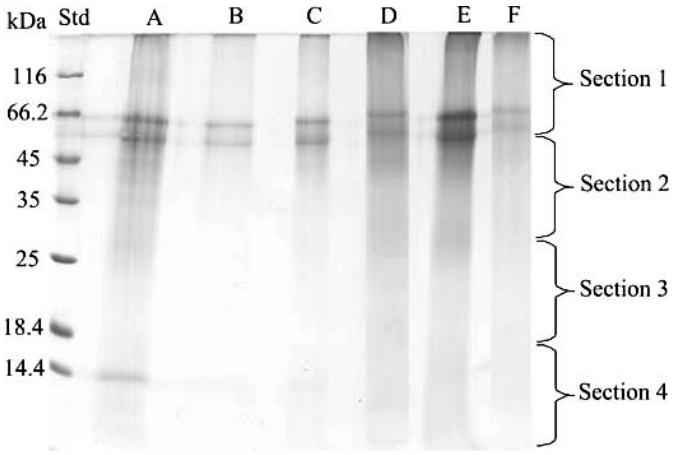

Fig. 2. 1-D SDS-PAGE electrophoretogram of the dissolved proteins in DOM. The left lane is a molecular weight standard. Lanes A, B, and C represent 5-kDa–0.2-μm size fractions from the 10-, 75-, and 3000-m layers. Lanes D, E, and F represent 0.2–0.7-μm size fractions from the 75-, 3000-, and 10-m layers. The gel was stained with colloidal Coomassie blue and cut into four pieces, i.e., sections 1, 2, 3, and 4.

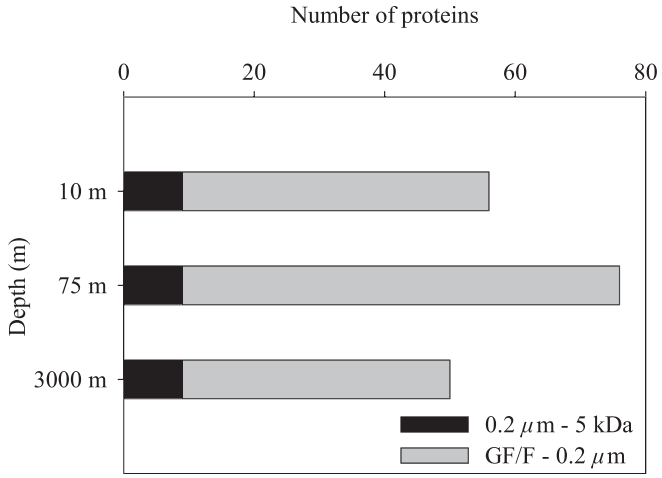

Fig. 3. The number of dissolved proteins identified from the two DOM fractions collected at different water layers at Sta. SEATS in the SCS.

whereas the number of spectra increased with the water depth and exhibited higher abundance in the SDOM, suggesting that archaeal activities played an important role in the dynamics of DOM from the bathypelagic zone. The number of spectra from Oomycetes in the DOM from the surface waters was much higher than that in the DOM of the bathypelagic zone, which was consistent with the aerobic feature of the Oomycetes as decomposers. Spectra from proteins of the Bacteroidetes, Planctomycetes, Firmicutes, and Actinobacteria were detected only in surface DOM samples, and almost all were detected in the LDOM. Spectra from phytoplanktonic proteins in the surface waters were more abundant than those in the bathypelagic zone: 471 spectra in the surface samples but only 132 spectra in the bathypelagic sample. However, a considerable number of spectra from the bacillariophytes, cryptophytes, chlorophytes, and stramenopiles were also detected in DOM samples at the 3000-m layer, accounting for 35%, 28%, 50%, and 21% of the total spectra of each phylogenetic group within the water column, respectively. The result indicated that the phytoplankton was also a major source of dissolved proteins in the water column and that only a few proteins were protected from degradation and so transferred into the bathypelagic DOM pool.

Subcellular location of dissolved proteins—For the LDOM fractions of DOM, cytoplasmic proteins were the major dissolved protein component ($\sim$ 51–79%) in each sample, and the highest spectrum number was observed in the 10-m layer before decreasing with depth (Fig. 4A). Spectra matching the prokaryotic outer membrane, periplasm, and inner membrane proteins were more abundant in the LDOM sample from the 75-m layer, accounting for approximately 35% of the total spectra, and including outer membrane lipoprotein, urea ABC transporter substrate binding protein, and porin. However, spectra that matched this group of proteins accounted for only approximately 3.5% and 10% in the 10- and 3000-m layers, respectively. Some low-abundance nuclear proteins were detected in all water layers, including adenosine triphosphate–dependent ribonucleic acid helicase eIF4A, 26S protease regulatory subunit, ubiquitin, and histone H2A, indicating that these picoeukaryote proteins were protected from degradation because of some unknown reasons, possibly such as association with other cellular components or inorganic matter. A few spectra from chloroplastic proteins were identified only in the LDOM from the 75-m water layer, including photosystem II protein D2, photosystem II reaction center protein D1, and ribulose 1.5-bisphosphate carboxylase oxygenase, probably suggesting that chloroplastic proteins from photosynthetic organisms were prone to be degraded, and few were preserved as refractory proteins in the DOM.

Compared to the LDOM, proteins from the SDOM were classified into fewer subcellular compartments (Fig. 4B). Cytoplasmic proteins were still the main components in all the SDOM samples, and their spectra accounted for approximately 46–87% of the total spectra in each SDOM sample. A highly abundant periplasmic protein, urea ABC transporter substrate binding protein, was detected in the SDOM sample from the 10-m water layer. Spectra from cytoplasmic protein, 5,10-methylenetetrahydromethanopterin reductase (MR) from the Archaea, presented high abundance in the SDOM samples from the 75- and 3000-m water layers, accounting for approximately 46% and 58% of the total spectra. Spectra from periplasmic proteins were more abundant in the SDOM sample from the 10-m layer, whereas only a small number of spectra from inner membrane proteins were detected in this layer.

COG functional classification of dissolved proteins in DOM—The identified dissolved proteins were functionally categorized based on COG annotation terms using the Basic Local Alignment Search Tool, with standard of reciprocal best hits, against the COG database. In total, approximately 8–21% of the total spectra for each LDOM and $\sim$ 32–43% of the total spectra for each SDOM could

Table 2. The number of spectra that matched phylogenetic groups contributing to dissolved proteins in the two DOM fractions from the three water layers in the SCS.

	Depth (m) and fraction					
	10		75		3000	
Species groups	0.2–0.7 μm	5 kDa–0.2 μm	0.2–0.7 μm	5 kDa–0.2 μm	0.2–0.7 μm	5 kDa–0.2 μm
Archaea	0	0	7	40	10	69
Proteobacteria	51	21	176	10	113	6
Bacteroidetes	11	1	10	0	0	0
Planctomycetes	50	0	0	0	0	0
Firmicutes	4	0	0	0	0	0
Actinobacteria	0	0	1	0	0	0
Cyanobacteria	13	33	92	1	3	2
Bacillariophyta	50	0	4	3	22	9
Prasinophyta	44	0	9	1	5	0
Dinophyta	9	0	17	0	0	0
Cryptophyta	23	8	15	5	16	4
Haptophyta	29	0	10	0	3	0
Chrysophyta	0	0	0	0	0	0
Chlorophyta	1	19	12	24	25	30
Stramenopiles	30	0	19	0	13	0
Oomycetes	59	0	24	0	1	0

not be assigned a functional annotation, and most of them were hypothetical proteins or predicted proteins.

In the LDOM fractions, the maximum number of spectra matching cytoskeleton proteins from Oomycetes and eukaryotic algae, including tubulins and actins, was observed in the 10-m layer, accounting for approximately 59% of the total spectra in the three water layers, and decreasing with depth (Fig. 5). In addition, spectra from cytoskeleton proteins also had the highest percentage in the 10-m layer, suggesting that these proteins might be accumulated as detrital proteins in the DOM. The abundance of proteins involved in energy production and conversion and posttranslational modification, protein turnover, and chaperones in the DOM sample from the 10-m water layer was significantly higher than that from the other DOM samples (Fig. 5). In the LDOM sample from the 10-m water layer, a total of seven proteins matched by 77 spectra were classified into energy production and conversion, including ATP synthase, malate dehydrogenase, N-ethylmaleimide reductase, and pyruvate carboxylase. Among them, spectra matching pyruvate carboxylase derived from Planctomycetes were very abundant, accounting for $\sim$ 65% of the total spectra of the group. In addition, 72 spectra corresponding to posttranslational modification, protein turnover, and chaperones were detected, accounted for approximately 19% of the total spectra in the 10-m layer, and mainly including different types of heat shock proteins derived from Stramenopiles, Oomycetes, Cryptophyta, and Firmicutes, together with T-complex protein from the Oomycetes. It is noteworthy that spectra matching proteins involved in lipid transport and metabolism, amino acid transport and metabolism, cell wall or membrane or envelope biogenesis, inorganic ion transport and metabolism, and photosynthesis were more abundant in the LDOM sample from the 75-m layer compared to the other water layers, whereas spectra from proteins involved in translation, ribosomal structure and biogenesis, and defense mechanisms were more abundant in the 3000-m layer (Fig. 5).

For the SDOM fractions, a total of 27 proteins that matched 286 spectra were identified. Most of the COG functional groups from the three water layers displayed minor differences in the abundance of spectra, except for proteins involved in amino acid transport and metabolism and in energy production and conversion (Fig. 6). Urea ABC transporter assigned to amino acid transport and metabolism was the most frequently detected cyanobacterial protein in the 10-m SDOM sample, whereas it was absent in the SDOM samples from the other two water layers. MR involved in energy production and conversion was the most abundant archaeal protein in the SDOM samples from the 75- and 3000-m water layers. Cytoskeleton proteins such as actins and proteins involved in posttranslational modification, protein turnover, and chaperones were identified in the SDOM samples from the three water layers, suggesting that they might be ubiquitous components of DOM. In addition, a considerable number of spectra from "unknown" proteins were identified in the three water layers, accounting for approximately 32–43% of the total spectra in each SDOM sample, indicating that the dissolved proteins in this fraction might become complex because of being modified. Totally, these results suggested that a few specific proteins were protected and preserved in the SDOM as refractory proteins.

Discussion

A widespread occurrence of dissolved proteins has been reported from the surface to deep oceanic waters in many regions of the world (Powell et al. 2005; Yamada and Tanoue 2006). However, so far, few proteins have been identified from DOM because of the limitation of the analytical methods for complex marine samples, and this

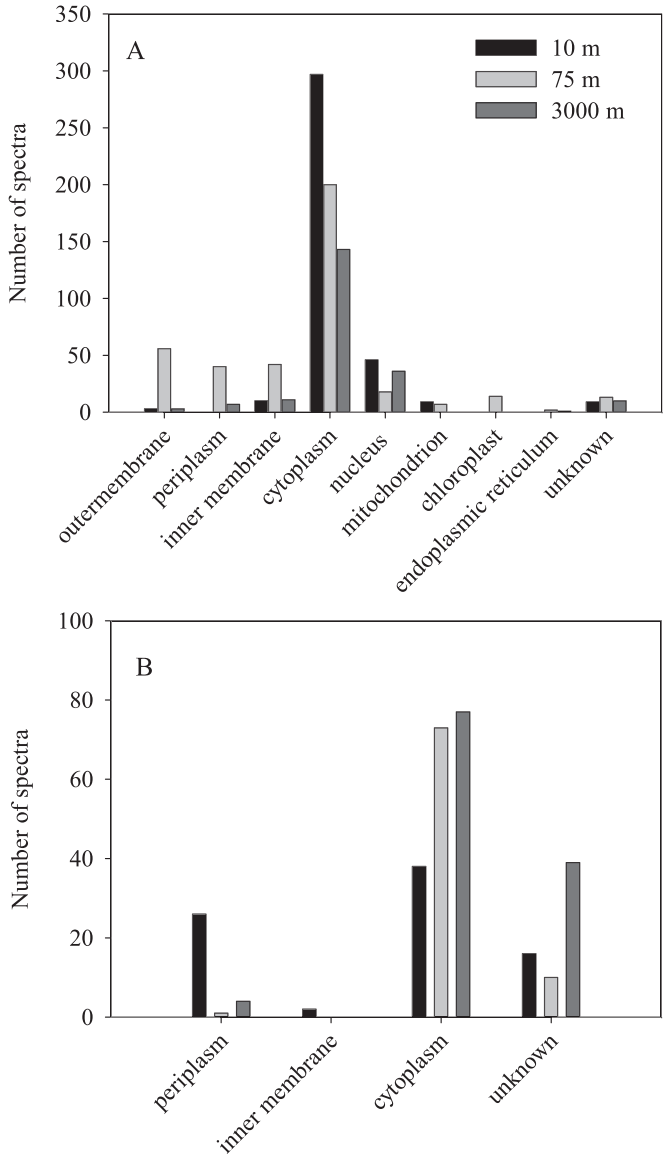

Fig. 4. Classification of all identified proteins by subcellular localization. (A) Subcellular localization for the 0.2–0.7-μm size fractions. (B) Subcellular localization for the 5-kDa–0.2-μm size fractions.

has impeded our understanding of marine DOM. Recently, metagenomic and metaproteomic approaches have been successfully applied to various environmental samples and have shown their power in identifying mixed proteins. Using the MS-based shotgun proteomic approach, thousands of proteins are identified from marine POM (Dong et al. 2010). In the present study, the MS-based proteomic approach was applied to analyze dissolved proteins in DOM and hundreds of proteins were first identified from marine DOM samples taken from the surface down to the bathypelagic zone. These proteins exhibited different biological sources and distribution patterns in the water column, which provided new insights into DOM with regard to its sources, chemical composition, and protection mechanism.

Source of dissolved proteins in DOM—It is reported that dissolved proteins in DOM are not intact proteins from living organisms but result from extracellular release by phytoplankton or from the cellular lysis of bacteria and phytoplankton by viruses or grazers (Tanoue 1995; Yamada and Tanoue 2003; Powell et al. 2005). According to the studies of Yamada and Tanoue (2003, 2009), the dissolved proteins detected in DOM are detrital proteins from the bacterial outer membrane and not from the living organisms. They conclude that living organisms, such as bacteria, passing through the GF/F filter cause insignificant effects on the dissolved proteins in the DOM, because the majority of the bacteria are filtered out by the GF/F filter, and the remainder produce invisible protein bands or spots in 1-D or 2-D gels visualized using silver staining. It is difficult to gain sufficient MS and MS^2 spectra in order to identify these proteins using present MS techniques. Another study also verifies that only 22–38% of the total bacterial biomass can pass through a GF/F filter (Lee et al. 1995). In our study, we investigated the protein profiles of DOM samples collected from different water layers using the same sampling procedure reported by Yamada and Tanoue (2003, 2006) and, although a considerable number of phytoplankton-originating proteins were found in both surface and bathypelagic DOM samples, we suggested that the dissolved proteins in the DOM were not derived from living organisms, because it is impossible that such a high density of phytoplankton cells could pass through the GF/F filter or live in the bathypelagic zone. Thus, it could be deduced that the proteins in the DOM were derived from detrital proteins originating from bacteria and phytoplankton more than from the living organisms retained on a 0.2-μm Durapore membrane. These proteins were tentatively protected by some unknown mechanisms or preserved as refractory proteins and accumulated in the DOM.

Previous work demonstrates that bacteria are an important source of dissolved proteins in DOM, and the genus *Pseudomonas* is one of the specific sources of dissolved proteins (Yamada and Tanoue 2003, 2006). In our study, highly abundant dissolved proteins originating from the Bacteriodetes, Planctomycetes, Firmicutes, Cyanobacteria, and Oomycetes were found in surface DOM samples, whereas almost no proteins from these phylogenetic groups were detected in the bathypelagic DOM sample (Table 2), indicating that these groups are the major contributors to dissolved proteins in the surface DOM but not the bathypelagic DOM. The prevalence of proteins derived from Proteobacteria in the DOM samples at different water layers indicated that the Proteobacteria was the most dominant phylogenetic group of the microbial community within the water column in the SCS, which was in agreement with the findings reported by Mering et al. (2007) making use of a set of protein-coding marker genes. It is interesting that no spectra from the Archaea were found in the 10-m layer DOM. However, the number of spectra matching archaeal proteins increased with water depth and reached its highest in the 3000-m layer. The Archaea are an important group in the marine ecosystem and are distributed widely from surface to deep sea. The highly abundant Archaea-originated proteins in the deep sea reflected the fact that Archaea

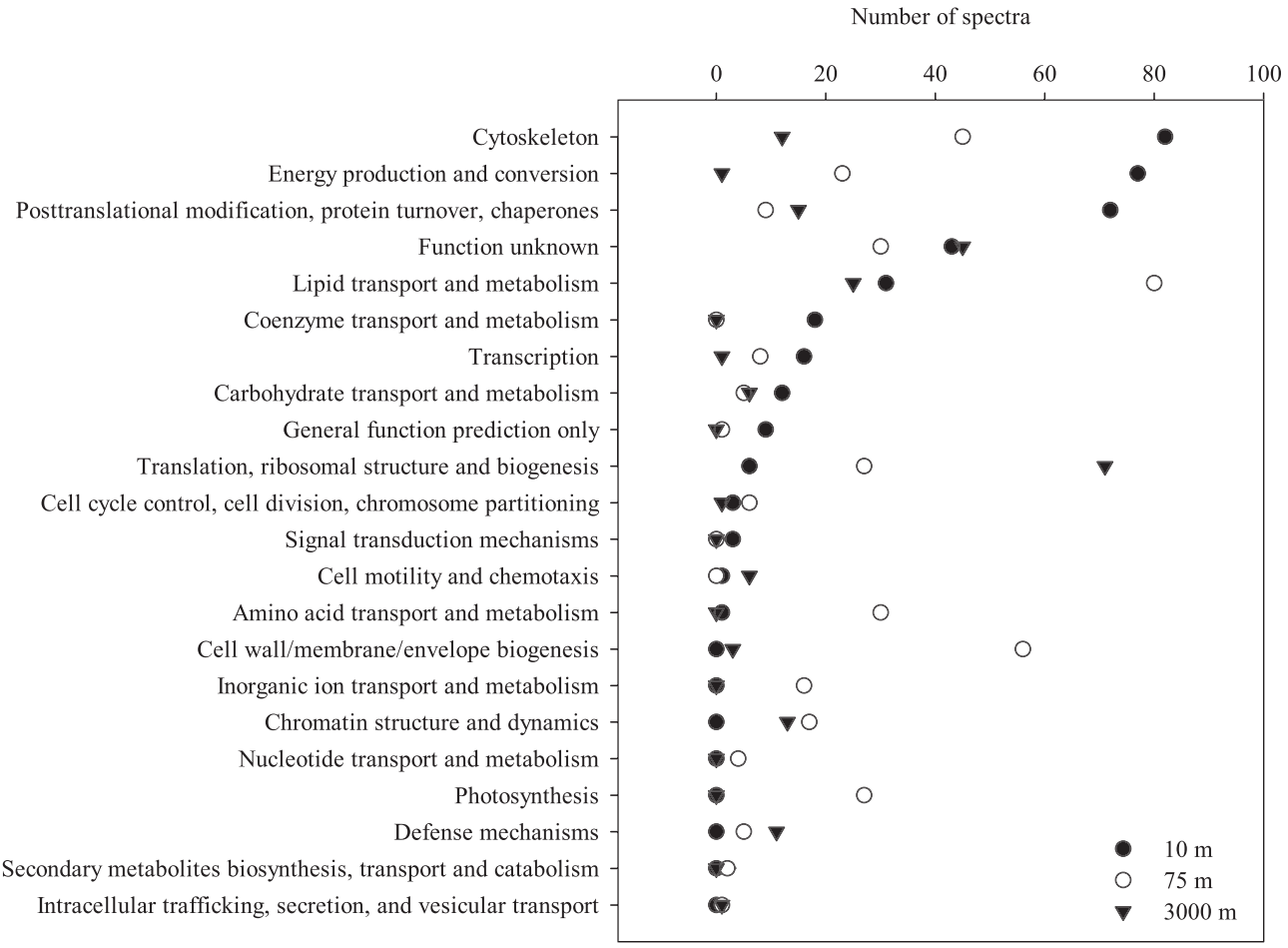

Fig. 5. COGs for dissolved proteins in the 0.2–0.7-μm size fractions from 10-, 75-, and 3000-m water depths.

play important roles in carbon biogeochemical cycling in the SCS. Based on the study of Karner et al. (2001), archaeal groups in the Pacific Ocean, such as pelagic crenarchaeota, are a consistent and significant component of the deep-sea microbiota, and they may rival bacterial abundances in the meso- and bathypelagic zones, and even gradually become the dominant microbial group.

In addition to bacterial sources, it is demonstrated that photosynthetic organisms are also an important source of dissolved proteins (Powell et al. 2005). Our study showed that spectra matched to the major phytoplankton groups were observed in both surface and bathypelagic DOM samples, although a declining trend was presented from the surface to the deep sea. It should be noted that a relatively abundant spectra of bacillariophytes, cryptophytes, chlorophytes, and stramenopiles were found in the DOM sample from the 3000-m layer, indicating that some proteins from these phylogenetic groups were protected from degradation and were transferred to the DOM, and accumulated in the deep sea. More interestingly, the number of spectra from chlorophyte proteins presented high abundance in the water column, especially in terms of SDOM in the bathypelagic zone, suggesting that this phytoplankton group might be one of the major sources of dissolved proteins in the DOM of the SCS.

Overall, the present results demonstrated that the sources of dissolved proteins were diverse in the surface DOM whereas the Archaea, Proteobacteria and some phytoplankton groups were major sources of dissolved proteins in the bathypelagic DOM, which further confirmed the conclusion that the dissolved proteins within the water column might mainly be regulated by in situ community structure at depth.

Features of dissolved proteins in DOM from different water layers—In our study, tubulins and actins assigned to cytoskeleton were frequently detected in each size fraction of DOM samples from the three water layers, and their abundances in both surface waters and the LDOM fraction were higher than in the deep sea and the SDOM fraction (Figs. 5, 6). In surface waters, flagella or cilia and microtubules or microfilaments from the Oomycetes and eukaryotic algae might be degraded directly into smaller protein molecules, which would enter the DOM pool because of a lack of aggregation. It is possible that highly abundant Oomycetes and eukaryotic algae from the surface waters served as a source, which favored the net accumulation of cytoskeleton proteins. Furthermore, slow decomposition of fecal pellet packages and aggregation in sinking POM might also be an important mechanism

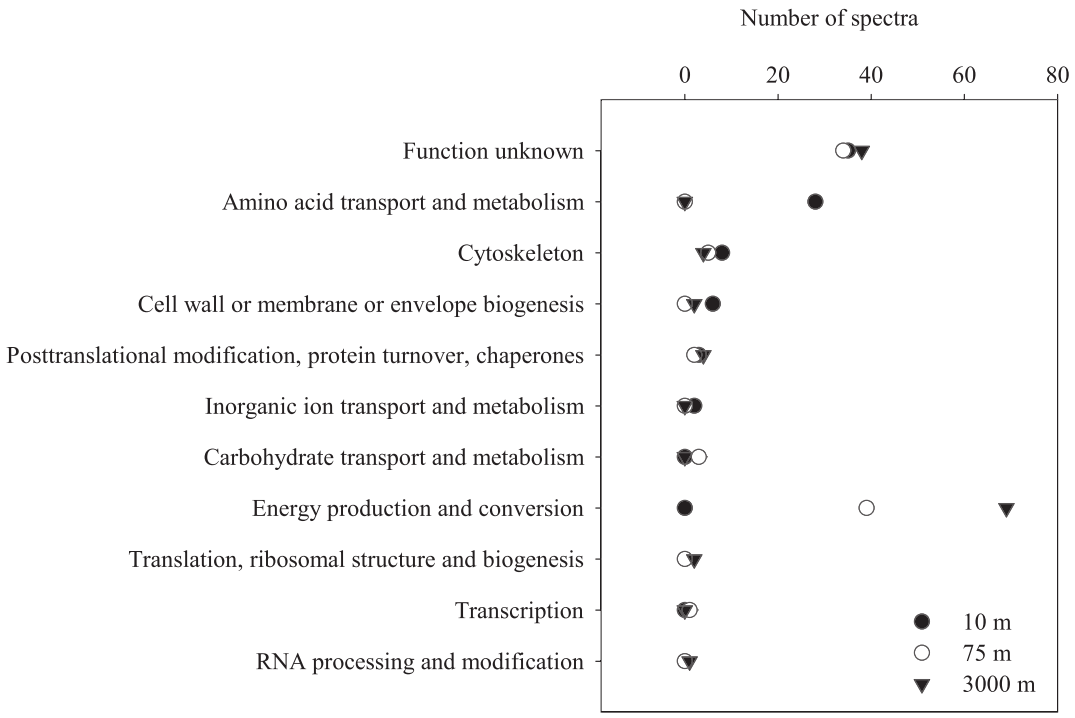

Fig. 6. COGs for dissolved proteins in the 5-kDa–0.2-μm size fractions from 10-, 75-, and 3000-m water depths.

for the occurrence of tubulins or actins in bathypelagic DOM.

So far, only a few membrane-associated enzymes (such as alkaline phosphatase, long-chain fatty acyl synthetase, and luminal binding protein) and uncombined enzymes (anthranilate synthase and ribulose bisphosphate carboxylase) are reported in marine DOM (Powell et al. 2005; Yamada and Tanoue 2006, 2009). Our present work also found that bacterial enzymes involved in energy production and conversion were more abundant in the LDOM fraction from surface waters (Fig. 5), in particular from the 10-m water layer. Some of them participate in electron transport or synthesis of energy and others take part in the citric acid cycle. Pyruvate carboxylase derived from Planctomycetes was the most abundant individual protein detected from the 10-m water layer. It is known that pyruvate carboxylase in bacteria synthesizes oxaloacetate from pyruvate and acts in gluconeogenic, glycerogenic, and anaplerotic roles, and its formation is regulated by the ambient carbon source (Peters-Wendisch et al. 1997). It might be inferred that the Planctomycetes were a dominant population in the 10-m water layer and produced abundant pyruvate carboxylase as a result of being subjected to environmental stresses, and could be used as potential protein marker of the upper layer surface DOM. These results also suggested that some bacteria in the surface layer might need more energy to adapt to the extreme environment of intense irradiance, high temperature, and ultraviolet radiation.

Interestingly, the most abundant protein involved in energy production and conversion found in the SDOM fraction from the bathypelagic zone was MR from methanogens in the Archaea (Fig. 6; Web Appendix). In methanogens, MR is involved in CO_2 reduction to CH_4, in methyl group oxidation to CO_2, and in autotrophic CO_2 fixation (Ma et al. 1991). The high abundance of MR in the bathypelagic zone suggested that this enzyme might play an important role in carbon cycling in the bathypelagic zone, which results in the accumulation of massive biogenic methane in the form of methane clathrates. On the other hand, MR might be used as a potential indicator of DOM from the deep sea.

Compared to DOM samples from the other two water layers, more protein types were detected abundantly in the LDOM sample from the 75-m layer (Fig. 5), including proteins involved in lipid, amino acid, and inorganic ion transport and metabolism; cell wall or membrane or envelope biogenesis; photosynthesis; and chromatin structure and dynamics, indicating that high biological activity occurred in this layer. It was noted that the 75-m layer was the deep chlorophyll *a* maximum layer (DCM) characterized by high abundance of phytoplankton and maximal photosynthetic activity during the sampling period. The high diversity and abundance of dissolved proteins in this layer might have originated from their release by phytoplankton or bacteria lysed by viruses. More recently, metagenomic analysis of DCM in the Mediterranean reveals highly abundant phages in the DCM layer, which play important roles in the lysis of microbes (Ghai et al. 2010). It is noteworthy that a few photosynthesis-associated proteins, such as photosystem II protein D2, photosystem II reaction center protein D1, and ribulose 1.5-bisphosphate carboxylase oxygenase, were found in this layer but not in the other two layers, reflecting the unique characteristics of DOM in the DCM layer.

It is surprising that more abundant spectra of bacterial proteins involved in translation, ribosomal structure, and

biogenesis were detected in the DOM sample from the 3000-m layer (Fig. 5; Web Appendix). Studies with *Escherichia coli* indicate that elevated pressure can inhibit many cellular functions, including substrate transport and biosynthesis of deoxyribonucleic acid, RNA, and protein (Yayanos and Pollard 1969). However, marine bacteria inhabiting the deep sea have evolved various strategies to adapt to the deep-sea environment with high pressures and low temperature, including synthesis of more heat shock proteins and unsaturated fatty acid, and introduction of the SS9 *rec*D gene (Bidle and Bartlett 1999; Pradillon and Gaill 2007). These studies suggest that the cellular metabolism of deep-sea bacteria including barotolerant and barophilic strains is slowed down, but that protein synthesis in their cells leads compared to other metabolic processes. As the factory of protein synthesis, ribosomes might be abundant in these bacteria. In addition, ribonucleoproteins, which are composed of ribosomal RNA and ribosomal proteins, might assemble into larger polymers in the deep-sea environment, and so be protected.

Overall, the above results indicated that some specific proteins were ubiquitous in the water column whereas a few unique proteins were only present in certain water layers, and this could be used as a potential biomarker to examine the dynamics of the DOM from the different water layers as well as the biological activity of in situ community in the environment.

Preservation of the dissolved proteins—Studies have demonstrated that the porins and OmpA-like proteins have rigid structures with β-barrel form and are resistant to biodegradation because of their being embedded in the bacterial outer membrane (Yamada and Tanoue 2006). Our study also found porins and Omp2b porin in DOM samples. In addition, it is likely that other proteins, such as actins, transporters and MR, which were detected in both surface and bathypelagic SDOM, were protected owing to their inherent molecular structure. The MR purified from *Methanopyrus kandleri* can withstand both high temperatures and high intracellular salt concentrations (Shima et al. 1998) because of the large number of negatively charged amino acid residues and the increased oligomeric interactions. Identification of abundant MR in SDOM from the 75- and 3000-m water layers should be ascribed to its special structure that may be resistant to biodegradation. Survival protein SurA and other molecular chaperones were also detected frequently in the SDOM. Survival protein SurA is a periplasmic molecular chaperone that facilitates correct folding of outer membrane porins (Bitto and McKay 2003), and so the protein might acquire resistance to degradation by association with detrital materials in the DOM, as do other molecular chaperones.

A few relatively abundant phytoplankton-originating proteins as well as nuclear proteins were observed in the 3000-m layer, suggesting that these proteins might be preserved and transferred to the deep-sea DOM. The mechanisms protecting these proteins from degradation are not clear. Two hypotheses are postulated to explain the existence of refractory DOM: physical protection and selective preservation (Hedges et al. 2000). Physically protected DOM molecules may be shielded by their matrix rather than their structure, as in the above-mentioned survival protein SurA, whereas selective preservation might come from the refractory nature of particular DOM molecules to their structure, such as porins and OmpA-like proteins. It is known that the marine DOM is divided into a pool of colloids or submicron particles and true solutes, or into a low-molecular-weight and a high-molecular-weight fraction. Chin et al. (1998) revealed that polymeric DOM in 0.2-μm–filtered surface seawater can spontaneously assemble, forming nanogels that then diffuse into the bulk water, annealing with each other to form microgels. The resulting microgels contain polysaccharides, proteins, and nucleic acid chains, forming noncovalently cross-linked physical polymer networks with random tangled topology. The size range of the nanogels and microgels varies between 1 nm and 1 μm (Verdugo et al. 2004). Our studies showed that the protein number in LDOM (0.2–0.7-μm size fraction) was greater than that in SDOM for the three water layers, suggesting that the LDOM might comprise a large quantity of nanogels or microgels formed from polymeric DOM, which protects dissolved proteins from degradation. It is postulated that proteins released from plankton might be more prone to assemble owing to their abundance, dimensional structures or charges, and membrane or liposome or other organic matrix encapsulation. On the other hand, it is likely that physical and chemical factors in different water depths, such as pH, temperature, salinity, and ionic concentration, might also determine which protein species were selected to form the nanogels or microgels.

Much effort has been devoted to the study of dissolved proteins of marine DOM, but only a few proteins have been identified so far because of the limitation of the analytical methods. This study is the first to characterize hundreds of proteins from DOM samples collected from three water layers using an MS-based ''shotgun'' proteomic approach, and these proteins presented different biological sources and distribution patterns in the water column. Bacteria and phytoplankton were the two major sources of dissolved proteins in DOM, although their contributions varied in different water layers. The detection of phytoplankton-originating proteins in DOM from the bathypelagic zone demonstrated that some phytoplankton proteins were protected and preserved in the DOM. The unique proteins found in certain water depths could be used as a potential biomarker to indicate the dynamics of the DOM from each different water layer as well as the biological activity of the ambient community. Certain proteins were observed throughout the water column, suggesting they were refractory components of DOM. Nanogels and microgels, acting as a physical barrier, might play important roles in protecting dissolved proteins from degradation as well as inherent molecular protection.

It must be pointed out that the proteins in the natural systems are likely more than what was identified and characterized in this study because of the fact that a significant fraction of the DOM might have been lost during the sampling process using the membrane filtration and ultrafiltration methods, which might hinder our

comprehensive understanding of marine DOM. We contend that more efforts should be devoted to develop more efficient and speedier sampling techniques to improve the sampling recovery of marine DOM for subsequent characterization of dissolved proteins. Moreover, there are a large set of proteins still unidentified because of the limitations of the genomic and proteomic databases of marine organisms, i.e., phytoplankton, zooplankton, and viruses. With an increasing metagenomic database of marine organisms and fast development of mass spectrometric techniques, it is likely that we will gain more information about the processes involved in DOM remineralization and the biological mechanisms occurring in DOM on a proteomic scale, and this will certainly improve our understanding of the sources, dynamics, and protection mechanisms of the dissolved proteins. Furthermore, the large-scale comparative studies of dissolve proteins in different regional ecosystems will also provide new insights into carbon cycling and biological activities in the ocean.

Acknowledgments

We thank the captain and crew of the R/V *DongFangHong 2* for their assistance during the sampling cruise. This study was supported by the Ministry of Science and Technology through grant 2009CB421203, and the National Natural Science Foundation of China through grants 40821063, 40776068, and 40876059. D.-Z. Wang was also supported by the Program for New Century Excellent Talents in University of China. We also thank John Hodgkiss from The University of Hong Kong for his help with English. The authors also appreciate the comments and helpful suggestions of two anonymous reviewers.

References

Aebersold, R., and M. Mann. 2003. Mass spectrometry-based proteomics. Nature **422:** 198–207, doi:10.1038/nature01511

Benner, R., and K. Kaiser. 2003. Abundance of amino sugars and peptidoglycan in marine particles and dissolved organic matter. Limnol. Oceanogr. **48:** 118–128, doi:10.4319/lo.2003.48.1.0118

Bidle, K. A., and D. H. Bartlett. 1999. RecD function is required for high-pressure growth of a deep-sea bacterium. J. Bacteriol. **181:** 2330–2337.

Bitto, E., and D. B. McKay. 2003. The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. J. Biol. Chem. **278:** 49316–49322, doi:10.1074/jbc.M308853200

Carlson, C. A. 2002. Production and removal process, p. 91–139. *In* D. A. Hansell and C. A. Carlson [eds.], Biogeochemistry of marine dissolved organic matter. Academic.

Chin, W. C., M. V. Orellana, and P. Verdugo. 1998. Spontaneous assembly of marine dissolved organic matter into polymer gels. Nature **391:** 568–572, doi:10.1038/35345

Dai, M. H., K. O. Buesseler, P. Ripple, J. Andrews, R. A. Belastock, O. Gustafsson, and S. B. Moran. 1998. Evaluation of two cross-flow ultrafiltration membranes for isolating marine organic colloids. Mar. Chem. **62:** 117–136, doi:10.1016/S0304-4203(98)00028-0

Dittmar, T., H. P. Fitznar, and G. Kattner. 2001. Origin and biogeochemical cycling of organic nitrogen in the eastern Arctic Ocean as evident from D- and L-amino acids. Geochim. Cosmochim. Acta **65:** 4103–4114, doi:10.1016/S0016-7037(01)00688-3

Dong, H. P., D. Z. Wang, M. H. Dai, L. L. Chan, and H. S. Hong. 2009. Shotgun proteomics: Tools for the analysis of marine particulate proteins. Limnol. Oceanogr.: Methods **7:** 865–874, doi:10.4319/lom.2009.7.865

———, ———, ———, and H. S. Hong. 2010. Characterization of particulate organic matter in the water column of the South China Sea using shotgun proteomic approach. Limnol. Oceanogr. **55:** 1565–1578, doi:10.4319/lo.2010.55.4.1565

Ducklow, H. W. 2002. Foreword, p. xv–xviii. *In* D. A. Hansell and C. A. Carlson [eds.], Biogeochemistry of marine dissolved organic matter. Academic.

Ghai, R., and others. 2010. Metagenome of the Mediterranean deep chlorophyll maximum studied by direct and fosmid library 454 pyrosequencing. ISME J. **4:** 1154–1166, doi:10.1038/ismej.2010.44

Hansell, D. A., and C. A. Carlson. 2001. Biogeochemistry of total organic carbon and nitrogen in the Sargasso Sea: Control by convective overturn. Deep-Sea Res. II **48:** 1649–1667, doi:10.1016/S0967-0645(00)00153-3

He, P., and others. 2005. The human plasma proteome: Analysis of Chinese serum using shotgun strategy. Proteomics **5:** 3442–3453, doi:10.1002/pmic.200401301

Hedges, J., and others. 2000. The molecularly-uncharacterized component of nonliving organic matter in natural environments. Org. Geochem **31:** 945–958, doi:10.1016/S0146-6380(00)00096-6

Jiao, N. Z., and others. 2010. Microbial production of recalcitrant dissolved organic matter: Long-term carbon storage in the global ocean. Nature Rev. Microbiol. **8:** 593–8599, doi:10.1038/nrmicro2386

Kan, J., T. E. Hanson, J. M. Ginter, K. Wang, and F. Chen. 2005. Metaproteomic analysis of Chesapeake Bay microbial communities. Saline Syst. **1:** 1–13, doi:10.1186/1746-1448-1-7

Karner, M. B., E. F. DeLong, and D. M. Karl. 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. Nature **409:** 507–510, doi:10.1038/35054051

Keller, A., A. I. Nesvizhskii, E. Kolker, and R. Aebersold. 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. **74:** 5383–5392, doi:10.1021/ac025747h

Lee, S. H., Y. C. Kang, and J. A. Fuhrman. 1995. Imperfect retention of natural bacterioplankton cells by glass fiber filters. Mar. Ecol. Prog. Ser. **119:** 285–290, doi:10.3354/meps119285

Lo, I., and others. 2007. Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria. Nature **446:** 537–541, doi:10.1038/nature05624

Lu, Z., and others. 2004. Predicting subcellular localization of proteins using machine-learned classifiers. Bioinformatics **20:** 547–556, doi:10.1093/bioinformatics/btg447

Ma, K., D. Linder, K. O. Stetter, and R. K. Thauer. 1991. Purification and properties of N^5,N^{10}-methylenetetrahydromethanopterin reductase (coenzyme F_{420}-dependent) from the extreme thermophile *Methanopyrus kandleri*. Arch Microbiol. **155:** 593–600, doi:10.1007/BF00245355

Mering, C., and others. 2007. Quantitative phylogenetic assessment of microbial communities in diverse environments. Science **315:** 1126–1130, doi:10.1126/science.1133420

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

Nesvizhskii, A. I., A. Keller, E. Kolker, and R. Aebersold. 2003. A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. **75:** 4646–4658, doi:10.1021/ac0341261

Peters-Wendisch, P. G., V. F. Wendisch, S. Paul, B. J. Eikmanns, and H. Sahm. 1997. Pyruvate carboxylase as an anaplerotic enzyme in *Corynebacterium glutamicum*. Microbiology **143:** 1095–1103, doi:10.1099/00221287-143-4-1095

Powell, M. J., J. N. Sutton, C. E. Del Castillo, and A. I. Timperman. 2005. Marine proteomics: Generation of sequence tags for dissolved proteins in seawater using tandem mass spectrometry. Mar. Chem. **95:** 183–198, doi:10.1016/j.marchem.2004.09.004

Pradillon, F., and F. Gaill. 2007. Pressure and life: Some biological strategies. Rev. Environ. Sci. Biotechnol. **6:** 181–195, doi:10.1007/s11157-006-9111-2

Ram, R. J., and others. 2005. Community proteomics of a natural microbial biofilm. Science **308:** 1915–1920, doi:10.1126/science. 1109070

Ryu, S., B. Gallis, Y. A. Goo, S. A. Shaffer, D. Radulovic, and D. R. Goodlett. 2008. Comparison of a label-free quantitative proteomic method based on peptide ion current area to the isotope coded affinity tag method. Cancer Inform. **6:** 243–255.

Shima, S., C. Tziatzios, D. Schubert, H. Fukada, K. Takahashi, U. Ermler, and R. K. Thauer. 1998. Lyotropic-salt-induced changes in monomer/dimer/tetramer association equilibrium of formyltransferase from the hyperthermophilic *Methanopyrus kandleri* in relation to the activity and thermostability of the enzyme. Eur. J. Biochem. **258:** 85–92, doi:10.1046/j.1432-1327.1998.2580085.x

Siegenthaler, U., and J. L. Sarmiento. 1993. Atmospheric carbon dioxide and the ocean. Nature **365:** 119–125, doi:10.1038/365119a0

Sowell, S. M., and others. 2009. Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. ISME J. **3:** 93–105, doi:10.1038/ismej.2008.83

Tanoue, E. 1995. Detection of dissolved protein molecules in oceanic waters. Mar. Chem. **51:** 239–252, doi:10.1016/0304-4203(95)00061-5

———, S. Nishiyama, M. Kamo, and A. Tsugita. 1995. Bacterial membranes: Possible source of a major dissolved protein in seawater. Geochim. Cosmochim. Acta **59:** 2643–2648, doi:10.1016/0016-7037(95)00134-4

Tatusov, R. L., M. Y. Galperin, D. A. Natale, and E. V. Koonin. 2000. The COG database: A tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28:** 33–36, doi:10.1093/nar/28.1.33

Verdugo, P., A. L. Alldredge, F. Azam, D. L. Kirchman, U. Passow, and P. H. Santschi. 2004. The oceanic gel phase: A bridge in the DOM-POM continuum. Mar. Chem. **92:** 67–85, doi:10.1016/j.marchem.2004.06.017

Wilm, M., A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, and M. Mann. 1996. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature **379:** 466–469, doi:10.1038/379466a0

Wilmes, P., M. Wexler, and P. L. Bond. 2008. Metaproteomics provides functional insight into activated sludge wastewater treatment. PLoS One **3:** e1778, doi:10.1371/journal.pone.0001778

Yamada, N., and E. Tanoue. 2003. Detection and partial characterization of dissolved glycoproteins in oceanic waters. Limnol. Oceanogr. **48:** 1037–1048, doi:10.4319/lo.2003.48.3.1037

———, and ———. 2006. The inventory and chemical characterization of dissolved proteins in oceanic waters. Prog. Oceanogr. **69:** 1–18, doi:10.1016/j.pocean.2005.11.001

———, and ———. 2009. Similarity of electrophoretic dissolved protein spectra from coastal to pelagic seawaters. J. Oceanogr. **65:** 223–233, doi:10.1007/s10872-009-0021-0

Yayanos, A. A., and E. C. Pollard. 1969. A study of the effects of hydrostatic pressure on macromolecular synthesis in *Escherichia coli*. Biophysics **9:** 1464–1482, doi:10.1016/S0006-3495(69)86466-0

Zou, L., X. C. Wang, J. Callahan, R. A. Culp, R. F. Chen, M. A. Altabet, and M. Y. Sun. 2004. Bacterial roles in the formation of high-molecular weight dissolved organic matter in estuarine and coastal waters: Evidence from lipids and the compound-specific isotopic ratios. Limnol. Oceanogr. **49:** 297–302, doi:10.4319/lo.2004.49.1.0297

Associate editor: Wade H. Jeffrey

Received: 23 December 2010
Accepted: 02 May 2011
Amended: 17 May 2011