Title
The Genome of the Epsilonproteobacterial Chemolithoautotroph Sulfurimonas dentificans

Permalink
https://escholarship.org/uc/item/36v0r7br

Authors
Sievert, Stefan M.
USF Genomics Class

Publication Date
2008-09-23
The genome of the epsilonproteobacterial chemolithoautotroph

*Sulfurimonas denitrificans*

Stefan M. Sievert‡*, Kathleen M. Scott‡*, Martin G. Klotz³, Patrick S. G. Chain⁴,⁵, Loren J. Hauser⁶, James Hemp⁷, Michael Hügler¹,⁸, Miriam Land⁶, Alla Lapidus⁵, Frank W. Larimer⁶, Susan Lucas⁵, Stephanie A. Malfatti⁴,⁵, Folker Meyer⁹, Ian T. Paulsen¹⁰#, Qinghu Ren¹⁰, Jörg Simon¹¹, and the USF Genomics Class²,Δ

Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts¹; Biology Department, University of South Florida, Tampa, Florida²; Departments of Biology and Microbiology & Immunology, University of Louisville, Louisville, Kentucky³; Lawrence Livermore National Laboratory, Livermore, California⁴; Joint Genome Institute, Walnut Creek, California⁵; Oak Ridge National Laboratory, Oak Ridge, Tennessee⁷; Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois⁷; Leibniz-Institut für Meereswissenschaften, Kiel, Germany⁸; Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois⁹; The Institute for Genomic Research, Rockville, Maryland¹⁰; Institute of Molecular Biosciences, Johann Wolfgang Goethe University, Frankfurt am Main, Germany¹¹

# Current address: Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, Australia

‡ SMS and KMS contributed equally to this work.

§ SMS and KMS contributed equally to this work.

Running title: *Sulfurimonas denitrificans* genome

Journal section: Evolutionary and Genomic Microbiology
Sulfur-oxidizing epsilonproteobacteria are common in a variety of sulfidogenic environments. These autotrophic and mixotrophic sulfur-oxidizing bacteria are believed to contribute substantially to the oxidative portion of the global sulfur cycle. In order to better understand the ecology and roles of sulfur-oxidizing epsilonproteobacteria, in particular the widespread genus *Sulfurimonas*, in biogeochemical cycles, the genome of *Sulfurimonas denitrificans* DSM1251 was sequenced. This genome has many features, including a larger size (2.2 Mbp), that suggest a greater degree of metabolic versatility or responsiveness to the environment than most of the other sequenced epsilonproteobacteria. A branched electron transport chain is apparent, with genes encoding complexes for the oxidation of hydrogen, reduced sulfur compounds, and formate, and the reduction of nitrate and oxygen. Genes are present for a complete, autotrophic reductive citric acid cycle. Many genes are present that could facilitate growth in the spatially and temporally heterogeneous sediment habitat from where *Sulfurimonas denitrificans* was originally isolated. Many resistance-nodulation-development-family transporter genes (11 total) are present, several of which are predicted to encode heavy metal efflux transporters. An elaborate arsenal of sensory and regulatory protein-encoding genes is in place, as well as genes necessary to prevent and respond to oxidative stress.
Epsilonproteobacteria are represented as molecular isolates from a vast array of habitats, including brackish, marine, and subsurface (e.g., (3, 11, 25, 39, 45, 71, 74); reviewed in (6)). Over the last years, quite a few cultured representatives of this group have been obtained from these environments, and currently all cultured members are chemolithoautotrophs or mixotrophs, capable of using sulfur and/or hydrogen as electron donors, while denitrifying and/or growing as microaerophiles (reviewed in (6)). These organisms use the reductive citric acid cycle for carbon fixation (24, 67). Given their abundance, sulfur-oxidizing epsilonproteobacteria, in particular members of the genus *Sulfurimonas*, are believed to be relevant to the function of the global sulfur cycle (6).

Genome data from these organisms would be key to metagenome sequencing efforts in habitats where they are abundant, and would also, by comparison to other epsilonproteobacteria, be helpful for determining the traits unique to a free-living, autotrophic lifestyle versus a host-associated, heterotrophic lifestyle. Recently, the genome sequences of *Sulfurovum* sp. NBC37-1 and *Nitratiruptor* sp. SB155-2, two sulfur-oxidizing epsilonproteobacteria from deep-sea hydrothermal vents were published, which revealed that these organisms share many features with their pathogenic (e.g., *Campylobacter* and *Helicobacter* spp.) epsilonproteobacterial relatives (41). Given the remarkable variety of habitats where sulfur-oxidizing epsilonproteobacteria are found, it was of great interest to also conduct these analyses on non-vent epsilonproteobacteria. To represent the abundant sulfur-oxidizing epsilonproteobacteria present in coastal marine sediments, we chose to sequence and analyze the genome of the sulfur-oxidizing chemolithoautotroph *Sulfurimonas denitrificans* DSM1251. Based on its phenotype, *S. denitrificans* was originally named *Thiomicrospira denitrificans* (70). Subsequent
sequencing revealed the polyphyletic nature of the *Thiomicrospiras*, with members from both the gammaproteobacteria and epsilonproteobacteria (40). As a result, *Thiomicrospira denitrificans* was eventually removed from the genus *Thiomicrospira* and placed within the genus *Sulfurimonas* (68). In addition to marine sediments, bacteria belonging to this genus have been isolated or detected in a variety of sulfidogenic environments, including deep-sea hydrothermal vents, the oxic-anoxic interface of marine anoxic basins, and oilfield (6, 18, 25, 32), making organisms of this genus globally significant.

**MATERIALS AND METHODS**

**Genome sequencing, annotation, and analysis.** DNA libraries were created and sequenced to approximately 13X depth of coverage at the Production Genomics Facility of the Joint Genome Institute (JGI) using the whole-genome shotgun method as previously described (7, 57). Gaps were closed and base quality problems were addressed by sequencing finishing reads, and PHRED/PHRAP/CONSED were used for assembly (12, 13, 20). Automated and manual annotation were conducted by ORNL similarly to (7, 57). Results were collated and presented via GenDB (37) for manual verification. The prediction of membrane transporters was based on a transporter annotation pipeline that uses several predictive approaches such as BLAST, COG, PFAM and TIGRFAM HMM searches, transmembrane topology prediction algorithms, and takes advantage of a curated database of transporters. Details of this pipeline and database have been published in (50, 51, 57). The main limitation of this approach is the
ability to accurately predict precise transporter specificities. Based on both internal and external testing, this methodology is highly successful at identifying putative transporters and predicting approximate substrate specificity. However, making precise substrate predictions, eg., serine transport rather than amino acid transport, is more problematic. The other related limitation is that the approach is dependent on comparison with known experimentally characterized transporters, so completely novel transporters which have never had homologues experimentally characterized will not be predicted by this methodology.

To uncover genes involved in oxidative stress, the *S. denitrificans* genome was examined with a series of BLAST queries, using genes known to be involved in oxidative stress response in *H. pylori* (73).

**Identification of genes encoding signal transduction and regulatory proteins.**

The complement of genes that encode signal transduction and regulatory proteins were compared among *S. denitrificans* DSM1251, *Thiomicrospira crunogena* XCL-2, and *Nitrosococcus oceani* ATCC 19707. To compare signal transduction and regulatory protein genes among these obligate autotroph genomes, genes were identified by querying the predicted gene products to the InterPro (PRINTS, PFAM, TIGRFAM, PRODOM and SMART) and COGs databases (via HMM search – InterPro, or RPSblast – COGs) to identify domains indicative of a role in these processes (e.g., EAL, GGDEF, PAS/PAC). Genes with predicted domains above the trusted cutoff score (InterPro) or an e value less than $e^{-5}$ (COGs) were assigned a product description and classified using a
set of rules based on the domain architecture of the protein. The final results were manually verified.

**Nucleotide sequence accession number.** The complete sequence of the *S. denitrificans* genome is available from GenBank (accession number NC_007575).

**RESULTS AND DISCUSSION**

**Genome structure.** The *S. denitrificans* DSM1251 genome is one of the largest epsilonproteobacterial genomes yet sequenced, consisting of a single 2.2 Mbp chromosome (Table 1). The coding density and G+C content are similar to the other epsilonproteobacteria (Table 1). Four rRNA operons are present, which, due to their elevated G+C content (~50%) relative to the genome average (34.5%), are visible as positive G+C content anomalies on the genome map (Fig. 1). Three of these operons (*16S-tRNA^{Ala}-23S-5S*) are 100% identical, and are oriented in the same direction, while the fourth (*16S-tRNA^{Ile}-23S-5S*) is in the opposite orientation, and its 5S and 23S genes each have a single nucleotide substitution compared to the others. The free-living *S. denitrificans*, *Sulfurovum* sp. NBC37-1, and *Nitratiruptor* sp. SB155-2 have more rRNA operons than those epsilonproteobacteria that are known to be exclusively host-related (Table 1), which is likely a reflection of an adaptation to fluctuating environmental conditions and the necessity for versatility (30, 41, 65).

Two large (17,627 bp), identical transposons are apparent as negative G+C content anomalies (30.0%; Fig. 1). Flanked by identical 12 bp inverted repeats, these
transposons (*Suden_0690 – 0702; Suden_1587 - 1599*) include genes encoding transposases, as well as proteins similar to the TniB (46%) and TniQ (47%) transposase accessory proteins found in mercury resistance-transposons in *Xanthomonas* sp. W17 and other systems (29). These transposons also include genes encoding a type I restriction-modification methyltransferase and restriction enzyme (see Supplemental Materials). Interestingly, one of the copies of this transposon interrupts a flagellin biosynthetic operon, which may explain why, unlike close relatives (25, 68), this strain of *S. denitrificans* is nonmotile (70). Other transposase and integrase genes are described in Supplemental Materials.

**Transporters.** *S. denitrificans* has a modest complement of genes (97 total) predicted to encode transporters. This number is similar to other sequenced heterotrophic epsilonproteobacteria (75-124 genes), which is surprising given that nutrient requirements for *S. denitrificans*, believed to be an obligate autotroph, are simple, compared to the others. This similarity in transporter numbers is due in part to the expansion of a few transporter families in this species compared to other epsilonproteobacteria. Amt-family transporters, are encoded in the *S. denitrificans* (*Suden_0641 and 0643*), *Sulfurovum* sp. NBC37-1 (2 copies), *Nitratiruptor* sp. SB155-2 (1 copy) and *Wolinella succinogenes* (1 copy) genome, but not in *Helicobacter pylori*, *Helicobacter hepaticus*, or *Campylobacter jejuni*, and are likely to facilitate ammonium uptake. Perhaps their absence in *Helicobacter* and *Campylobacter* spp. is due to nitrogen requirements for these species being met primarily from exogenous urea and/or amino acids (35, 54, 69). A FNT (formate-nitrite transporter) gene is present in *S. denitrificans*
(Suden_0716), and absent from other sequenced epsilonproteobacteria. Neither gene context nor sequence comparison clarifies the substrate for this transporter. Also notable is the presence of an abundance of Resistance-Nodulation-Cell Division (RND) Superfamily genes (11 genes; Suden_0270; Suden_0536; Suden_0799; Suden_0876; Suden_0877; Suden_0883; Suden_1281; Suden_1440; Suden_1499; Suden_1574; Suden_1575; Suden_2011), compared to other epsilonproteobacteria, including the two hydrothermal vent species (2-6 genes). Many of these are predicted to encode transporters for metal efflux in S. denitrificans. As in W. succinogenes and the two hydrothermal vent epsilonproteobacteria, an apparent operon is present that encodes a cytoplasmic arsenate reductase (Suden_0314), arsenite permease (Suden_0313), and regulatory protein ArsR (Suden_0315; (41, 61). Apparently the sediment ecosystem inhabited by S. denitrificans requires a similar or perhaps enhanced level of resistance to metals and other toxins, than the digestive tract habitats and hydrothermal vents favored by the other sequenced species.

Electron donors. S. denitrificans was originally isolated in a chemostat with thiosulfate as electron-donor and nitrate as electron-acceptor (70). However, prior to this study, the pathways and complexes involved were not identified. Neutrophilic sulfur-oxidizing bacteria use two types of sulfur oxidation pathways: one involving a multienzyme complex catalyzing the complete oxidation of reduced sulfur compounds to sulfate (Sox pathway) (15, 28), and another implementing sulfite and elemental sulfur as important intermediates (27, 47, 59). The genome of S. denitrificans reveals that the oxidation of reduced sulfur compounds proceeds via the Sox pathway (Fig. 2). Homologs for genes
encoding all components that are required for a fully functional complex in vitro, i.e.,
SoxB, SoxXA, SoxYZ, and SoxCD (15), could be identified. As in other obligate
sequenced autotrophs (41, 57), the sox genes in *S. denitrificans* do not occur in one
cruster, as in the model organism *Paracoccus pantotrophus* GB17 (15), but in different
parts of the genome. *S. denitrificans* has basically two clusters, one containing soxXYZAB
(*Suden_260-264*) and another one containing soxZYCD (*Suden_2057-2060*). SoxZY are
known to interact with both SoxAB and SoxCD and their duplication could possibly
indicate differential regulation of these two loci. SoxCD has homologies to sulfite
dehydrogenase (SorAB), but has been shown to act as a sulfur dehydrogenase (15). In
addition, it has recently been shown that organisms that lack soxCD, but do have soxB,
soxXA, and soxYZ use the Sox system to oxidize thiosulfate to sulfur, which is stored
either inside the cell or excreted (21). However, elemental sulfur formation by *S.
denitrificans* has not been reported. Recently, sulfur oxidation enzymes were also
measured in the closely related bacteria *Sulfurimonas autotrophica* and *Sulfurimonas
paralvinellae* (67). In this case, sulfite dehydrogenase was detected using an assay that
would not be expected to measure such activity were these organisms to use the Sox
system (C. G. Friedrich, pers. comm.), indicating that other *Sulfurimonas spp.* might
either not use the Sox system or use a modified version of it. In this regard it is
interesting to note that the SoxC sequence identities of *S. denitrificans* to sequences of
those organisms that have a contiguous sox gene set are significantly lower (44%) than
when SoxC sequences from organisms in which sox genes occur in one cluster are
compared among themselves (>63%). Both soxB and soxC genes exhibit highest
similarities with genes from *Sulfurovum* sp. NBC37-1 (41), which suggests that both
clusters of sox genes are not recent additions to this epsilonproteobacterial lineage. In fact, a phylogenetic analysis based on a large number of SoxB sequences from a variety of sulfur-oxidizing bacteria is even suggestive of an origin of the sox system in epsilonproteobacteria (36).

Besides the Sox system, S. denitrificans also has a gene encoding a sulfide:quinone oxidoreductase (SQR; Suden_0619). SQR catalyzes the oxidation of sulfide to elemental sulfur in Rhodobacter capsulatus (56), leading to the deposition of sulfur outside the cells. At present its role in S. denitrificans is unclear as this species has not been shown to deposit elemental sulfur, though this possibility has not been exhaustively explored with differing cultivation conditions.

The genome also provided evidence for the ability to also use H₂ and formate as electron donors (Fig. 2). Based on this information, S. denitrificans was successfully cultivated with H₂ as its electron donor and nitrate as the electron acceptor (Sievert and Molyneaux, unpublished data). The S. denitrificans genome encodes two Ni-Fe hydrogenase systems: one cytoplasmic, and one membrane-bound hydrogenase complex. The genes encoding the two subunits of the cytoplasmic enzyme (which lack TAT motifs; Suden_1437-8) are adjacent to genes encoding the periplasmic hydrogenase (its small subunit has a TAT motif, and a b-type cytochrome subunit would function to anchor it to the membrane and shunt electrons to the quinone pool; Suden_1434-6). The small subunit of the cytoplasmic hydrogenase of S. denitrificans forms a cluster with sequences from the two deep-sea hydrothermal vent epsilonproteobacteria and Aquifex aeolicus and is distantly related to H₂-sensing hydrogenases of alphaproteobacteria and cyanobacteria (41). Nakagawa et al. (41) suggest that the cytoplasmic hydrogenase acts
as a H₂-sensing hydrogenase in *Sulfitro Irving* NBC37-1 and *Nitratiruptor* SB155-2.

However, an alternative, and in our view more likely, function for the cytoplasmic enzyme as a catalytically active hydrogenase is suggested by the sequence similarity of both subunits to the enzyme from *A. aeolicus*. In *A. aeolicus*, the cytoplasmic hydrogenase can reduce electron acceptors with very negative redox midpoint potentials, and therefore has been suggested to provide electrons with low potential electrons to the reductive citric acid cycle (4), which would circumvent the necessity for reverse electron transport and thus increase its growth efficiency, similar to what has been found in certain Knallgas bacteria using the Calvin cycle for CO₂-fixation. Further experiments are needed to confirm the actual role of the cytoplasmic hydrogenase. Following these hydrogenase genes are several genes encoding hydrogenase-assembly related functions

A formate dehydrogenase complex is encoded in an operon similar in gene order to one found in *W. succinogenes* (*Suden_0816-24*). Formate dehydrogenase α subunits contain a selenocysteine residue (26) which is encoded by a stop codon. A putative selenocysteine codon (TGA) followed by a palindromic region was found between two open reading frames (ORFS) that are homologous to the amino and carboxy ends of formate dehydrogenase; accordingly, these ORFS have been combined into a single coding sequence for the α subunit of this enzyme, *Suden_0820*, which includes the molybdopterin-binding and iron-sulfur cluster domains typically found in this subunit, as well as the TAT-pathway signal sequence which would shunt this subunit to the periplasm. The β and γ subunits are encoded by *Suden_0819* and *Suden_0818*, respectively, with the latter having an unprecedented N-terminal addition with two more
predicted transmembrane segments (making it six in total). *Suden_0824* encodes a ferredoxin, which may shuttle the electrons from formate oxidation to cellular processes. Formate dehydrogenase maturation is likely facilitated by the products of *Suden_0823*, which encodes a TorD family protein that functions in molybdoprotein formation, and *Suden_0817*, which encodes a FdhD/NarQ family maturation protein. *Suden_0816* encodes a protein belonging to the aminotransferase class V PFAM, as does the SelA protein, which catalyzes a step in selenocysteinyl-tRNA synthesis. Other genes likely to be involved in selenocysteine synthesis are encoded nearby (*Suden_0830*: selenophosphate synthase; *Suden_0831*: L-seryl-tRNA selenium transferase; *Suden_0832*: selenocysteine-specific translation elongation factor SelB). Interestingly, the hydrothermal vent epsilonproteobacteria do not appear to have this complex. Though a homolog to the α subunit of formate dehydrogenase was present in both *Nitratiruptor* SB155-2 and *Sulfurovum* sp. NBC37-1, genes encoding the other subunits were not apparent from BLASTp searches of their genomes. In addition, *S. denitrificans* has a gene encoding the large subunit of a formate dehydrogenase H (*fdhF*, *Suden_1902*), which is most similar to the one in *W. succinogenes* (WS0126). However, its function and substrate are not apparent based on its sequence or genomic context.

**Electron acceptors.** All genes required for the complete reduction of nitrate to N\(_2\) are present (Fig. 2). However, *S. denitrificans* has some notable modifications compared to the canonical denitrification pathway. Similar to *Bradyrhizobium japonicum*, *S. denitrificans* appears to have only a periplasmic nitrate reductase (Nap) and not a cytoplasmic membrane-bound nitrate reductase (Nar) (2), which is present in most
organisms producing $\text{N}_2$ from nitrate (52). The \textit{nap} gene cluster (\textit{NapABHDFGLD}; \textit{Suden\_1514-1519, 1521}) has the same arrangement as the one identified in \textit{W. succinogenes}, which is unusual in that it lacks a gene encoding the NapC subunit (64). Possibly, \textit{S. denitrificans} is able to denitrify under microaerobic conditions, as the Nap enzyme has been implicated in aerobic denitrification (38). Nitrite reduction to nitric oxide is likely catalyzed by a cytochrome cd$_1$-dependent nitrite reductase (\textit{nirS, nirF}; \textit{Suden\_1985, 1988}), whose genes are present in a gene cluster (\textit{Suden\_1976-1989}) that also includes siroheme synthesis genes and two genes annotated as \textit{norCB} (\textit{Suden\_1983-1984}). The \textit{norCB} genes, which encode the small and large subunits of nitrous oxide-forming nitric oxide reductase (cNOR), a member of the heme-copper oxidase (HCO) superfamily (16), are usually clustered together with additional genes required for enzyme assembly and activation (76). These latter genes, \textit{norD} and \textit{norQ}, are missing from the \textit{S. denitrificans} genome. While there is evidence for functional cNOR in bacteria that lack the \textit{norQ/cbbQ} gene, there are presently no experimental reports that demonstrate functionality of cNOR in bacteria that also lack the \textit{norD} gene. Attempts to test this for \textit{Hydrogenobacter thermophilus} using \textit{Pseudomonas aeruginosa} or \textit{Escherichia coli} as an expression hosts were inconclusive (66). The \textit{S. denitrificans} \textit{norCB} genes are closest related by sequence similarity to the \textit{norCB} genes in the genomes of \textit{H. thermophilus}, \textit{Methylococcus capsulatus} as well as \textit{Sulfurovum} sp. \textit{NBC37-1} and \textit{Nitratiruptor} sp. SB155-2, the latter of which are two newly sequenced marine epsilonproteobacteria (41). All of these genomes also lack the \textit{norD} gene. Even though the functionality of cNOR is questionable, \textit{S. denitrificans} is a
complete denitrifier and must be able to reduce NO. Attempts to find other inventory implicated in NO reduction were successful and yielded additional candidate systems. One of them, NADH:flavorubredoxin-NOR, also known as the NorVW complex (17, 19), was also not complete and thus likely non-functional because a NorW-encoding gene was not identified. Interestingly, the genome encodes NorV in the unusual form of two genes: one encoding a rubredoxin (Suden_1582), which is succeeded by a flavodoxin gene (Suden_1581). Although both cNOR and NADH:flavorubredoxin-NOR may not have catalytic activity, it is possible that their NO-binding capacity has a function in NO sequestration and detoxification.

Interestingly, the *S. denitrificans* genome also encodes a previously unidentified member of the HCO superfamily that is also a candidate for catalyzing nitric oxide reduction. These HCO genes follow a set of pseudogenes normally involved in nitrate reduction (Suden_0100-0102). Based on structural modeling and genome analysis it is expected that this new HCO family is a novel non-electrogenic quinone-oxidizing nitric oxide reductase, gNOR (J. Hemp, M. G. Klotz, L. Y. Stein, R. B. Gennis, unpubl. data). The gNOR family, encoded by the norGHJ genes (Suden_0103-0105), is unique within the HCO superfamily in that it exhibits a novel active-site metal ligation, with one of the three conserved histidine ligands being replaced with an aspartate. This ligation pattern strongly suggests that the active-site metal is an iron. Structural modeling of members of the gNOR family has identified three conserved acidic residues, which form a charged pocket within the active site, a feature shared with the cNOR family (49). Besides *S. denitrificans*, gNOR also appears to be present in *Sulfovorum* sp. NBC37-1 and *Persephonella marina* strain EX-H1, whereas it is missing in *Nitratiruptor* sp. SB155-2.
Since phylogenetic analysis demonstrates that nitric oxide reductase activity has evolved multiple times independently within the heme-copper superfamily, these shared features between the distantly related gNOR and cNOR families are interesting examples of convergent evolution (J. Hemp, M. G. Klotz, L. Y. Stein, R. B. Gennis, unpubl. data).

Nitrous oxide reduction is carried out by nitrous oxide reductase encoded by an unusual nos gene cluster (Suden_1298-1308) similar to one previously identified in W. succinogenes (63). As in Wolinella, the NosZ in S. denitrificans (Suden_1298) contains a C-terminal extension of about 200 residues that carries a monoheme cytochrome c binding motif (CXGCH), suggesting it, too, functions as a cytochrome c nitrous oxide reductase (cNOS; Fig. 2). This feature is also shared by NosZ of Sulfurovum sp. NBC37-1 and Nitratiruptor sp. SB155-2, which form a cluster with NosZ from S. denitrificans (Fig. S3), possibly allowing the design of primers to screen for the presence of denitrifying epsilonproteobacteria in the environment. It has been hypothesized that the nos gene cluster in W. succinogenes codes for proteins involved in an electron transport chain from menaquinol to cytochrome c nitrous oxide reductase (63), and it is likely that the same holds true for S. denitrificans, as well as Sulfurovum sp. NBC37-1 and Nitratiruptor sp. SB155-2. In addition, S. denitrificans has an almost identical copy of nosZ next to a c553-type monoheme cytochrome c (Suden_1770, 1769), but its function is at this point unknown. Interestingly, S. denitrificans also has a gene coding for a large subunit of a ferredoxin-nitrite reductase (nirB; Suden_1241), which could be involved in nitrite assimilation or detoxification (8). However, no gene coding for the small subunit was identified, raising questions about its function.

Additional electron acceptors are suggested by this organism’s gene complement.
Like some other delta- and epsilonproteobacteria, the *S. denitrificans* genome contains a cluster of four genes (*Suden_0081-0084*) that encode the FixNOQP proteins, which constitute a proton-pumping *cbb*$_3$-type cytochrome c oxidase, suggesting an ability to use oxygen as a terminal electron acceptor. This is somewhat unexpected, since *S. denitrificans* was originally described as an obligate denitrifier and is quite sensitive to oxygen (S. Sievert, unpubl. data). *Cbb3*-type cytochrome c oxidase complexes have extremely high affinities for oxygen (48), which might allow this organism to use oxygen as an electron acceptor under extremely low oxygen tensions, or alternatively, to scavenge oxygen to prevent poisoning. Another possibility might be the involvement of *cbb3*-type cytochrome c oxidase in the catalytic reduction of NO (14). It is interesting, that, in line with other epsilonproteobacteria, the genome does not contain genes encoding FixGHIJ, which are present in all other bacteria that express a *cbb*$_3$-type cytochrome c oxidase and are involved in assembly and maturation of the *cbb*$_3$-type cytochrome c oxidase complex (33, 44). The reasons for the absence of *fixGHIJ* in epsilonproteobacteria remain unknown, and it needs to be tested experimentally what the actual role of *cbb3*-type cytochrome c oxidase in *S. denitrificans* is.

Additionally, *S. denitrificans, Nitratiruptor* sp. SB155-2, and *Sulfurovum* sp. NBC37-1 may be able to use sulfur compounds as terminal electron acceptors. Genes *Suden_0498-0500* encode an Fe/S-protein, a NrfD-related membrane anchor (8 helices), and an unusually long molybdopterin-containing oxidoreductase that contains a twin-arginine translocation pathway signal, respectively; homologs of these three genes are also encoded adjacent to each other in the two hydrothermal vent epsilonproteobacteria. Together these proteins could form a periplasm-facing membrane-bound complex that is
most likely involved in the reduction of sulfur compounds like elemental sulfur, polsulfide, thiosulfate or tetrathionate. However, only a limited number of these molybdopterin-containing oxidoreductases have been characterized and it is presently not possible to infer substrate specificity from the primary structure. Further experiments are needed to confirm the substrate used by this complex.

**Carbon fixation and central carbon metabolism.** Genes encoding the enzymes of the reductive citric acid cycle are apparent (Fig. 3), which is consistent with prior biochemical and genetic analyses of this organism (24). Of particular note are genes encoding the enzymes necessary for the cycle to operate in the reductive direction: pyruvate:acceptor oxidoreductase (Suden_0096 – 0099, based on similarity to biochemically characterized orthologs in *H. pylori* (23)), pyruvate carboxylase (Suden_0622, Suden_1259, based on biochemically characterized orthologs from *C. jejuni* (72), but see below), 2-oxoglutarate:ferredoxin oxidoreductase (Suden_1052 – 55, as for pyruvate:acceptor oxidoreductase, based on (23)), and ATP-dependent citrate lyase (Suden_0570, Suden_0571) (24). Also noteworthy is the presence of two copies of succinate dehydrogenase/fumarate reductase, one of which has a subunit that would anchor it to the membrane (Suden_1028 - 1030), while the other lacks this subunit and may be cytoplasmic (Suden_0037, Suden_0038). The membrane-bound form is unusual in that it has a cysteine-rich, type-E membrane anchor. It is similar to SdhABE from *W. succinogenes* which has been characterized as a membrane-bound fumarate-reducing complex with subunits SdhAB facing the periplasm. In contrast to *W. succinogenes*, the *S. denitrificans* gene encoding subunit A is about 43 residues shorter and lacks the TAT
signal peptide present in its *W. succinogenes* homolog, something that appears to be shared with *Sulfurovum* sp. NBC37-1 and *Nitratiruptor* sp. SB155-2, both of which are also chemolithoautotrophic epsilonproteobacteria using the reductive citric acid cycle for carbon fixation. Thus, it appears that in these organisms the membrane-bound fumarate-reducing complex faces into the cytoplasm. At present the exact function of the two fumarate reductases is unknown, although it is likely that the membrane-bound one, due to its potential for additional energy generation, might be involved in the reductive citric acid cycle for autotrophic carbon fixation. The intriguing possibility that *S. denitrificans* might also be able to carry out fumarate respiration has to await further experimentation. However, *S. denitrificans* does not contain a *frdCAB* operon typical for epsilonproteobacteria.

The acetyl-CoA and oxaloacetate produced by the reductive citric acid cycle could be funneled to central carbon metabolism: acetyl-CoA could be converted to pyruvate via pyruvate:acceptor oxidoreductase (see above) and oxaloacetate could be used to form phosphoenolpyruvate via phosphoenolpyruvate carboxykinase (*Suden_1696*). Acetyl-CoA could also be directed to fatty acid synthesis (acetyl-CoA carboxylase; *Suden_1174; Suden_1608*). Genes are present that suggest an ability to supplement autotrophic growth with acetate assimilation in all three sulfur-oxidizing epsilonproteobacteria. Two possible systems for converting acetate to acetyl-CoA are present: acetyl-CoA ligase (*Suden_1451*), as well as phosphate acetyltransferase (*Suden_0055*) and acetate kinase (*Suden_0056*), and are also present in *Sulfurovum* sp. NBC37-1, while acetate kinase is absent in *Nitratiruptor* sp. SB155-2. Perhaps the two
systems have different affinities for acetate, as has been demonstrated in methanogens (60), and are differentially expressed depending on environmental concentrations of this organic acid.

In order for S. denitrificans to grow autotrophically using the reductive citric acid cycle, there must be a means of carboxylating pyruvate to form oxaloacetate. In some organisms this is accomplished by the tandem activities of phosphoenolpyruvate synthetase and phosphoenolpyruvate carboxylase, while others use pyruvate carboxylase (55). In contrast to the two autotrophic epsilonproteobacteria Sulfurovum NBC37-1 and Nitratiruptor SB155-2, which have genes encoding phosphoenolpyruvate synthase as well as pyruvate kinase (41), the latter of which is usually involved in ATP synthesis during glycolysis, S. denitrificans does not appear to have any genes that might encode an enzyme that could interconvert phosphoenolpyruvate and pyruvate. Instead, it may use pyruvate carboxylase, as genes encoding both the biotin carboxylase subunit (Suden_0622) and biotin carboxyl carrier subunit (Suden_1259) of this enzyme are present. The biotin carboxyl carrier subunit gene (Suden_1259) occurs in an apparent operon with other genes homologous to the subunits of sodium-transporting oxaloacetate decarboxylase (Suden_1258 – 60). Suden_1259, which encodes the α subunit of this complex, has a high level of sequence similarity with pyruvate carboxylase genes from various Campylobacter species (including one from C. jejuni which has been biochemically characterized) (72), while the β and γ subunits (Suden_1258 and 1260), which are absent from the heterotrophic epsilonproteobacteria, are similar to those found from Sulfurovum sp. NBC37-1, Nitratiruptor sp. SB155-2, many gammaproteobacteria, many Chlorobia, and Desulfotalea psychrophila, a deltaproteobacterial sulfate reducer.
Heterotrophic organisms that have this complex ferment citrate. After cleaving citrate to acetate and oxaloacetate, they use the oxaloacetate decarboxylase complex to couple the exothermic decarboxylation of this organic acid to the extrusion of sodium ions. For these organisms, the other genes necessary for citrate fermentation (e.g., citrate transporter) are encoded nearby (9). This is not the case in *S. denitrificans*. An alternative function for *Suden_1259* is suggested by phylogenetic analysis, which places it within a clade with the biochemically characterized pyruvate carboxylase from *C. jejuni* (Fig. 4) and separate from biochemically characterized oxaloacetate decarboxylase genes from *Klebsiella pneumoniae* and *Vibrio cholerae* (5). Other members of this clade include the genes from the *Chlorobia, Sulfurovum* sp. NBC37-1, and *Nitratiruptor* sp. SB155-2, which also use the reductive citric acid cycle for carbon fixation and are not known to ferment citrate. It is possible to operate the oxaloacetate decarboxylase complex as a pyruvate carboxylase by imposing a sodium gradient across the membrane (10). It is tempting to speculate that in the autotrophic epsilonproteobacteria, the *Chlorobia*, and possibly *D. psychrophila*, this complex functions as a pyruvate carboxylase. Interestingly, the sequenced autotrophic epsilonproteobacteria and *Desulfotalea psychrophila* are all marine organisms, and the *Chlorobia* evolved in the marine environment (1). Only five of the ten sequenced *Chlorobia* have the sodium-transporting oxaloacetate decarboxylase/pyruvate carboxylase complex, while the other five, including *C. tepidum*, have the alpha subunit (on which Fig. 4 is based), but not the three-subunit pump. With the exception of *C. limicola*, which was isolated from a mineral hot spring, all *Chlorobia* containing the sodium pump have a requirement for sodium. Thus, it is likely that these organisms have found a way to couple pyruvate carboxylation, which is energetically unfavorable, to a
sodium gradient, something that was previously proposed, but never shown in an organism (10). Whether the complex encoded by *Suden_1258-1260* functions as a pyruvate carboxylase or an oxaloacetate decarboxylase is a key point begging clarification, which will be nontrivial, given that a genetic system has not been developed in this organism.

Carbon fixed by the reductive citric acid cycle can be shunted through gluconeogenesis, as all genes necessary for this pathway are present. The presence of genes encoding citrate synthase (*Suden_2100*) and ATP-dependent (irreversible) phosphofructokinase (*Suden_0549*) are enigmatic as their roles are unclear in this obligate autotroph.

Genes are apparent whose products could utilize the carbon skeletons synthesized by central carbon metabolism for ammonia assimilation (see below) and amino acid, nucleotide, fatty acid, and phospholipids synthesis. Cysteine biosynthesis is notable, in that the reduction of sulfate proceeds via adenosine 5′-phosphosulfate (APS) rather than 3′-phosphoadenylylsulfate (PAPS) in a pathway that was until recently only known from plants (42). Genes encoding assimilatory sulfate reduction co-occur in an apparent operon (*Suden_0154-0160*). Most likely this operon-like structure is turned on or off depending on whether *S. denitrificans* is inhabiting an environment with a high concentration of reduced inorganic sulfur compounds. Interestingly, sulfate assimilation in *Sulfurovum* sp. NBC37-1 and *Nitratiruptor* SB155-2 appears to proceed via PAPS.

**Nitrogen assimilation.** The *S. denitrificans* genome contains *nirC* (*focA*) and *nirB* genes encoding the formate-nitrite transporter (*Suden_0716*) and the large subunit of
NAD(P)H-dependent ammonia-forming siroheme nitrite reductase (*Suden_1241*), respectively, along with the inventory for siroheme synthesis (*Suden_1977, cysG, Suden_1988, cobA-cysG, uroporphyrinogen III methylase); however, it lacks the *nirD* gene, which encodes the small subunit of siroheme nitrite reductase. Because the genome also lacks *nrfHA* genes, which encode respiratory nitrite ammonification capacity in many delta- and epsilonproteobacteria (62), it appears that *Sulfurimonas* is solely dependent on ammonia uptake from the environment. The genome contains, indeed, two genes encoding different ammonia permeases (noted above) (22, 43), one AmtB-like (*Suden_0641*) and one Rhesus factor-like (*Suden_0643*) permease, which are clustered together with the gene encoding nitrogen regulatory protein PII (glnK, *Suden_0642*). Whereas AmtB proteins function as ammonia gas uptake channels, the substrate for Rh-like protein channels is still debated and includes ammonia as well as CO₂ (43). The genome contains also all the additional genes needed for 2-oxoglutarate-sensing and regulation of nitrogen assimilation (reviewed in (34)).

**Chemotaxis and other regulatory and signaling proteins.** Close relatives of *S. denitrificans* are motile, while this particular strain is nonmotile, probably due to the interruption of a flagellar biosynthetic operon by a transposon (see Genome Structure, above). Based on the presence of all of the genes necessary to encode the flagellar apparatus, none of which display any evidence of degeneration, an abundant sensory apparatus necessary to detect the presence of chemoattractants or repellants, and communicate this information to the flagellar motor, as well as the sequence identity of this transposon with a duplicate in the genome (see above), it is likely that nonmotility is
a recently acquired property. Interestingly, many of the genes encoding the chemotaxis
components are in a large cluster with multiple kinases and response regulators (Fig. 5),
as in *Nitratiruptor* sp. SB155-2 (41), suggesting interconnectivity between the
chemotaxis and other signal transduction systems. Perhaps the original enrichment and
isolation procedure for this strain (in a chemostat) might have selected for a non-motile
strain.

The *S. denitrificans* genome encodes a relative abundance of signaling proteins. Particularly well-represented in these genomes are genes encoding proteins with EAL
and GGDEF domains (based on hits to PFAMS, 16 and 38 genes, respectively), which
likely function in the synthesis and hydrolysis of the intracellular signaling compound
cyclic diguanylate (53). Further, six proteins with PAS/PAC-domains are encoded, which
may function as redox sensors (75). The genomic repertoire of signaling and regulatory
genes was compared with two other free-living, obligate chemolithoautotrophs for which
these data are available (Table 2). Some features are similar to *Thiomicrospira crunodena*: both of these species have a relative abundance of signal transduction
proteins compared to *Nitrosococcus oceani* (31), which may be a response to more
spatially (sediments; *S. denitrificans*) or temporally (hydrothermal vents; *T. crunodena*)
heterogeneous habitats, compared to the open ocean (*N. oceani*). Both have a large
number of genes encoding methyl-accepting chemotaxis proteins (MCPs; Table 2).
Unlike *T. crunodena*, none of the MCPs from *S. denitrificans* are predicted to have
PAS/PAC domains that could bind redox-sensitive cofactors (57), but a gene encoding a
protein with a PAS/PAC domain is present in the large cluster of chemotaxis genes
described above (Fig. 5), suggesting potential communication between sensing cellular or
environmental redox conditions and the chemotactic apparatus. Another similarity between *T. crunogena* and *S. denitrificans* is an abundance of genes predicted to be involved with cyclic nucleotide signal transduction, and many of these are predicted to have EAL and/or GGDEF domains, indicating a role for cyclic diguanylate in intracellular signaling in this organism. Many of these predicted proteins also have PAS/PAC domains, as in *T. crunogena* (57).

Unique among the three species compared here, *S. denitrificans* has a relative abundance of signal transduction histidine kinases and an expanded complement of winged helix family two component transcriptional regulators (Table 2). Both *T. crunogena* and *N. oceani* can use a rather limited variety of electron donors and acceptors, compared to *S. denitrificans*. Perhaps this expansion in histidine kinases and transcriptional regulators coordinate expression of the complexes necessary for the oxidation of multiple electron donors (e.g., H₂, reduced sulfur compounds, formate), and reduction of multiple electron acceptors (O₂ and NOₓ).

**Oxidative stress.** *S. denitrificans* has reasonably elaborate defenses against oxidative stress, on par with what has been observed for pathogenic epsilonproteobacterium *H. pylori* (73). Initially this was considered surprising to us, because this nonpathogenic species does not have to elude the oxidative arsenal of a host immune system, nor, as a microaerophile capable of growth via denitrification, does it grow in the presence of high concentrations of oxygen. However, given the presence of enzymes with labile iron-sulfur clusters with irreplaceable roles in central carbon metabolism in this organism (e.g., pyruvate:acceptor oxidoreductase; *Suden_0096* –...
defenses are a necessary part of survival.

Several genes are present whose products could prevent the buildup of intracellular Fe^{2+} that can spur hydroxyl radical generation via the Fenton reaction. A gene is present that encodes the Fur protein (\textit{Suden\_1272}), which regulates iron uptake (14). A gene encoding iron-binding ferretin (\textit{Suden\_1760}) is also present.

Many enzymes to detoxify reactive oxygen and nitrogen species are encoded in this genome. An iron/manganese superoxide dismutase (\textit{Suden\_1129}) is present in this species, but not the other sulfur-oxidizing epsilonproteobacteria (41). This enzyme and ruberythrin (\textit{Suden\_0739}) could convert superoxide to hydrogen peroxide, and this superoxide could in turn be dispatched by catalase (\textit{Suden\_1323}), peroxiredoxins (\textit{Suden\_0132, Suden\_0630, Suden\_1778, Suden\_1803}), or cytochrome c peroxidase (\textit{Suden\_0214; Suden\_1585}). Peroxiredoxins, particularly alkylhydroperoxide reductases, might be specifically targeted towards low levels of H_{2}O_{2} or organic peroxides (58, 73).

To dispense with nitric oxide that escapes the periplasmic and membrane-associated respiratory nitrogen reduction complexes, a truncated hemoglobin gene is present (\textit{Suden\_0993}), which may convert nitric oxide to nitrate (46). Thioredoxins (\textit{Suden\_0342; Suden\_0501; Suden\_1867; Suden\_2099}) and thioredoxin reductase (\textit{Suden\_1869}) could funnel electrons to these oxidative stress proteins, as glutathione does not appear to play this role in this organism, since genes encoding glutathione synthetase or gamma-glutamate-cysteine ligase are absent.

Genes are apparent whose products could enable a cell to cope with damage inflicted by any reactive oxygen or nitrogen species (ROS and RNS) that escape cellular
defenses. Endonuclease III (Suden_0516) and IV (Suden_1835) and MutS (Suden_0755) could repair oxidative DNA damage. Methionine sulphoxide reductase (Suden_0012) and alkylhydroxide reductase (Suden_1778) could contend with any methionine residues or lipids that had been oxidized by interaction with ROS or RNS, while nitroreductases (Suden_0519; Suden_1158) could prevent oxidized cellular nitrogenous and other compounds from generating peroxide.

Conclusions. *S. denitrificans* has several unique features which differentiate it from the other epsilonproteobacteria that have been sequenced to date. It has a larger genome than most of the others, which likely provides the sensory, regulatory, and metabolic versatility necessary for survival in a habitat more heterogeneous than found in a metazoan host. For example, the numerous genes whose products have redox sensory domains likely function to position these cells in the redoxcline to enable them to obtain the electron donors and acceptors needed for growth. Furthermore, these cells are quite versatile with respect to electron donors and acceptors, as the genome data suggest a capability of using donors and acceptors beyond those based on cultivation studies. Although *S. denitrificans* has been isolated from coastal marine sediments, its genome shares many features with two recently described autotrophic deep-sea hydrothermal vent epsilonproteobacteria, including the potential to utilize a variety of redox substrates (hydrogen gas, reduced sulfur compounds, oxygen, and nitrate), its responses to oxidative stress and high metal content, and a genome size intermediate between the two. This suggests that while these habitats appear at first strikingly different, they require similar adaptations on the scale of the microbes. Several additional features, which are absent in
their hydrothermal vent relatives, are present in *S. denitrificans* that may be particularly valuable in the sediment habitat. Their formate dehydrogenase complex would enable *S. denitrificans* to utilize a major by-product of fermentation that would co-occur with it should sediment organic carbon loads be high. The presence of the additional oxidative stress protein (superoxide dismutase) may enable *S. denitrificans* to cope with diurnal shifts in sediment oxygen concentration, and several additional RND-family efflux pumps relative to hydrothermal vent epsilonproteobacteria suggest that survival in marine sediments requires a degree of versatility and defense against environmental insult beyond what is necessary at moderate temperatures at hydrothermal vents.

**ACKNOWLEDGMENTS**

This work was performed under the auspices of the United States Department of Energy by Lawrence Livermore National Laboratory, University of California, under contract W-7405-ENG-48. Genome closure was funded in part by a USF Innovative Teaching Grant (KMS). SMS received partial support through a fellowship from the Hanse Wissenschaftskolleg in Delmenhorst, Germany (http://www.h-w-k.de) and NSF grant OCE-0452333. KMS is grateful for support from the NSF grant MCB-0643713. MH was supported by a WHOI postdoctoral scholarship. M.G.K was supported, in part, by incentive funds provided by the UofL-EVPR office, the KY Science and Engineering
Foundation (KSEF-787-RDE-007), and the National Science Foundation (EF-0412129).

We further thank Marga Bauer and Frank-Oliver Glöckner for helpful advice and for sharing their annotation guidelines at an early stage of this project, as well as three anonymous reviewers for their insightful suggestions.
REFERENCES


FIGURE LEGENDS

FIG. 1. Map of the *Sulfurimonas denitrificans* DSM1251 genome. The two outer rings include protein-encoding genes, which are color-coded based on their membership in COG categories. Ring 3 depicts the deviation from the average G+C (%), while the innermost ring is the GC skew (=\([G-C]/[G+C]\)). R1, R2, R3, and R4 are ribosomal RNA operons (with their orientation indicated with arrows), and the two regions marked T are identical large transposons. The G+C and GC skew rings were calculated with a sliding window of 10,000 bp with a window step of 100.

FIG. 2. Model for electron transport in *Sulfurimonas denitrificans*. Abbreviations: I—NADH dehydrogenase; II—succinate dehydrogenase/fumarate reductase; III—bc$_1$ complex; IV—cbb$_3$ cytochrome c oxidase; Cyt$_n$ c—cytochrome c; CCP—cytochrome c peroxidase; FDH—formate dehydrogenase; HYD—hydrogenase; MK—menaquinone; MKH$_2$—menaquinol; Mo—molybdodenum containing cofactor; NAP—periplasmic nitrate reductase; NIR—nitrite reductase; NOR—nitric oxide reductase; NOS—nitrous oxide reductase; PM—plasma membrane; PSR—polysulfide reductase; SOX—sulfur oxidation system. Candidate monoheme cytochromes c listed as “Cyt$_x$ c” are *Suden_0904, Suden_0741, Suden_0578* (all COG2863) as well as *Suden_0865, Suden_1329* and *Suden_1112* (no COG assignment).
FIG. 3. Central carbon metabolism in *Sulfurimonas denitrificans*. Abbreviations:

- 2-OG—2-oxoglutarate; acCoA—acetyl-Coenzyme A; ACL—ATP-citrate lyase;
- APFK—ATP-dependent phosphofructokinase; CS—citrate synthase; F6P—fructose 6-phosphate; FBP—fructose 1,6-bisphosphate; FBPP—fructose 1,6 bisphosphate phosphatase; FR—fumarate reductase; fum—fumarate; isocit—isocitrate; mal—malate;
- MQ—menaquinone; OAA—oxaloacetate; PC—pyruvate carboxylase; PEP—phosphoenolpyruvate; pyr—pyruvate; suc—succinate; suCoA—succinyl-Coenzyme A.

FIG. 4. Phylogenetic relationships of alpha-subunits of oxaloacetate decarboxylase (OAD), pyruvate carboxylase (PVC), oxoglutarate carboxylase (OGC), and type III pyruvate carboxylase to the product of *Suden_1259* of *Sulfurimonas denitrificans*. Sequences were aligned using the program package MacVector. Neighbor-joining and Parsimony trees based on the predicted amino acid sequences were calculated using PAUP 4.0b10. Bootstrap values (1,000 replicates) for the major nodes are given for the neighbor-joining (first value) and parsimony analyses (second value).

FIG. 5. A large gene cluster from the *Sulfurimonas denitrificans* genome that includes many of the genes for chemotaxis signal transduction.
TABLE 1. Comparative genome features for epsilonproteobacteria‡

<table>
<thead>
<tr>
<th>Species</th>
<th>Size (Mbp)</th>
<th>% coding</th>
<th>%GC</th>
<th>rRNA operons</th>
<th># CDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sulfurimonas denitrificans DSM1251</td>
<td>2.20</td>
<td>93.8</td>
<td>34.5</td>
<td>4</td>
<td>2104</td>
</tr>
<tr>
<td>Sulfurovum sp. NBC37-1</td>
<td>2.56</td>
<td>90.1</td>
<td>43.8</td>
<td>3</td>
<td>2466</td>
</tr>
<tr>
<td>Nitratiruptor sp. SB155-2</td>
<td>1.88</td>
<td>95.1</td>
<td>39.7</td>
<td>3</td>
<td>1857</td>
</tr>
<tr>
<td>Campylobacter fetus 82-40</td>
<td>1.80</td>
<td>90.0</td>
<td>33.3</td>
<td>3</td>
<td>1719</td>
</tr>
<tr>
<td>Campylobacter jejuni NCTC 11168</td>
<td>1.64</td>
<td>95.4</td>
<td>30.6</td>
<td>3</td>
<td>1629</td>
</tr>
<tr>
<td>Campylobacter jejuni RM1221</td>
<td>1.78</td>
<td>91.8</td>
<td>30.3</td>
<td>3</td>
<td>1838</td>
</tr>
<tr>
<td>Helicobacter hepaticus ATCC 51449</td>
<td>1.80</td>
<td>93.4</td>
<td>35.9</td>
<td>1</td>
<td>1875</td>
</tr>
<tr>
<td>Helicobacter acinonychis Sheeba</td>
<td>1.55</td>
<td>89.0</td>
<td>38.2</td>
<td>2</td>
<td>1618</td>
</tr>
<tr>
<td>Helicobacter pylori26695</td>
<td>1.67</td>
<td>90.0</td>
<td>38.9</td>
<td>2*</td>
<td>1576</td>
</tr>
<tr>
<td>Helicobacter pylori J99</td>
<td>1.64</td>
<td>90.7</td>
<td>39.2</td>
<td>2*</td>
<td>1491</td>
</tr>
<tr>
<td>Helicobacter pylori HPAG1</td>
<td>1.59</td>
<td>91.0</td>
<td>39.1</td>
<td>2*</td>
<td>1544</td>
</tr>
<tr>
<td>Wolinella succinogenes DSM1740</td>
<td>2.11</td>
<td>94.5</td>
<td>48.5</td>
<td>3</td>
<td>2043</td>
</tr>
</tbody>
</table>

‡Data for all taxa, except for Sulfurivom sp. NBC37-1 and Nitratiruptor sp. SB155-2 were collated from the Integrated Microbial Genomes webpage and had been generated using consistent methodology. For Sulfurivom sp. NBC37-1 and Nitratiruptor sp.
sp. SB155-2, data were collected from (41), for which slightly different methodologies were used to identify coding sequences (CDS).

*In *H. pylori*, the 16S gene is not collocated with the 23S and 5S genes in an operon. Additionally, an orphan 5S sequence is found in strain 26695.
TABLE 2. Regulatory and signaling proteins of *Sulfurimonas denitrificans* and other obligate chemolithoautotrophs*

<table>
<thead>
<tr>
<th>Functional Description</th>
<th>S. denitrificans</th>
<th>T. crunogena</th>
<th>N. oceani</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcription/Elongation/Termination Factors</td>
<td>56</td>
<td>72</td>
<td>104</td>
<td></td>
</tr>
<tr>
<td>Two component transcriptional regulator, winged helix family</td>
<td>19</td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Signal Transduction proteins</td>
<td>146</td>
<td>128</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Chemotaxis Signal Transduction proteins</td>
<td>28</td>
<td>27</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Methyl-accepting chemotaxis proteins</td>
<td>13</td>
<td>14</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Non-Chemotaxis Signal Transduction</td>
<td>118</td>
<td>101</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Signal Transduction Histidine Kinase</td>
<td>36</td>
<td>17</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Cyclic nucleotide signal transduction</td>
<td>42</td>
<td>49</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>202</td>
<td>200</td>
<td>179</td>
<td></td>
</tr>
</tbody>
</table>
**Sulfurimonas denitrificans** DSM1251 is compared to gammaproteobacteria *Thiomicrospira crunogena* XCL-2 and *Nitrosococcus oceani* ATCC 19707.
Tmden_0963
winged-helix transcription regulator

histidine kinase

histidine kinase

diguanylate cyclase/phosphodiesterase
with PAS/PAC domain

response regulator receiver protein

methyl-accepting chemotaxis protein

hypothetical

cheA

cheR

histidine kinase

cheD

cheB

histidine kinase

conserved hypothetical protein

methyl-accepting chemotaxis protein

response regulator receiver histidine kinase

methyl-accepting chemotaxis protein
diguanylate cyclase/phosphodiesterase
with a response regulator receiver

histidine kinase

histidine kinase

response regulator receiver protein

methyl-accepting chemotaxis protein

methyl-accepting chemotaxis protein

Tmden_0990
Further details of genome structure. In addition to the large transposon interrupting one of the flagellar biosynthetic operons (Fig. S1), another transposase gene (*Tmden_1713*) is located near a tRNA<sup>Thr</sup> gene, adjacent to a hypothetical protein gene (*Tmden_1712*), whose 3’-end is 82% identical, at the nucleotide level, with two transposase genes located downstream (*Tmden_1724* and *Tmden_1725*). These two genes, and regions 5’ and 3’ of each (totaling 1302 bp apiece), are 100% identical to each other. This region also includes genes encoding a recombinase and phage integrase (*Tmden_1723*; Fig. S2); the presence of the phage integrase gene, identical repeats, and juxtaposition to a tRNA gene suggest that this portion of the genome may be a remnant of a degraded or partially excised prophage. Other potential transposase genes are present (*Tmden_0961*; *Tmden_1698*; *Tmden_1708*), but have insufficient sequence similarity to known proteins for deducing their function convincingly.

Twelve phage integrase genes are present. As expected for phage genes, six of these are near tRNA genes (*Tmden_0248, Tmden_0779, Tmden_0800, Tmden_1618, Tmden_1723, Tmden_1743*) which are common insertion sites for lysogenic phages (1), and three are flanked by clusters of genes encoding hypothetical proteins (*Tmden_1247; Tmden_1618; Tmden_1633*), which is consistent with the observation that many phage genes are unique and uncharacterized (2). Two phage integrase genes (*Tmden_0938; Tmden_0959*) flank genes encoding a type I restriction modification system gene cluster and are part of a larger region (bp 977850 – 1002764; 5 o’clock, Fig. 1), that have a negative G+C anomaly (31.1%). The remaining two are included in the large transposons described above (*Tmden_0693 and Tmden_1590*).
Another negative G+C anomaly, which also has a GC skew anomaly, is visible on the genome map at approximately 1 o’clock (bps 192095 – 210856; 30.5% G+C; Fig. 1). Although this region contains several genes encoding hypothetical proteins, as one might expect were it derived from a phage, it does not appear to include any transposase or integrase genes, nor does it include any repeated sequences that might suggest recent gene rearrangement in this region.

**Restriction-modification systems.** *S. denitrificans* has numerous restriction-modification (RM) systems encoded in its genome. Eleven DNA methyltransferase genes are present, and encode methyltransferases similar to those found in Type I (Tmden_0697, Tmden_0942, Tmden_1594); Type II (Tmden_0121; Tmden_0129; Tmden_0130; Tmden_0478; Tmden_0537; Tmden_1565; Tmden_1839; Tmden_1855) and Type III (Tmden_1355) restriction-modification systems. For 6 of these methyltransferases, genes encoding restriction enzymes are nearby (Type I: Tmden_0700; Tmden_0948; Tmden_1597; Type II: Tmden_128; Tmden_1854; Type III: Tmden_1350) and for two of them the genes appear to encode fused methylase/restriction enzymes (Tmden_0478, Tmden_0537)(5). Based on genome sequence data, such large numbers of RM systems are not unusual for epsilonproteobacteria: *Helicobacter pylori* has 24 RM systems (3), *C. jejuni* has 10, and *W. succinogenes* has 5 (4, 5). If active, perhaps in *S. denitrificans* these systems provide a robust defense against the introduction of phage and other ‘non-native’ DNA into the genome.
REFERENCES


**FIGURE LEGENDS.**

FIG. S1. Two large identical transposons from the *Sulfurimonas denitrificans* genome. Numbers indicate the position of the regions, in nucleotides, with respect to the origin of replication, and the arrows indicate the presence of the inverted repeat sequences at each end: $> = \text{TGT CATTTACAA;} < = \text{TTGTAATGACA}.$

FIG. S2. Map of a region from the *Sulfurimonas denitrificans* genome that includes a small repeated region. The duplicate copies of this repeat include the two adjacent transposase genes (shaded grey), while a third region with a high level of identity (82%) is included within a hypothetical gene upstream (also shaded grey). Numbers indicate the position of the regions, in nucleotides, with respect to the origin of replication.

FIG. S3. Phylogenetic relationships of NosZ from different bacteria and the archaeon *Pyrobaculum calidifontis*. Suden_1298 is part of a novel nos cluster previously identified in *Wolinella succinogenes* (63). All epsilonproteobacterial sequences have a C-terminal extension and contain a heme $c$-binding motif. The sequences from *Dechloromonas aromatica* and *Magentospirillum magnetotacticum* also have a (somewhat shorter) C-terminal extension, but are lacking a heme $c$-binding motif (63). Sequences were aligned using the program package MacVector. Neighbor-joining and Parsimony trees based on the predicted amino acid sequences were calculated using
PAUP 4.0b10. Bootstrap values (1,000 replicates) for the major nodes are given for the neighbor-joining (first value) and parsimony analyses (second value).
FIG. S1
FIG. S2
FIG. S3