Draft genome sequence of a psychrotolerant sulfur-oxidizing bacterium *Sulfuricella denitrificans* skB26 and proteomic insights into cold adaptation

Tomohiro Watanabe, Hisaya Kojima*, Manabu Fukui

The Institute of Low Temperature Science, Hokkaido University, Sapporo, Japan

Running title: Genome and proteome of *Sulfuricella denitrificans*

*Corresponding author.
Postal address: The Institute of Low Temperature Science, Kita-19, Nishi-8, Kita-ku, Sapporo 060-0819, Japan
Tel: +81-11-706-5460
Fax: +81-11-706-5460
E-mail: kojimah@pop.lowtem.hokudai.ac.jp
Abstract

Except for several conspicuous cases, very little is known about sulfur oxidizers living in natural freshwater environments. *Sulfuricella denitrificans* skB26 is a psychrotolerant sulfur oxidizer recently isolated from a freshwater lake, as a representative of a new genus in the class *Betaproteobacteria*. In this study, approximately 3.2 Mb of draft genome sequence of the strain skB26 was obtained. In the draft genome consisting of 23 contigs, a single rRNA operon, 43 tRNA genes, and 3,133 coding sequences were identified. The identified genes include those required for sulfur oxidation, denitrification, and carbon fixation. Comparative proteomic analysis was conducted to assess cold adaptation mechanisms of this organism. From the cells grown at 22°C and 5°C, proteins were extracted for analysis with nano-liquid chromatography-electrospray ionization-tandem mass spectrometry. In the cells cultured at 5°C, relative abundances of ribosomal proteins, cold shock proteins, and DEAD/DEAH box RNA helicases were increased in comparison to 22°C. These results suggest that maintenance of proper translation is critical for growth under low-temperature conditions, similar to other cold-adapted prokaryotes.
INTRODUCTION

Chemolithoautotrophic sulfur-oxidizing bacteria, widely distributed in aquatic environments, have attracted interest among microbial ecologists. With their functions of inorganic carbon fixation and sulfur oxidation, they thrive as primary components of ecosystems fueled by reduced sulfur species, represented by those in terrestrial hot springs and hydrothermal vents. On the other hand, significant influences of sulfur oxidizers on cycles of nitrogen and phosphorous have been shown in several nitrate-storing bacteria, which are supplied with sulfide from sulfate reduction in organic-rich marine sediments (20, 27, 28). Except for these conspicuous cases, however, ecology of chemolithoautotrophic sulfur oxidizers in natural environments remains largely unknown. In marine environments, several key players in sulfide oxidation have been identified only recently, in intertidal sediment (15) and water columns (7, 14, 32). A trial to identify major planktonic sulfur oxidizers has been also performed in a stratified freshwater lake (1), but very little is known about chemolithotrophic sulfur oxidizers living in general freshwater environments.

Recently, a chemolithotrophic sulfur oxidizer *Sulfuricella denitrificans* skB26 was isolated from a stratified freshwater lake (13), as a representative of a new genus in the class Betaproteobacteria. Occurrences of the close relatives of this bacterium have been reported in several studies, as 16S rRNA gene clones detected in libraries constructed with general primer pairs (3, 16, 19, 24). The strain skB26 was isolated from cold anoxic hypolimnion, and has ability to grow at temperatures as low as 0°C. In general aquatic ecosystems, primary source of reduced sulfur species is sulfate reduction in anoxic nitrate-depleted zones, which are often associated with low temperature. Therefore, cold adaptation may have essential significance for sulfur oxidizers in natural environments for access to sufficient supplies of growth substrate. Although several sulfur oxidizers have been isolated from low-temperature environments (7, 12, 26), their cold adaptation mechanisms are poorly understood. Here, we
report the draft genome sequence of this novel psychrotolerant chemolithoautotroph, along with results of proteomic analysis to gain insight into cold adaptation.

MATERIALS AND METHODS

Genome sequence and annotation

*Sulfuricella denitrificans* skB26 was grown in the NaCl-free defined medium previously described (13) containing 20 mM thiosulfate and 20 mM nitrate at 22°C. Genomic DNA was extracted by using AquaPure Genomic Isolation kit (Bio-Rad Laboratories, CA), and then fragmented by shearing with a Hydroshear (GeneMachines, CA). From the resulting DNA fragments, 8 kbp paired-end library was prepared using GS Titanium Library Paired End Adaptors (Roche Diagnostics, CH). The library was sequenced by using a Genome Sequencer FLX System (Roche Diagnostics). The paired-end reads were assembled by using GS De Novo Assembler version 2.3 with default parameters. Gaps between contigs belonging to the same scaffold were closed by Sanger sequencing of PCR products. The draft genome sequence was automatically annotated using Microbial Genome Annotation Pipeline (www.migap.org/). In the pipeline, open reading frames (ORFs) were identified by MetaGene Annotator, and then predicted ORFs were searched against reference databases including RefSeq, TrEMBL, and Clusters of Orthologous Groups of proteins (COGs) dataset. Genes for tRNAs and rRNAs were identified by tRNAscan-SE and rRNAmer, respectively (30). Further genome analysis was performed using IMC-GE software (In Silico Biology, Japan). In the draft genome, homolog genes involved in sulfur oxidation, denitrification, carbon dioxide fixation were identified by BLASTP search using the sequence data of *Thiobacillus denitrificans* as query (threshold, >30%), and normalized gene names were assigned. To predict function of two homolog genes encoding Csp family proteins, the amino acid sequences were subjected to PSI-BLAST search considering identity above 90%, and then
manually annotated.

Genome sequence accession number

The *Sulfuricella denitrificans* skB26 draft genome sequence data have been deposited to DDBJ/EMBL/GenBank databases under accession numbers BAFJ0100001- BAFJ0100023. The version described in this paper is the first version.

Protein extraction and SDS-PAGE

*S. denitrificans* was cultured in the same medium at the temperature for optimum growth (22°C) and *in situ* temperature of isolation source (5°C). Cells were harvested at stationary phase (35 and 68 days, respectively) and washed once with phosphate buffered saline by centrifugation at 10,000 × g for 20 min at 4°C. The pellets were frozen at -80°C until protein extraction. Protein extraction was performed using a kit intended to extract total cellular proteins including membrane proteins, i.e., ReadyPrep Protein Extraction Kit (Total Protein) (Bio-Rad Laboratories, CA). The cell pellets were suspended in a buffer included in the kit which contains strong chaotropic extraction reagent. After 10 times sonications for 10 sec at 30-sec intervals, resulting lysates were centrifuged for 20 min at 10,000 rpm at 20°C. The soluble fractions were recovered and used for further analysis. The protein content was quantified by the Bradford method using a Bio-Rad protein assay kit (Bio-Rad Laboratories). The protein samples were mixed with sodium dodecyl sulfate sample buffer. The resulting samples were incubated at 99°C for 5 min, and then 50 µg of the denatured protein samples were subjected to SDS-PAGE using a pre-cast gel, SuperSep™ Ace 12.5% (Wako pure Chemical industries, Japan). After electrophoresis, the gel was stained with Coomassie Brilliant Blue R-250, destained, and the lanes were cut out for trypsin in-gel protein digestion.
Trypsin in-gel protein digestion and mass analysis

Proteins were digested with trypsin as previously described (29). Nano-liquid chromatography-electrospray ionization-tandem mass spectrometry (nano-LC-ESI-MS/MS) experiment was performed on a multidimensional HPLC Paradigm MS2 (AMR Inc, Japan) and nanospray electrospray ionization device (Michrom Bioresources Inc, CA) connected to a LTQ ion-trap MS (Thermo Fisher Scientific, MA). The digested peptides were loaded on an L-column 2 ODS (Chemicals Evaluation & Research Inst., Japan) packed with C18 (5 µm, 12 nm pore size). Solvent system consisted of solvent A (2% acetonitrile and 0.1% formic acid in H2O) and solvent B (90% acetonitrile and 0.1% formic acid in H2O). Flow rate was maintained at 1 µl/min, and solvent B was increased in a linear gradient from 5% to 65% over 40 min. Then, it was further increased to 95% and kept at the concentration for 5 min before returning to 5% for analysis of the next sample. Peptide spectra were recorded in a range of m/z 450-1800. Mass spectra were acquired in data-dependent scan mode. One MS/MS spectrum of the one most intense peak was collected following on the full spectrum scan.

MASCOT search and semi-quantitative analysis of proteins

For peptide identification from MS/MS data, a protein sequence database was constructed from the S. denitrificans draft genome. By using the MASCOT (ver.2.3.01), MS/MS spectra were searched against the generated database. Search parameters were set as follows; tryptic digest with a maximum of two missed cleavage; fixed modifications, carbamidomethyl cysteine; variable modifications, methionine oxidation; peptide masses, monoisotopic; positive charge (+1, +2, +3) of peptide; mass tolerance of 1.2 Da for the precursor ion and 0.8 Da for productions. To assess false-positive identifications, an automatic decoy search was performed against a randomized database with significance threshold of p<0.05, and the estimated false discovery rate was below 3.4 %. Protein identification was
judged as positive when two or more different peptides derived from it were detected.

To evaluate protein content, exponentially modified protein abundance index (emPAI) was determined as previously described (10). For comparative analysis, normalized protein content (PC) value was calculated as follows: PC = \( \frac{\text{emPAI}}{\sum \text{emPAI}} \times 100 \) (%), where \( \sum \text{emPAI} \) is the summation of emPAI values for all identified proteins.

RESULTS AND DISCUSSION

Genome sequence and annotation

By pyrosequencing, we sequenced 77,574,798 bp from 386,840 reads. The generated sequences were assembled into 177 contigs, and two scaffolds containing multiple contigs were constructed. There were 124 gaps within the scaffolds, and 117 of them were closed by Sanger sequencing of PCR products. As a result, *S. denitrificans* skB26 draft genome consisting of 23 contigs was obtained. Basic characteristics of the draft genome are shown in Table 1. The functional classification of the identified CDSs based on the COG designations is shown in Supplementary Fig. S1.

Genes required for sulfur oxidation, denitrification, and carbon fixation are summarized in Table S1. As reported previously, strain skB26 oxidizes thiosulfate and elemental sulfur to sulfate. The draft genome contains the *soxXYZAB* gene cluster but lacks *soxCD*, suggesting generation of elemental sulfur as an intermediate of thiosulfate oxidation (18). In *Allochromatium vinosum*, elemental sulfur is oxidized to sulfite by the dissimilatory sulfite reductase system encoded by *dsr* genes (2). Accordingly, the gene cluster *dsrABEFHCMKLJOPNR* occurs in the genome of skB26. We also identified the putative *hdrAABC* gene cluster encoding the heterodisulfide reductase complex, hypothesized to oxidize disulfide intermediates to sulfite (21). It has been proposed that adenosine-5’-phosphosulfate reductase (*apr*) and ATP sulfurylase (*sat*) oxidize sulfite to
sulfate (17); these genes were also found in the strain. Furthermore, sulfide:quinone oxidoreductase (*sqr*), which catalyzes sulfide oxidation (22), occurred in the genome.

Strain skB26 contains genes necessary for denitrification, including nitrite reductase (*nir*), nitric oxide reductase (*nor*), and nitrous oxide reductase (*nos*) (23). This draft genome lacks cytoplasmic membrane nitrate reductase (*nar*), thought to be generally possessed by denitrifiers (23), but contains periplasmic nitrate reductase (*nap*).

The effects of low temperature on relative protein abundance

For proteomic analysis, *S. denitrificans* was grown on thiosulfate under denitrifying conditions. Cells cultured at 22°C (MT) and 5°C (LT) were harvested in stationary phase and subjected to comparative proteomic analysis using nano-LC-ESI-MS/MS. Using the MASCOT search, 1450 and 1579 proteins were identified in MT and LT respectively. The detected proteins were classified into functional categories based on COG designations, but effect of temperature was not apparent in this qualitative comparison (Fig. S1). For the comparisons based on semi-quantitative analysis, normalized protein content (PC) values were calculated.

The most abundant protein in LT was ribosomal protein L5 (PC = 47%), which increased about 5-fold compared to MT (Table S2). Similarly, a 20-fold increase was observed for ribosomal protein L7/L12 in LT, and 6 ribosomal proteins were identified only in LT. The protein L7/L12 has been reported as a cold acclimation protein (6), and increased ribosomal protein abundance at low temperatures has been observed in other prokaryotes (11, 33). The high cellular levels of ribosomal proteins observed in *S. denitrificans* may be associated with cold adaptation.

To eliminate interference from abundant ribosomal proteins, PC values were recalculated after excluding these proteins for comparison of other proteins. In order to
examine effects of temperature specifically observed in anaerobic sulfur oxidizers, PC values of proteins putatively involved in sulfur oxidation, denitrification, and carbon dioxide fixation were compared between LT and MT (Fig. 1). Among these three functions, differences in relevant proteins were apparent in sulfur oxidation as generally lower contents in LT (Fig. 1A). These sulfur oxidation-related proteins, including heterodisulfide reductase (Hdr) which lacks direct evidence for involvement in thiosulfate oxidation, were the second most abundant following the ribosomal proteins in both MT and LT. Sum of the PC values of proteins shown in Fig. 1A reached 40% in MT, but it decreased to 17% in LT. These results may be relevant to suppressed energy production at low temperature, as shown in other cold-adapted bacterium (31). It should be noted, however, decrease in total sulfur oxidation-related proteins was largely attributed to exceptionally drastic changes in DsrH and SoxZ (Fig. 1). In the process of sulfur oxidation, DsrH and SoxZ work in conjunction with DsrEF and SoxY, respectively (4). The imbalanced changes of DsrH and SoxZ may have no connection to energy production.

As candidates for components of the cold adaptation mechanism, proteins with greater PC values in LT (more than 5 times those in MT) after the recalculation are listed in Table 2. One of the candidate proteins for cold adaptation, cold shock protein (CspD) has been reported as a cold-inducible protein in Bacillus subtilis (5). Another cspD homolog was also found in the genome (SCD_00462), and its product was identified only in LT as the most abundant LT-specific protein (Table S3). One of the major negative effects of low temperature is the inhibition of translation caused by stabilization of the mRNA secondary structure. It has been proposed that Csps work in conjunction with DEAD/DEAH box helicase to maintain proper conformation of mRNA in B. subtilis (8). In the draft genome of strain skB26, 3 copies of the DEAD/DEAH box helicase gene were identified. Protein coded by one of these genes was more abundant (about 9-fold) at 5°C (Table 2), and product of another copy was
identified only in LT (Table S3). The greatest change (32-fold) was observed in sulfur relay protein DsrC (SCD_00181), although other putative sulfur relay proteins encoded by five genes were not increased in LT (Table S4). In *Escherichia coli*, the DsrC homolog is involved in thiolation of tRNA, which facilitates ribosome binding (9). Increased abundance of these proteins and ribosomal proteins consistently suggest maintenance of proper translation at low temperatures is important for cold adaptation of *S. denitrificans* skB26.

Another candidate protein for cold adaptation was beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ), a protein involved in fatty acid biosynthesis (Table 2). It is well known that modification of membrane fatty acid composition is critical to maintain membrane fluidity under cold conditions, and various bacteria tend to increase content of unsaturated fatty acids at low temperatures (25). In a cycle of hydrocarbon chain elongation for fatty acids biosynthesis, FabZ catalyzes dehydration to generate unsaturated fatty acids (34). Although fatty acid profile of strain skB26 has not been revealed, it may change in response to low temperatures with involvement of the FabZ.

Other candidate proteins shown in Table 2 were classified into various functional categories including cell motility, nucleic acid metabolism, energy production, and signal transduction. As to these proteins, it was impossible to deduce how they contribute to growth or survival at low temperatures. Their functional diversity may reflect universal effects of temperature on various biological processes and respective responses of this organism to cope with them.

From the comparative proteomic analysis, it was suggest that mechanisms to maintain proper translation and membrane fluidity play central roles in cold adaptation of *S. denitrificans*. It should be noted that the protein identification was performed on the basis of draft genome, and thus some important genes and proteins may be still left undiscovered. For further detailed analyses, a project to obtain complete genome sequence of strain skB26 is
ongoing. Considering that studies on sulfur oxidizers tend to place a disproportionate emphasis on extremophiles living in harsh environments, knowledge from studies on S. denitrificans will be valuable for a better understanding of the sulfur cycle in general freshwater environments.

Acknowledgments

We thank Y. Kasahara and R. Kadoya for assistance in proteomic analysis. This study was supported by a grant from the Ministry of Education, Culture, Sports, Science and Technology, Japan, to Fukui (2237005).

References


Table 1. Basic characteristics of *S. denitrificans* skB26 draft genome.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size (bp)</td>
<td>3238714</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>56.1</td>
</tr>
<tr>
<td>No. of contigs</td>
<td>23</td>
</tr>
<tr>
<td>Max contig length (bp)</td>
<td>737595</td>
</tr>
<tr>
<td>Total no. of coding sequences</td>
<td>3133</td>
</tr>
<tr>
<td>Coding sequence density (gene/kbp)</td>
<td>0.905</td>
</tr>
<tr>
<td>Total CDS length (bp)</td>
<td>2931503</td>
</tr>
<tr>
<td>Average gene length (bp)</td>
<td>935.7</td>
</tr>
<tr>
<td>No. of rRNA operons</td>
<td>1</td>
</tr>
<tr>
<td>No. of tRNA genes</td>
<td>43</td>
</tr>
</tbody>
</table>
Table 2. List of protein more abundant (fold change>5) at 5°C.

<table>
<thead>
<tr>
<th>Locus tag</th>
<th>Description</th>
<th>Functional category</th>
<th>Normalized protein content (PC)</th>
<th>Ratio of PC (5°C/22°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCD_00181</td>
<td>Sulfur relay protein, TusE/DsrC/DsvC family</td>
<td>P</td>
<td>0.03548 1.12914 31.8</td>
<td></td>
</tr>
<tr>
<td>SCD_01927</td>
<td>Glutaredoxin-like protein</td>
<td>O</td>
<td>0.03040 0.57336 18.9</td>
<td></td>
</tr>
<tr>
<td>SCD_00850</td>
<td>Flagellin and related hook-associated proteins</td>
<td>N</td>
<td>0.00107 0.01660 15.5</td>
<td></td>
</tr>
<tr>
<td>SCD_01247</td>
<td>FeS cluster assembly scaffold IscU</td>
<td>C</td>
<td>0.00835 0.08731 10.5</td>
<td></td>
</tr>
<tr>
<td>SCD_00930</td>
<td>DEAD/DEAH box helicase domain-containing protein</td>
<td>LKJ</td>
<td>0.00281 0.02504 8.9</td>
<td></td>
</tr>
<tr>
<td>SCD_01712</td>
<td>Hypothetical protein</td>
<td>S</td>
<td>0.17117 1.39561 8.2</td>
<td></td>
</tr>
<tr>
<td>SCD_02129</td>
<td>Anaerobic ribonucleoside-triphosphate reductase</td>
<td>F</td>
<td>0.00207 0.01552 7.5</td>
<td></td>
</tr>
<tr>
<td>SCD_01704</td>
<td>Hydrogenase maturation factor</td>
<td>O</td>
<td>1.45679 10.89136 7.5</td>
<td></td>
</tr>
<tr>
<td>SCD_02710</td>
<td>Fe(II) trafficking protein YggX</td>
<td>CO</td>
<td>0.01884 0.12361 6.6</td>
<td></td>
</tr>
<tr>
<td>SCD_01982</td>
<td>Nitroreductase</td>
<td>C</td>
<td>0.00247 0.01570 6.4</td>
<td></td>
</tr>
<tr>
<td>SCD_00192</td>
<td>Guanylate kinase</td>
<td>F</td>
<td>0.00909 0.05573 6.1</td>
<td></td>
</tr>
<tr>
<td>SCD_02164</td>
<td>Cytochrome c553</td>
<td>C</td>
<td>0.09928 0.60598 6.1</td>
<td></td>
</tr>
<tr>
<td>SCD_03038</td>
<td>Protein tyrosine phosphatase</td>
<td>X</td>
<td>0.00882 0.05151 5.8</td>
<td></td>
</tr>
<tr>
<td>SCD_00637</td>
<td>RNA chaperone (cspD: CspA family protein)</td>
<td>K</td>
<td>0.01951 0.10436 5.3</td>
<td></td>
</tr>
<tr>
<td>SCD_02624</td>
<td>Pilus protein</td>
<td>T</td>
<td>0.01971 0.10504 5.3</td>
<td></td>
</tr>
<tr>
<td>SCD_00876</td>
<td>Beta-hydroxacyl-(acyl-carrier-protein) dehydratase FabZ</td>
<td>I</td>
<td>0.00354 0.01772 5.0</td>
<td></td>
</tr>
</tbody>
</table>

aLocus tags indicate gene localizations in the first version of the draft genome BAFJ0100001- BAFJ01000023.

bOne-letter abbreviations indicate the COG functional category. cPC values were calculated excluding all ribosomal proteins.
Supporting information

**Table S1.** Genes for enzymes involved in sulfur oxidation, denitrification, carbon fixation in strain skB26 draft genome.

**Table S2.** List of all ribosomal proteins encoded in strain skB26 draft genome and their normalized protein contents (PC).

**Table S3.** List of all proteins identified only in cells cultured at 5°C. PC value was calculated except for all ribosomal proteins.

**Table S4.** All genes encoding sulfur relay proteins in strain skB26 draft genome and their normalized protein contents.
Figure 1

A

Normalized protein content (%)

10.6 (%) 7.1 (%) 8.5 (%)

(putative dsrC)

B

Normalized protein content (%)

1.8

1.6

1.4

1.2

1.0

0.8

0.6

0.4

0.2

0

dsr

sqr

arp

sat

hdr

sox

C

Normalized protein content (%)

1.8

1.6

1.4

1.2

1.0

0.8

0.6

0.4

0.2

0

nap

nir

nor

nos

Form I RubisCO  Form II RubisCO

Figure 1