

The genome sequence of the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough

John F Heidelberg^{1,2}, Rekha Seshadri¹, Shelley A Haveman³, Christopher L Hemme⁴, Ian T Paulsen^{1,5}, James F Kolonay¹, Jonathan A Eisen^{1,5}, Naomi Ward^{1,2}, Barbara Methe¹, Lauren M Brinkac¹, Sean C Daugherty¹, Robert T Deboy¹, Robert J Dodson¹, A Scott Durkin¹, Ramana Madupu¹, William C Nelson¹, Steven A Sullivan¹, Derrick Fouts¹, Daniel H Haft¹, Jeremy Selengut¹, Jeremy D Peterson¹, Tanja M Davidsen¹, Nikhat Zafar¹, Liwei Zhou¹, Diana Radune¹, George Dimitrov¹, Mark Hance¹, Kevin Tran¹, Hoda Khouri¹, John Gill¹, Terry R Utterback¹, Tamara V Feldblyum¹, Judy D Wall⁴, Gerrit Voordouw³ & Claire M Fraser^{1,6}

Desulfovibrio vulgaris Hildenborough is a model organism for studying the energy metabolism of sulfate-reducing bacteria (SRB) and for understanding the economic impacts of SRB, including biocorrosion of metal infrastructure and bioremediation of toxic metal ions. The 3,570,858 base pair (bp) genome sequence reveals a network of novel c-type cytochromes, connecting multiple periplasmic hydrogenases and formate dehydrogenases, as a key feature of its energy metabolism. The relative arrangement of genes encoding enzymes for energy transduction, together with inferred cellular location of the enzymes, provides a basis for proposing an expansion to the ‘hydrogen-cycling’ model for increasing energy efficiency in this bacterium. Plasmid-encoded functions include modification of cell surface components, nitrogen fixation and a type-III protein secretion system. This genome sequence represents a substantial step toward the elucidation of pathways for reduction (and bioremediation) of pollutants such as uranium and chromium and offers a new starting point for defining this organism’s complex anaerobic respiration.

Sulfate-reducing bacteria (SRB) are anaerobic prokaryotes found ubiquitously in nature. SRB were the first nonphotosynthetic, anaerobic bacteria shown to generate energy (ATP) through electron transfer-coupled phosphorylation. For this process, the SRB typically use sulfate as the terminal electron acceptor for respiration of hydrogen or various organic acids, which results in the production of sulfide, a highly reactive and toxic end-product. Beyond their obvious function in the sulfur cycle, SRB play an important role in global cycling of numerous other elements¹. For example, in the carbon cycle, the SRB form part of microbial consortia that completely mineralize organic carbon in anaerobic environments; polymeric materials (e.g., cellulose) are first depolymerized and metabolized by fermentative microorganisms, and the resulting organic acid and reduced gas (that is, CO and H₂) end-products are further fermented or oxidized by other microbes, including SRB. The latter are particularly active in sulfate-rich (e.g., marine) environments, where they effectively link the global sulfur and carbon cycles^{1,2}.

Beyond these ecological roles, SRB also have a major economic impact because of their involvement in biocorrosion of ferrous metals

in anaerobic environments³, described as “industrial venereal disease—it’s expensive, everybody has it, and nobody wants to talk about it”⁴. For example, because SRB are abundant in oil fields, their metabolism has many negative consequences for the petroleum industry (e.g., corrosion of drilling and pumping machinery and storage tanks, souring of oil by sulfide production, plugging of machinery and rock pores with biomass and sulfide precipitates). The SRB also contribute to bioremediation of toxic metal ions^{5,6}. Their metabolism increases the pH, causing toxic metal ions like copper (II), nickel (II) and cadmium (II) to precipitate as metal sulfides in acidic aquatic environments (e.g., mine effluents). Additionally, SRB can deliver electrons directly to oxidized toxic metal ions, including uranium (VI), technetium (VII), and chromium (VI), converting these into less soluble, reduced forms. Hence, SRB-mediated reduction represents a potentially useful mechanism for the bioremediation of metal ion contaminants from anaerobic sediments⁶.

Most research on the metabolism and biochemistry of SRB has been done on the genus *Desulfovibrio*, a member of the δ -proteobacteria^{7,8}. Here we report the genome sequence of *D. vulgaris* Hildenborough, a

¹The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA. ²The Center for Marine Biotechnology, 701 East Pratt Street, Baltimore, Maryland 21202, USA. ³Department of Biological Sciences, University of Calgary, 2500 University Dr. NW, Calgary, Alberta T2N 1N4, Canada. ⁴Biochemistry Department, University of Missouri-Columbia, 117 Schweitzer Hall, Columbia, Missouri 65211, USA. ⁵Johns Hopkins University, Charles and 34th Streets, Baltimore, Maryland 21218, USA. ⁶George Washington University Medical Center, 2300 I Street NW, Washington, DC 20037, USA. Correspondence should be addressed to J.F.H. (dvo@tigr.org).

Table 1 General features of the *D. vulgaris* Hildenborough genome

	Chromosome	Megaplasmid
Size (base pairs)	3,570,858	202,301
G+C percentage	63.2	65.7
Predicted protein-coding genes (CDS)		
No. similar to known proteins	1,894	96
No. similar to proteins of unknown function ^a	372	13
No. of conserved hypothetical proteins ^b	273	12
No. of hypothetical proteins ^c	856	31
Total CDSs	3,395	152
Average CDS size (bp)	908	1,131
Percentage coding	86.4	85.0
Number of rRNA operons	5	0
Number of tRNA	68	0
Phage ^d	4	0
IS transposases	26	2

^aUnknown function. Substantial sequence similarity to a named protein for which no function is currently attributed. ^bConserved hypothetical protein. Sequence similarity to a translation of another open reading frame; however, no experimental evidence for protein expression exists.

^cHypothetical protein. No substantial similarity to any other sequenced protein. ^dTwo copies of nearly identical mu-like bacteriophage (DVU0189–221 and DVU2847–79) are present in the chromosome. DVU2688–733 represent a third lambdoid bacteriophage and DVU1498–1508 are remnants of a bacteriophage genome.

well-studied strain of this genus. The availability of this genome sequence greatly expands our understanding of its energy transduction and electron transport mechanisms. This paves the way for elucidation of the mechanisms by which SRB contribute to metal ion bioremediation and biocorrosion, as well as of their key roles in biogeochemical cycles.

General genome features

The *D. vulgaris* Hildenborough genome was sequenced by the whole genome sequencing method. Genome features are listed in Table 1. The *D. vulgaris* plasmid, known to contain the *nif* genes, can be lost when the organism is cultivated in ammonium-containing media (G. Voordouw, unpublished observation). The plasmid lacks homologs to previously characterized plasmid replication or partitioning genes. In addition to the genes for MoFe nitrogenase, the plasmid encodes all essential components of a type III secretion apparatus (DVUA0106–22). Type III secretion systems have typically been associated with pathogens and symbionts (e.g., *Rhizobium* spp. and *Yersinia* spp.⁹) where they translocate bacterial proteins (effectors) into the eukaryotic cells across three biological membranes to influence host responses. Preliminary analysis suggests the presence of putative effectors of this type III secretion system¹⁰, but it will be interesting to see which of these putative effector genes are authentic and determine the conditions under which their products are secreted.

Much of our historical knowledge on SRB physiology derives from a desire to understand the bacteria's role in microbially influenced corrosion, which may involve electron transport from the metal surface to the SRB sulfate reduction pathway through a hydrogen intermediate³. The genome sequence reveals multiple candidate hydrogenases and cytochromes that may be involved in the removal of hydrogen from metals through electron transfer. It has been postulated that microbial metabolism may indirectly contribute to the accumulation of corrosive sulfide and organic acid end-products causing localized pitting of metals. Maximum corrosion activity occurs at sites of intermittent oxygenation, thus supporting the conclusion that corrosion by these anaerobes occurs along an oxygen gradient, directly or

through a series of redox intermediates³. The genome sequence indicates that *D. vulgaris* has a large family of 27 methyl-accepting chemotaxis proteins, including the proteins DcrA and DcrH¹¹, which sense oxygen or redox potential and may be important in positioning the SRB in the gradients generated across the oxic/anoxic interface. The roles of these functions in stimulating microbially influenced corrosion can now be tested.

Energy metabolism

Desulfovibrio spp. use hydrogen, organic acids or alcohols as electron donors for sulfate reduction. Although many key components of the *Desulfovibrio* energy metabolism are known, the genome sequence indicates many novel features, including the existence of additional periplasmic *c*-type cytochromes and formate dehydrogenases as well as previously unknown cytoplasmically-oriented, membrane-bound hydrogenases.

Hydrogen oxidation/sulfur reduction. During chemolithotrophic growth, *D. vulgaris* derives energy only from oxidative phosphorylation by coupling the reduction of sulfate (or other sulfur oxyanions, e.g. sulfite or thiosulfate) to sulfide with the oxidation of hydrogen (Fig. 1 and Supplementary Table 1 online). The first step in the process is periplasmic hydrogen oxidation (Fig. 1 and Supplementary Table 1 online). Although three *D. vulgaris* periplasmic hydrogenase complexes were previously described¹², genome analysis revealed a fourth (HynBA-2 isozyme). The exact roles of these hydrogenases have not yet been established; each of them may have a different specificity or this redundancy may allow compensation under conditions of stress. For example, deletion of the genes for Fe-only hydrogenase impairs chemolithotrophic growth¹³, suggesting at least some redundancy in function.

The electrons generated by periplasmic hydrogen oxidation are possibly stored within the periplasm in multiheme *c*₃-type cytochromes until they are passed through the inner membrane via electron shuttles (Fig. 1 and Supplementary Table 1 online). The tetrahemic cytochrome *c*₃ (DVU3171) is generally regarded as the primary electron acceptor from periplasmic hydrogen oxidation and accounts for the majority of the *c*-type cytochromes of the periplasm¹⁴. However, genome analysis indicates the presence of three alternate *c*₃-type cytochromes (DVU2524, DVU2809 and DVU0263); localization of two of these (DVU2524 and DVU2809) suggests roles in dedicated pathways. Specifically, DVU2524 may serve as a dedicated electron acceptor for HynBA-2 (NiFe-hydrogenase isozyme 2), and DVU2809 is part of an operon encoding a formate dehydrogenase, suggesting that this tetraheme cytochrome *c*₃ accepts electrons arising from formate oxidation (see below).

Intermolecular electron transfer between the various *c*-type cytochromes creates a vast network of interconnected hemes¹⁵. This network likely provides the electrical wiring for connecting multiple periplasmic redox proteins and may also serve as a capacitor for storage of low-potential electrons originating from hydrogen or formate oxidation. Furthermore, the cytochrome network provides electrons for the reduction of metal ions, (Fig. 1 and Supplementary Table 1 online) Cr(VI) or U(VI). Historically, electrons for metal ion reduction were believed to be transferred solely by cytochrome *c*₃ (DVU3171). When the corresponding gene in *Desulfovibrio desulfuricans* G20 was knocked out, over 90% of the reduction of U(VI) from externally supplied H₂ was blocked; however, there was only a 50% decrease in the reduction of U(VI) with lactate or pyruvate as the electron donor⁶. Therefore, although this cytochrome is involved in metal ion reduction, the alternative tetraheme cytochromes likely play a role as well.

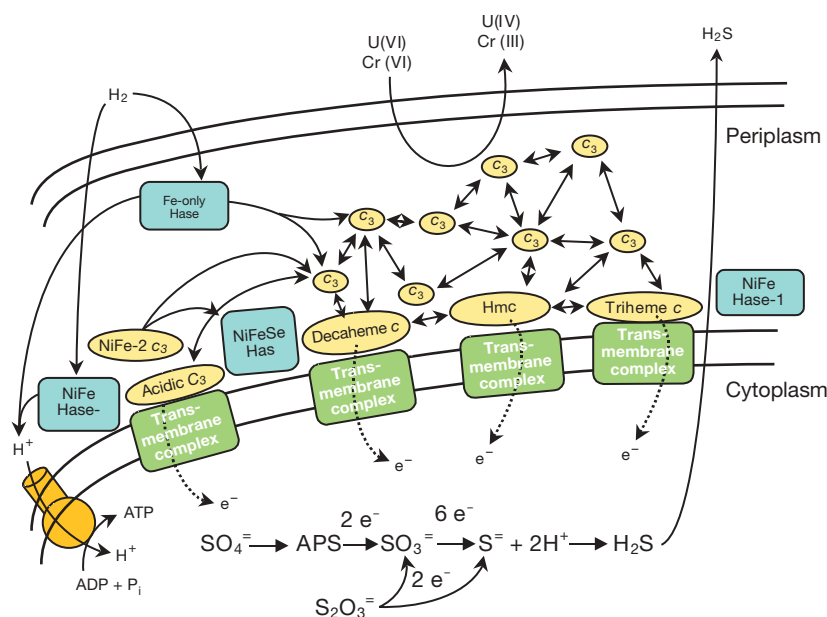


Figure 1 *D. vulgaris* c-type cytochrome network. Diagrammatic view of the c-type cytochrome network potentially present in the periplasm of *D. vulgaris* Hildenborough, and the associated periplasmic hydrogenases and transmembrane complexes. Multiple pathways are predicted for electrons from hydrogen supplied externally or produced in the cytoplasm. The protons generated from hydrogen oxidation could then be used to drive ATP synthesis through the F₁F₀ ATP synthase (DVU0774–80) pictured in the cytoplasmic membrane (far left). The electrons generated from hydrogen oxidation are transferred into the c-type cytochrome network for delivery through the cytoplasmic membrane via membrane-bound electron carriers for reduction of the terminal electron acceptors sulfate or thiosulfate. For locus identification of genes, see **Supplementary Table 1** online.

Hydrogenase activity and electron storage occur in the periplasm whereas sulfate reduction occurs in the cytoplasm, therefore, electrons must cross the cytoplasmic membrane. Before genome analysis, the Hmc complex (DVU0531–36, containing a hexadeca-heme c-type cytochrome) represented the only known *D. vulgaris* transmembrane electron circuit^{15–17}. However, genome analysis indicates the presence of four alternate transmembrane electron conduits (Fig. 1 and **Supplementary Table 1** online). The presence of multiple conduits provides an explanation for the observation that deletion of the *hmc* operon only partially inhibits growth of *D. vulgaris* in the presence of hydrogen and sulfate¹⁸.

In the cytoplasm, *D. vulgaris* uses sulfate as the principal terminal electron acceptor, generating sulfide, in a process designated as dissimilatory sulfate reduction (Fig. 1 and **Supplementary Table 1** online). In summary, the genome sequence adds detail to our understanding of *D. vulgaris* growth on hydrogen and sulfate through physically separated hydrogen oxidation and sulfate reduction reactions, resulting in a proton gradient that is consumed by the F₁F₀-ATP synthase complex (DVU0774–80)¹⁹.

Carbon metabolism and hydrogen cycling. *D. vulgaris* belongs to the group of incompletely-oxidizing sulfate reducers, where acetic acid is often an end-product of organic acid and alcohol oxidations providing electrons for sulfate reduction (Fig. 2 and **Supplementary Table 2** online). The preferred substrate, lactate, is oxidized to pyruvate by a membrane-bound lactate dehydrogenase (DVU0600) and pyruvate is further oxidized primarily by pyruvate:ferredoxin oxidoreductase (DVU3025, POR)²⁰. However, there are genes encoding a number of alternative oxo-organic acid oxidoreductases that could react with pyruvate (DVU1569–70, DVU1944–47, DVU1950–51, DVU3347–50

and DVU0374). Enzymatic action of phosphate acetyltransferase (DVU3029) and acetate kinase (DVU3030) result in substrate-level ATP synthesis coupled to the conversion of acetyl-CoA to CoA and acetate. Thus, *D. vulgaris* can ferment pyruvate, but the thermodynamics of lactate oxidation to pyruvate with hydrogen production is highly endergonic^{13,21}.

A chemiosmotic model for energy transduction from lactate oxidation termed “hydrogen cycling” has been proposed²²: the protons and electrons produced in lactate and pyruvate oxidation react with a cytoplasmic hydrogenase to form hydrogen, which diffuses across the membrane where it is reoxidized by periplasmic hydrogenases to form a proton gradient (Fig. 2 and **Supplementary Table 2** online). This proton gradient is used for additional ATP synthesis. This model requires both periplasmic and cytoplasmic hydrogenase isozymes. Whereas periplasmic hydrogenases have been demonstrated^{12,23}, cytoplasmically oriented hydrogenases have remained controversial. The genome sequence indicates the presence of genes for two membrane-bound hydrogenases, (*echABCDEFGHIJ*²⁴ and *cooMKLXUHF*^{13,25}), which both have active sites facing the cytoplasm. Either of these could be involved in H₂ production from lactate. Additionally, *cooMKLXUHF* could be involved, together

with CO dehydrogenase (DVU2098), in hydrogen production from CO according to the overall equation $\text{CO} + \text{H}_2\text{O} \rightarrow \text{CO}_2 + \text{H}_2$ (ref. 13). In any case, the hydrogen produced could then be captured by periplasmic hydrogenases to form a proton gradient. Thus, the genome sequence indicates the presence of novel complexes that are involved in energy transduction (e.g., ‘hydrogen cycling’) for which the exact mechanism remains to be worked out^{13,26}.

The mechanism of energy conservation associated with formate oxidation is more straightforward than for lactate (Fig. 2 and **Supplementary Table 2** online). Electrons from formate oxidation are probably donated to the periplasmic tetraheme cytochrome network and used for cytoplasmic sulfate reduction after import through the multiheme-containing transmembrane electron conduits (Fig. 1 and **Supplementary Table 1** online). In contrast to the cellular organization of hydrogenases and formate dehydrogenases in *Escherichia coli*^{27,28}, the operon structure of the *D. vulgaris* formate dehydrogenases suggests that the subunits are not membrane-bound and that they interact with the cytochrome network instead. The multiple periplasmic formate dehydrogenases may be specialized for the oxidation of formate that diffuses into the periplasm from external sources, or of formate that diffuses into the periplasm from the cytoplasm. The latter might be part of a mechanism to form a proton gradient from lactate oxidation through formate ‘cycling’ analogous to the model for hydrogen ‘cycling’ (Fig. 2 and **Supplementary Table 2** online). Thus, formate production would be the second example of a mechanism to move protons through the membrane, by diffusion of an uncharged species, with release of the protons in the periplasm by oxidation and subsequent recapture of the associated electrons for sulfate reduction.

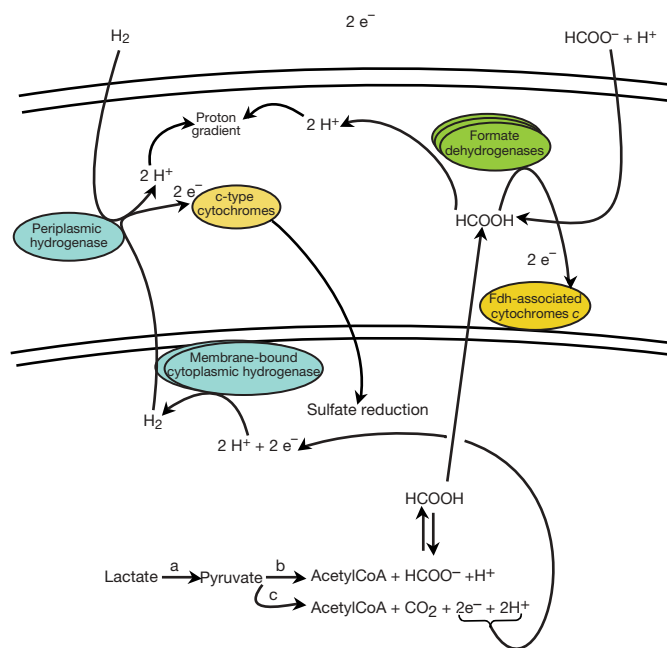


Figure 2 Hydrogen and formate cycling in *D. vulgaris*. Diagrammatic representation of the contribution of hydrogen and formate cycling proposed to occur during oxidation of organic acids by *D. vulgaris* Hildenborough. For hydrogen cycling, reducing equivalents ($2\text{H}^+ + 2\text{e}^-$) generated from lactate or pyruvate oxidation are suggested to be a substrate for one of two membrane-bound hydrogenases that have a cytoplasmic orientation. The gaseous hydrogen diffuses to the periplasm where any of several candidate hydrogenases (Fig. 1) would oxidize the hydrogen with the electrons released being captured by the *c*-type cytochrome network. The electrons could then be channeled through the cytoplasmic membrane by one of several putative transmembrane protein conduits (Fig. 1). For formate cycling, a similar path is proposed. Formate generated from pyruvate would, when protonated, diffuse into the periplasm and be reoxidized by one of three isozymes of formate dehydrogenase. Formate supplied externally would also be oxidized in the periplasm. Electrons would be transferred directly or indirectly to the *c*-type cytochrome matrix for subsequent delivery to the cytoplasmic membrane where sulfate is then reduced. In all cases, protons generated would contribute to the proton gradient supporting transport processes and ATP synthesis. For locus identification of genes, see **Supplementary Table 2** online.

Although growth on hexoses has never been demonstrated in *D. vulgaris*, genes for enzymes catalyzing each of the steps in glycolysis are readily recognized in the genome sequence²⁹, suggesting enzymatic activities similar to those demonstrated for the close relative *Desulfovibrio gigas*³⁰. The metabolic capacity to store glycogen would also appear to be present in *D. vulgaris*, and this polymer has been proposed to be a source of reductant for oxygen consumption³¹. Clearly, during growth on organic acids, a full complement of enzymes for gluconeogenesis is needed for cell wall biosynthesis and modification.

D. vulgaris does not mineralize organic carbon substrates but produces acetate as an end-product. We did not identify two genes for tri-carboxylic acid (TCA) cycle enzymes in the genome. First, we did not identify a typical α -ketoglutarate dehydrogenase, but sequences with strong similarity to a ferredoxin-dependent α -ketoglutarate synthase³² were evident (DVU1569–70, DVU1944–47). Second, we did not find a gene homologous to citrate synthase. This is unexpected because earlier work demonstrated citrate synthase activity in extracts of *D. vulgaris*³³. This activity produced citrate with a stereochemistry

opposite to that of the *E. coli* enzyme, perhaps accounting for the lack of the identification of a citrate synthase in the *D. vulgaris* genome. Whether the TCA cycle functions oxidatively or only reductively remains to be determined; however, it is likely to have a biosynthetic function because acetate is not oxidized.

D. vulgaris is known to use lactate, pyruvate, ethanol, malate and fumarate³⁴ but not riboses or hexoses. Consistent with this, it has one of the largest sets of carboxylate transporters described in bacteria, including six paralogs of the *E. coli* LctP lactate transporter and three sets of tripartite ATP-independent periplasmic transporters. *D. vulgaris* appears to have some capacity for carbohydrate uptake, including a glycerol channel and a probable ribose ABC transporter. It also has an unusual set of phosphotransferase system (PTS) genes organized in three clusters: DVU0829–31, encoding the general energy coupling proteins HPr and Enzyme I, and a distant homolog of a sorbose Enzyme IID; DVU1630–34, encoding two Enzyme IIAs, Enzyme IIB, a distant homolog of Enzyme IIC and a putative kinase; and DVU0981 encoding a multidomain protein with both a HPr and an Enzyme I domain. Thus, it appears to contain sufficient PTS components to form at least one complete sugar PTS transporter, although the substrate specificity of such a system is unclear. *D. vulgaris* possesses a variety of systems for uptake of nitrogenous compounds including urea, ammonia, amino acids and oligopeptides.

Use of alternate electron acceptors. Although *D. vulgaris* is best known for using sulfate (or other sulfur oxyanions) as the terminal electron acceptor, the genome contains pathways for the reduction of other terminal electron acceptors (*i.e.*, oxygen, nitrite and metal ions). Perhaps *D. vulgaris* uses these to derive energy for growth, to prevent competition by bacteria using these other electron acceptors or to protect the cell from inhibition of the sulfate reduction pathway by these alternative electron acceptors. For example, each of the two subunits of a periplasmic nitrite reductase (DVU0624–25, NrfHA) have multiple *c*-type heme binding sites (PS00190 domain: NfrH, four sites; NfrA, five sites), indicating that electrons for nitrite reduction may be drawn from the periplasmic cytochrome *c*₃ network. However, the lack of experimental evidence that *D. vulgaris* can grow by using nitrite as a terminal electron acceptor suggests that NrfHA likely functions only to prevent inhibition of sulfate reduction by nitrite³⁵.

The possible use of oxygen as a terminal electron acceptor by *Desulfovibrio* species has been a topic of intense debate³⁶. Genome analysis has revealed that *D. vulgaris* does possess a high-affinity terminal oxidase (DVU3270–71) homologous to the *D. gigas* enzyme that allows oxygen respiration under microaerophilic conditions³⁷. Additionally, *D. vulgaris* contains genes for a low affinity *aa3*-type terminal cytochrome *c* oxidase (DVU1812–15, *cox*) adjacent to the genes for monoheme cytochrome *c*-553 (DVU1817, *cyf*). Although the presence of these genes suggests that *D. vulgaris* should have the potential to respire oxygen both under microaerophilic and fully aerobic conditions³⁶, sustainable growth of *D. vulgaris* in oxygen has never been demonstrated. These oxidases may have evolved for protection from the damaging effects of oxygen³⁸. Protection against oxygen inactivation may also be provided by a cytoplasmic, nonenergy-conserving respiratory chain that terminates with rubredoxin-oxygen oxidoreductase (DVU3185, *roo* (ref. 39)). Interestingly, the gene for this terminal oxidase is preceded by those for rubredoxin (DVU3184, *rub*) and superoxide reductase (DVU3183, *sor* (ref. 40)), which has an oxygen-stable ferrous iron in its active site. The gene organization suggests that Sor helps to remove oxidative stress by reducing superoxide anions formed in the Roo-catalyzed reaction. The hydrogen peroxide is then further reduced to water by rubrerythrin-1 (DVU3094, *rbr1*)^{41,42}.

The genome sequence indicates the presence of two additional Rbr-homologs, rubrerythrin-2 (DVU2318, *rbr2*) and nigerythrin (DVU0019, *ngr*), that could also contribute to the reduction of hydrogen peroxide. In addition to Sor and Rbr, oxygen defense proteins that are typical for the anaerobic world, *D. vulgaris* has a periplasmic iron superoxide dismutase (DVU2410, Fe-Sod) and a plasmid-encoded catalase (DVUA0091).

DISCUSSION

The genome sequence of *D. vulgaris* allowed the identification and initial characterization of the bacterium's complex, periplasmic cytochrome network, and has provided a more complete picture of its transmembrane electron transport and cytoplasmic sulfate reduction capabilities. The number of hydrogenase isozymes recognized has expanded and, for the first time, candidates for cytoplasmic isozymes have been revealed. This is the first evidence for cytoplasmic isozymes, which are a critical component of the 'hydrogen-cycling' model of energy conservation. In addition to hydrogen and hydrogenases, the unexpectedly high number of formate dehydrogenases, suggests a second system of chemiosmotic energy conservation by the diffusion of an uncharged metabolic intermediate, formate, from the cytosol with subsequent periplasmic oxidation. The contribution of these enzymes to the cellular energy budget can now be explored through mutagenesis, microarray or other gene expression experiments, combined with physiological characterization²⁶. The recognition of the importance of this mode of energy conservation in microbial metabolism lags far behind that of fermentation and traditional respiration but may be found almost universally. The sequence also allows in-depth characterization of mechanisms for SRB-mediated biocorrosion and metal ion bioremediation and provides the intellectual foundation necessary to either effectively contain this microorganism (and others like it), through the use of oxygen, nitrite or biocidal agents, or to realize its utility for environmental remediation purposes.

METHODS

Strain. The strain chosen for whole genome sequencing was *D. vulgaris* subsp. *vulgaris* Hildenborough (NCIMB 8303), which was originally isolated in 1946 from clay soil near Hildenborough, Kent (UK)⁸. Because its *nif* genes-containing plasmid⁴³ is easily lost, a freeze-dried culture obtained from the NCIMB was inoculated directly into 100 ml of Postgate's Medium C, a lactate- and sulfate-containing medium⁸. The grown culture was immediately used for DNA extraction using the method of Marmur, modified to include digestion with proteinase K⁴⁴. Before random libraries were made, the DNA extracted from this culture was used to confirm the presence of *nif* genes by Southern hybridization. The genomic sequence is thus as closely related to that of the originally deposited strain as possible.

Sequencing. Cloning, sequencing and assembly were as described for genomes sequenced by TIGR⁴⁵. Plasmid libraries with one small-insert (2–3 kbp) and one large-insert (10–12 kbp) were constructed in pUC-derived vectors after random mechanical shearing (nebulization) of genomic DNA. Sequencing was achieved at a success rate of 86% with an average read length of 599 bp. The plasmid sequences were jointly assembled using TIGR Assembler. The coverage criteria were that every position required at least double-clone coverage (or sequence from a PCR product amplified from genomic DNA), and have either sequence from both strands or two different sequencing chemistries. The sequence was edited manually, and additional PCR⁴⁶ and sequencing reactions were done to close gaps, improve coverage and resolve sequence ambiguities. All repeated DNA regions were verified by PCR amplification across the repeat and sequencing of the product. The final genome is based on 54,926 sequences.

Genome analysis. The replicative origin was determined by colocalization of genes (*dnaA*, *dnaN*, *recF* and *gyrA*) often found near the origin in prokaryotic genomes and GC nucleotide skew (G–C/G+C) analysis⁴⁷. On this basis, we

designated base pair one in an intergenic region that is located in the putative origin of replication.

An initial set of open reading frames likely to encode proteins (CDSs) were predicted as previously described⁴⁵. All predicted proteins larger than 30 amino acids were searched against a nonredundant protein database as previously described⁴⁵. Frameshifts and point mutations were detected and corrected where appropriate. Remaining frameshifts and point mutations are considered to be authentic and were annotated as 'authentic frameshift' or 'authentic point mutation.' Protein membrane-spanning domains were identified by TopPred⁴⁸. The 5' regions of each CDS were inspected to define initiation codons using homologies, and position of ribosomal binding sites and transcriptional terminators. Two sets of hidden Markov models were used to determine CDS membership in families and superfamilies: pfam v11.0 (ref. 49) and TIGRFAMs 3.0 (ref. 50). Pfam v11.0 hidden Markov models were also used with a constraint of a minimum of two hits to find repeated domains within proteins and mask them.

Domain-based paralogous families were then built by doing all-versus-all searches on the remaining protein sequences using a modified version of a previously described method⁴⁵.

Requests for material should be addressed to gdv@tigr.org. The annotated genome sequence and the gene family alignments are available on the World Wide Web at http://www.tigr.org/tigr-scripts/CMR2/GenomePage3.pl?data_base=gdv. The sequences have been deposited in GenBank with accession no. AE017285 (chromosome) and AE017286 (plasmid).

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

This work was supported by the United States Department of Energy office of biological and environmental research through the microbial genome programs. We thank J.K. Voordouw for DNA isolation, S. Salzberg, O. White, M. Heaney, S. Lo, M. Holmes, M. Covarrubias, J. Sitz, A. Resnick, Y. Zhao, M. Zhurkin, R. Deal, R. Karamchedu and V. Sapiro for informatics, database and software support, and the TIGR faculty and sequencing core for expert advice and assistance.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 17 November 2003; accepted 3 March 2004

Published online at <http://www.nature.com/naturebiotechnology/>

- Truper, H.G. in *Sulfur: Its Significance for Chemistry, for the Geo-, Bio- and Cosmospere, and Technology*, edn. 5. (eds. Müller, A. & Krebs, B.) 351–365 (Elsevier, Amsterdam, 1984).
- Widdel, F. in *Biology of Anaerobic Organisms* (ed. Zehnder, A.J.B.) 469–586 (John Wiley, New York, 1988).
- Hamilton, W.A. Microbially influenced corrosion as a model system for the study of metal microbe interactions: a unifying electron transfer hypothesis. *Biofouling* **19**, 65–76 (2003).
- Singleton, R. Jr. in *The Sulfate-Reducing Bacteria: Contemporary Perspectives*. (eds. Odom, J.M. & Singleton, R. Jr.) 189–210 (Springer-Verlag, New York, 1993).
- Michel, C., Brugna, M., Aubert, C., Bernadac, A. & Bruschi, M. Enzymatic reduction of chromate: comparative studies using sulfate-reducing bacteria. Key role of polyheme cytochromes *c* and hydrogenases. *Appl. Microbiol. Biotechnol.* **55**, 95–100 (2001).
- Payne, R.B., Gentry, D.M., Rapp-Giles, B.J., Casalot, L. & Wall, J.D. Uranium reduction by *Desulfovibrio desulfuricans* strain G20 and a cytochrome *c*₃ mutant. *Appl. Environ. Microbiol.* **68**, 3129–3132 (2002).
- Hansen, T.A. Metabolism of sulfate-reducing prokaryotes. *Antonie Van Leeuwenhoek* **66**, 165–185 (1994).
- Postgate, J.R. *The Sulphate-Reducing Bacteria*, edn. 2, vol. 130 (Cambridge University Press, London, 1984).
- Price, S.B., Leung, K.Y., Barve, S.S. & Straley, S.C. Molecular analysis of IcrGVH, the V antigen operon of *Yersinia pestis*. *J. Bacteriol.* **171**, 5646–5653 (1989).
- Petnicki-Ocwieja, T. et al. Genomewide identification of proteins secreted by the Hrp type III protein secretion system of *Pseudomonas syringae* pv. tomato DC3000. *Proc. Natl. Acad. Sci. USA* **99**, 7652–7657 (2002).
- Xiong, J., Kurtz, D.M. Jr., Ai, J. & Sanders-Loehr, J. A hemerythrin-like domain in a bacterial chemotaxis protein. *Biochemistry* **39**, 5117–5125 (2000).
- Lissolo, T., Choi, E.S., LeGall, J. & Peck, H.D. Jr. The presence of multiple intrinsic membrane nickel-containing hydrogenases in *Desulfovibrio vulgaris* (Hildenborough). *Biochem. Biophys. Res. Commun.* **139**, 701–708 (1986).
- Voordouw, G. Carbon monoxide cycling by *Desulfovibrio vulgaris* Hildenborough. *J. Bacteriol.* **184**, 5903–5911 (2002).

14. Aubert, C. *et al.* Characterization of the cytochromes C from *Desulfovibrio desulfuricans* G201. *Biochem. Biophys. Res. Commun.* **242**, 213–218 (1998).
15. Aubert, C., Brugna, M., Dolla, A., Bruschi, M. & Giudici-Orticoni, M.T. A sequential electron transfer from hydrogenases to cytochromes in sulfate-reducing bacteria. *Biochim. Biophys. Acta.* **1476**, 85–92 (2000).
16. Pereira, I.A.C., Romao, C.V., Xavier, A.V., LeGall, J. & Teixeira, M. Electron transfer between hydrogenases and mono- and multi-heme cytochromes in *Desulfovibrio* spp. *J. Biol. Inorg. Chem.* **3**, 494–498 (1998).
17. Rossi, M. *et al.* The hmc operon of *Desulfovibrio vulgaris* subsp. *vulgaris* Hildenborough encodes a potential transmembrane redox protein complex. *J. Bacteriol.* **175**, 4699–4711 (1993).
18. Dolla, A., Pohorelic, B.K., Voordouw, J.K. & Voordouw, G. Deletion of the hmc operon of *Desulfovibrio vulgaris* subsp. *vulgaris* Hildenborough hampers hydrogen metabolism and low-redox-potential niche establishment. *Arch. Microbiol.* **174**, 143–151 (2000).
19. Badziong, W. & Thauer, R.K. Growth yields and growth rates of *Desulfovibrio vulgaris* (Marburg) growing on hydrogen plus sulfate and hydrogen plus thiosulfate as the sole energy sources. *Arch. Microbiol.* **117**, 209–214 (1978).
20. Pieulle, L., Magro, V. & Hatchikian, E.C. Isolation and analysis of the gene encoding the pyruvate-ferredoxin oxidoreductase of *Desulfovibrio africanus*, production of the recombinant enzyme in *Escherichia coli*, and effect of carboxy-terminal deletions on its stability. *J. Bacteriol.* **179**, 5684–5692 (1997).
21. Pankhania, I.P., Spormann, A.M., Hamilton, W.A. & Thauer, R.K. Lactate conversion to acetate, CO₂, and H₂ in cell suspensions of *Desulfovibrio vulgaris* (Marburg): indications for the involvement of an energy driven step. *Arch. Microbiol.* **150**, 26–31 (1988).
22. Odom, J.M. & Peck, H.D. Jr. Hydrogen cycling as a general mechanism for energy coupling in the sulfate-reducing bacteria *Desulfovibrio* sp. *FEMS Microbiol. Lett.* **12**, 47–50 (1981).
23. Voordouw, G. & Brenner, S. Nucleotide sequence of the gene encoding the hydrogenase from *Desulfovibrio vulgaris* (Hildenborough). *Eur. J. Biochem.* **148**, 515–520 (1985).
24. Meuer, J., Bartoschek, S., Koch, J., Kunkel, A. & Hedderich, R. Purification and catalytic properties of Ech hydrogenase from *Methanosarcina barkeri*. *Eur. J. Biochem.* **265**, 325–335 (1999).
25. Kerby, R.L. *et al.* Genetic and physiological characterization of the *Rhodospirillum rubrum* carbon monoxide dehydrogenase system. *J. Bacteriol.* **174**, 5284–5294 (1992).
26. Haveman, S.A. *et al.* Gene expression analysis of energy metabolism mutants of *Desulfovibrio vulgaris* Hildenborough indicates an important role for alcohol dehydrogenase. *J. Bacteriol.* **185**, 4345–4353 (2003).
27. Böck, A. & Sawers, G. in *Escherichia Coli and Salmonella: Cellular and Molecular Biology*, edn. 2 (ed. Neidhardt, F.C.) 262–282 (ASM Press, Washington, DC, 1996).
28. Gennis, R.B. & Stewart, V. in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, edn. 2 (ed. Neidhardt, F.C.) 217–261 (ASM Press, Washington, DC, 1996).
29. Wall, J.D. *et al.* in *Biochemistry and Physiology of Anaerobic Bacteria* (eds. Ljungdahl, L.G., Adams, M.W., Ferry, J.G. & Barton, L.L.) 85–98 (Springer-Verlag, New York, 2003).
30. Fareleira, P., Legall, J., Xavier, A.V. & Santos, H. Pathways for utilization of carbon reserves in *Desulfovibrio gigas* under fermentative and respiratory conditions. *J. Bacteriol.* **179**, 3972–3980 (1997).
31. Santos, H. *et al.* Aerobic metabolism of carbon reserves by the “obligate anaerobe” *Desulfovibrio gigas*. *Biochem. Biophys. Res. Commun.* **15**, 551–557 (1993).
32. Buchanan, B.B. & Arnon, D.I. A reverse KREBS cycle in photosynthesis: consensus at last. *Photosynth. Res.* **24**, 47–53 (1990).
33. Gottschalk, G. & Barker, H.A. Presence and stereospecificity of citrate synthase in anaerobic bacteria. *Biochemistry* **6**, 1027–1034 (1967).
34. Widdel, F. & Bak, F. in *The Prokaryotes* (ed. Balows, *et al.*) 3352–3378 (Springer-Verlag, New York, 1992).
35. Greene, E.A., Hubert, C., Nemati, M., Jenneman, G.E. & Voordouw, G. Nitrite reductase activity of sulphate-reducing bacteria prevents their inhibition by nitrate-reducing, sulphide-oxidizing bacteria. *Environ. Microbiol.* **5**, 607–617 (2003).
36. Cypionka, H. Oxygen respiration by desulfovibrio species. *Annu. Rev. Microbiol.* **54**, 827–848 (2000).
37. Lemos, R.S. *et al.* The ‘strict’ anaerobe *Desulfovibrio gigas* contains a membrane-bound oxygen-reducing respiratory chain. *FEBS Lett.* **496**, 40–43 (2001).
38. Richardson, D.J. Bacterial respiration: a flexible process for a changing environment. *Microbiology* **146**, 551–571 (2000).
39. Dos Santos, W.G. *et al.* Purification and characterization of an iron superoxide dismutase and a catalase from the sulfate-reducing bacterium *Desulfovibrio gigas*. *J. Bacteriol.* **182**, 796–804 (2000).
40. Jenney, F.E. Jr., Verhagen, M.F., Cui, X. & Adams, M.W. Anaerobic microbes: oxygen detoxification without superoxide dismutase. *Science* **286**, 306–309 (1999).
41. Lumpio, H.L., Shenvi, N.V., Summers, A.O., Voordouw, G. & Kurtz, D.M. Jr. Rubrythrin and rubredoxin oxidoreductase in *Desulfovibrio vulgaris*: a novel oxidative stress protection system. *J. Bacteriol.* **183**, 101–108 (2001).
42. Fournier, M. *et al.* Function of oxygen resistance proteins in the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* hildenborough. *J. Bacteriol.* **185**, 71–79 (2003).
43. Postgate, J., Kent, H.M. & Robson, R.L. DNA from diazotrophic *Desulfovibrio* strains is homologous to *Klebsiella pneumoniae* structural nif DNA and can be chromosomal or plasmid-borne. *FEMS Microbiol. Lett.* **33**, 159–163 (1986).
44. Voordouw, G., Niviere, V., Ferris, F.G., Fedorak, P.M. & Westlake, D.W. Distribution of hydrogenase genes in *Desulfovibrio* spp. and their use in identification of species from the oil field environment. *Appl. Environ. Microbiol.* **56**, 3748–3754 (1990).
45. Heidelberg, J.F. *et al.* Genome sequence of the dissimilatory metal ion-reducing bacterium *Shewanella oneidensis*. *Nat. Biotechnol.* **20**, 1118–1123 (2002).
46. Tettelin, H., Radune, D., Kasif, S., Khouri, H. & Salzberg, S.L. Optimized multiplex PCR: efficiently closing a whole-genome shotgun sequencing project. *Genomics* **62**, 500–507 (1999).
47. Lobry, J.R. Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* **13**, 660–665 (1996).
48. Claros, M.G. & von Heijne, G. TopPred II: an improved software for membrane protein structure predictions. *Comput. Appl. Biosci.* **10**, 685–686 (1994).
49. Bateman, A. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **28**, 263–266 (2000).
50. Haft, D.H. *et al.* TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res.* **29**, 41–43 (2001).