LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-JRNL-220798

Whole-Genome Transcriptional Analysis of Chemolithoautotrophic Thiosulfate Oxidation by Thiobacillus denitrificans Under Aerobic vs. Denitrifying Conditions

H. R. Beller, T. E. Letain, A. Chakicherla, S. R. Kane, T. C. Legler, M. A. Coleman

April 24, 2006

Journal of Bacteriology

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Revised manuscript JB00568-06

Whole-genome transcriptional analysis of chemolithoautotrophic thiosulfate oxidation by *Thiobacillus denitrificans* under aerobic vs. denitrifying conditions

Harry R. Beller,* Tracy E. Letain, Anu Chakicherla, Staci R. Kane, Tina C. Legler, and Matthew A. Coleman

Lawrence Livermore National Laboratory, Livermore, CA 94551

Running title: *T. denitrificans* microarrays: aerobic vs. denitrifying

***Corresponding author:**

Lawrence Livermore National Laboratory
P.O. Box 808, L-542
Livermore, CA 94551-0808
E-mail beller2@llnl.gov
Phone (925) 422-0081
Fax (925) 423-7998

Abstract

Thiobacillus denitrificans is one of the few known obligate chemolithoautotrophic bacteria capable of energetically coupling thiosulfate oxidation to denitrification as well as aerobic respiration. As very little is known about the differential expression of genes associated with key chemolithoautotrophic functions (such as sulfur-compound oxidation and CO_2 fixation) under aerobic versus denitrifying conditions, we conducted whole-genome, cDNA microarray studies to explore this topic systematically. The microarrays identified 277 genes (approximately ten percent of the genome) as differentially expressed using Robust Multi-array Average statistical analysis and a 2-fold cutoff. Genes upregulated (ca. 6- to 150-fold) under aerobic conditions included a cluster of genes associated with iron acquisition (e.g., siderophore-related genes), a cluster of cytochrome *cbb*$_3$ oxidase genes, *cbbL* and *cbbS* (encoding the large and small subunits of form I ribulose 1,5-bisphosphate carboxylase/oxygenase, or RubisCO), and multiple molecular chaperone genes. Genes upregulated (ca. 4- to 95-fold) under denitrifying conditions included *nar*, *nir*, and *nor* genes (associated respectively with nitrate reductase, nitrite reductase, and nitric oxide reductase, which catalyze successive steps of denitrification), *cbbM* (encoding form II RubisCO), and genes involved with sulfur-compound oxidation (including two physically separated but highly similar copies of sulfide:quinone oxidoreductase and of *dsrC*, associated with dissimilatory sulfite reductase). Among genes associated with denitrification, relative expression levels (i.e., degree of upregulation with nitrate) tended to decrease in the order *nar* > *nir* > *nor* > *nos*. Reverse transcription, quantitative PCR analysis was used to validate these trends.

1 **INTRODUCTION**

2 *Thiobacillus denitrificans* is an obligately chemolithoautotrophic bacterium characterized
3 by its ability to conserve energy from the oxidation of inorganic sulfur compounds under either
4 aerobic or denitrifying conditions (5). As a facultative anaerobe, *T. denitrificans* may benefit
5 from modulating key components of its energy metabolism, such as sulfur-compound oxidation
6 or carbon dioxide fixation, according to whether oxygen or nitrate is the terminal electron
7 acceptor. For example, *T. denitrificans* can express both form I and form II ribulose 1,5-
8 bisphosphate carboxylase/oxygenase (RubisCO), which have different relative affinities for CO_2
9 and the competing substrate O_2 and therefore may differ in CO_2 fixation efficiency under aerobic
10 vs. denitrifying conditions. Also, among its large complement of genes associated with sulfur-
11 compound oxidation, *T. denitrificans* shares some genes with aerobic, chemolithotrophic sulfur-
12 oxidizing bacteria and some with anaerobic, phototrophic sulfur bacteria (5). There is very little
13 information on how (or whether) *T. denitrificans* modulates the expression of these sulfur-
14 oxidizing genes as a function of the prevailing terminal electron acceptor. The recent availability
15 of the complete genome sequence of *T. denitrificans* (5) and of high-density oligonucleotide
16 microarrays provided us an opportunity to address these and other questions by systematically
17 investigating differential expression across the entire *T. denitrificans* genome under aerobic vs.
18 denitrifying conditions.

19 **MATERIALS AND METHODS**

20 **Cell growth and exposure conditions.** To represent gene expression under denitrifying
21 conditions, *T. denitrificans* (ATCC strain 25259, obtained from the American Type Culture
22 Collection) was cultivated at 30°C under strictly anaerobic conditions as described previously (4)
23 with growth medium that contained 20 mM thiosulfate, 20 mM nitrate, and 30 mM bicarbonate

1 (pH ~7). For exposure immediately before harvesting of RNA, 1200 mL of cells in late
2 exponential phase (1 to 2 x10^8 cells/mL) were harvested anaerobically by centrifugation (13,400
3 x *g*, 15°C, 10 min), resuspended in modified growth medium (phosphate concentration reduced
4 to 1.5 mM), and three 10-mL replicates (ca. 7.3 mg protein each) in sealed vials (90% N_2 – 10%
5 CO_2 headspace) were incubated for ca. 35 min. Cell growth, resuspension, and incubation were
6 performed in an anaerobic glove box (4).
7 To represent gene expression under aerobic conditions, *T. denitrificans* was cultivated
8 (two successive transfers) with growth medium that differed from the denitrifying medium in
9 several noteworthy respects: it contained no nitrate, it was equilibrated with atmospheric oxygen
10 (rotating in a shake flask at 200 rpm), it contained 70 mM phosphate, 0.7 μM copper (as
11 compared to 1.2 μM in denitrifying medium), and 10 μM iron (as compared to 7.5 μM). The
12 reason for using a higher phosphate buffer concentration in the aerobic medium was that, when
13 lower phosphate concentrations were tested, the pH of aerobic growth medium dropped from ~7
14 to ~5 as *T. denitrificans* oxidized thiosulfate. This follows from the stoichiometry of thiosulfate
15 oxidation, which yields five-fold more protons per mole of thiosulfate under aerobic than
16 denitrifying conditions. For exposure immediately before harvesting of RNA, 1200 mL of cells
17 in late exponential phase were harvested by centrifugation, resuspended in aerobic growth
18 medium, and three 10-mL replicates (ca. 3.9 mg protein each) were incubated at 30°C in 125-mL
19 Erlenmeyer flasks rotating at 200 rpm for 60 min. The pH of the cell suspensions remained in
20 the circumneutral range throughout the incubation period.
21 Metabolic activity (thiosulfate oxidation to sulfate; nitrate consumption in anaerobic
22 cultures) was assessed in all anaerobic and aerobic suspensions by sampling each culture twice:
23 immediately upon resuspension and immediately before harvesting for RNA. Ion

chromatography was used to determine thiosulfate, sulfate, and nitrate concentrations (4). Previous experiments indicated that metabolic rates during suspensions were sufficiently linear throughout the incubation period that initial and final concentrations could be used to calculate representative specific rates. These analyses demonstrated that specific thiosulfate oxidation rates were comparable under denitrifying and aerobic conditions (0.43$\pm$0.005 and 0.56$\pm$0.006 μmol thiosulfate$\cdot$min$^{-1}\cdot$mg protein^{-1}, respectively).

RNA extraction. Immediately after exposures, two volumes of RNAprotect (Qiagen) were added to each culture. Samples were incubated at room temperature for 12 min, split in half, and centrifuged at 4,000 rpm for 10 min. The supernatant was decanted and the pellet was stored at –20°C until extraction. RNA extraction was carried out with a MasterPure Complete DNA and RNA Purification Kit (EpiCentre) using a modified protocol. Briefly, 300 μL of lysis solution containing 112 μg proteinase K was added to the cell pellet and the sample was incubated at 65°C for 20-25 min. The sample was placed on ice for 3-5 min and 200 μL of MPC solution was added to precipitate protein. The supernatant was recovered after centrifugation at >10,000 x *g*, at 4°C for 10 min. Nucleic acid was subsequently precipitated from the supernatant after addition of 500 μL 99% isopropanol and centrifugation at >10,000 x *g*, at 4°C for 10 min. The pellet was treated with DNase I for 20 min at 37°C. To this sample was added 200 μL each of 2X T&C lysis solution and MPC solution with vortexing after each addition. The samples were placed on ice for 3-5 min and centrifuged at >10,000 x *g*, at 4°C for 10 min. RNA in the supernatant was recovered by isopropanol precipitation as described. The RNA pellet was washed twice with 75% ethanol, dried briefly, suspended in water, and stored at -80°C until cDNA synthesis. Aliquots were analyzed with a Bioanalyzer (Agilent), which indicated minimal

1 degradation and concentrations ranging from 310 to 2000 ng/µL. 260/280 ratios ranged from 1.7
2 to 2.1.
3 **Preparation of labeled cDNA.** cDNA production and labeling were performed by
4 NimbleGen Systems, Inc. RNA samples were thawed on ice and 10 µg total RNA was used to
5 perform cDNA synthesis with SuperScript II reverse transcriptase and random hexamers. After
6 this reaction, RNase A and H were used to digest the RNA. Single-stranded cDNA was
7 subsequently purified by phenol extraction. Glycogen (10 µg) was added as a carrier prior to
8 precipitation with 1/10 volume ammonium acetate and 2.5 volumes of absolute ethanol. The
9 resulting pellet was suspended in 30 µL water. The cDNA yield was determined by UV/visible
10 spectrophotometry at 260 nm. The cDNA was partially digested with DNase I (0.2 U) at 37°C
11 for 13 min or until 50- to 200-base fragments were observed with the Bioanalyzer. The
12 fragmented cDNA was end-labeled using biotin-N_6-ddATP and terminal deoxynucleotidyl
13 transferase (51 U) with incubation at 37°C for 2 hours. The labeled product was concentrated to
14 20 µL using a Microcon YM-10 10,000 MWCO filter device (Millipore) and frozen at -20°C
15 prior to hybridization.
16 **Array design.** The genome sequence from *T. denitrificans* ATCC 25259 (5)(GenBank
17 accession no. CP000116) was submitted to NimbleGen Systems Inc. for microarray design and
18 manufacture using maskless, digital micromirror technology. High-density (approximately
19 400,000-spot) microarrays employed a randomized design and a 4-in-9 pattern to enhance
20 sensitivity. Three replicates of the genome were included per chip. An average of ten different
21 60-base oligonucleotides (60-mer probes) represented each ORF (open reading frame) in the
22 genome. 60-mer probes were selected such that each probe had at least three mismatches
23 compared to all other 60-mers in the target genome. A total of 28,320 probes were designed for

1 the genome, which was annotated to have 2,832 ORFs at the time of microarray design (the
2 finished genome is annotated to have 2,827 ORFs; 5). A quality control check (hybridization)
3 was performed for each array, which contained on-chip control oligonucleotides.
4 **Microarray hybridization and analysis.** NimbleGen Systems, Inc. performed array
5 hybridization using their Hybriwheel technology. The arrays were pre-hybridized at 45°C in a 50
6 mM MES (4-Morpholineethanesulfonic acid) buffer containing 500 mM NaCl, 10 mM EDTA,
7 and 0.005% Tween-20 with herring sperm DNA (0.1 mg/ml) to prevent non-specific binding to
8 the array. After 15 min, 4 µg of labeled cDNA in hybridization buffer was added and arrays
9 were incubated at 45°C for 16-20 h. Several wash steps (initially non-stringent and later
10 stringent conditions) removed free probe, followed by detection of bound probe with Cy3-
11 labeled streptavidin. To amplify the signal, biotinylated anti-streptavidin goat antibody was
12 hybridized to the array. The arrays were analyzed using an Axon GenePix 4000B Scanner with
13 associated software (Molecular Devices Corp., Sunnyvale, CA).
14 **Microarray data analysis.** Investigation of reproducible differences between treatments
15 was performed using the Bioconductor R software package. Data were processed using quantile
16 normalization (7) and background correction was performed using the RMA (Robust Multi-array
17 Average) method. Data were visualized with box-and-whisker plots and scatterplots (Volcano
18 plots). Intensities were adjusted to have the same interquartile range. A linear model fit was
19 determined for each gene using the LIMMA package (Linear Models for Microarray Data;
20 Gordon K. Smyth) and lists of genes with the most evidence of differential expression were
21 obtained.
22 **Reverse transcription, quantitative PCR analysis.** Confirmation of transcript levels
23 for modulated genes was performed by reverse transcription, quantitative PCR (RT-qPCR)

analysis of RNA samples representing each of the two experimental conditions. Total RNA from samples used for microarray analysis was reverse transcribed and amplified using a QuantiTect SYBR Green RT-PCR kit (Qiagen) with gene-specific primers. Each gene-specific PCR was performed in triplicate using 25-µL reactions containing ~20 ng of template on a Prism 7000 cycler (ABI). Calibration curves were performed with genomic DNA serially diluted over a range of four to five orders of magnitude. The PCR conditions were optimized to be performed as follows for all transcripts: 50°C for 30 min; 95°C for 15 min; 94°C for 15 s; 58°C for 30 s; 72°C for 30 s; 30-35 cycles. The primers are listed in Supplementary Table 1.

RT-PCR analysis of *sqr* and *dsrC* transcripts. Qualitative and quantitative RT-PCR studies were performed to investigate whether a gene associated with sulfur-compound oxidation, *dsrC* (Tbd1408), was co-transcribed with upstream genes associated with nitrate reduction (*nar* genes) and sulfide:quinone oxidoreductase (*sqr*; Tbd1407). Forward PCR primers were designed for Tbd1406 (*narI*) and Tbd1407 (Tbd1406F and Tbd1407F, respectively; Supplementary Table 1) and reverse primers were designed for Tbd1408 (Tbd1408R and Tbd1408R2; Supplementary Table 1). Control primers for the large transcript (targeting 1406 and 1407) and the Tbd1408 transcript were also designed and tested. The PCR conditions were optimized using *T. denitrificans* genomic DNA. cDNA was produced from RNA samples used in microarray experiments (aerobic and denitrifying conditions) with 150 - 250 ng RNA (pre-treated with DNase), 100 units Retroscript reverse transcriptase (Ambion), random decamers, and incubation at 43°C for 75 min. PCR products were visualized by gel electrophoresis on a 1% agarose 1X TAE gel with ethidium bromide staining and UV illumination. RT-qPCR analysis was performed with a forward primer from Tbd1407, the Tbd1407-1408 intergenic region, or Tbd1408 with a reverse primer for Tbd1408 (Supplementary Table 1). The primers

and template were added to SYBR Green Master Mix (Bio-Rad) and reactions were run on a Cepheid SmartCycler using the following program: 98°C for 15 s; 60°C for 60 s; 40 cycles. Controls for both RT-PCR and RT-qPCR analyses included trials without reverse transcriptase and trials without template.

Microarray data accession number. Microarray data have been deposited in the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo) under accession number XXXXXXX.

RESULTS AND DISCUSSION

Genome-wide observations. Microarray analysis identified 277 genes in *T. denitrificans* as differentially expressed under aerobic vs. denitrifying conditions using a 2-fold cutoff ($P<0.0001$). The top 50 upregulated genes under denitrifying conditions are listed in Table 1 and the top 50 upregulated genes under aerobic conditions are listed in Table 2. A large percentage of the most upregulated genes under either denitrifying or aerobic conditions occur as gene clusters and can be classified within a small number of functional categories. To illustrate, under denitrifying conditions, upregulated genes include those associated with nitrate reductase (Tbd1401-1406; *nar* cluster), nitrite reductase (Tbd0070-0077; *nir* cluster), nitric oxide reductase (Tbd0554-0562; *nor* cluster), and sulfur-compound oxidation (including Tbd1407-1408, adjacent to the *nar* cluster). Other gene clusters with less obvious functional associations are also included among the most upregulated genes (e.g., Tbd1499-1501; Tbd1835-1838, which includes divergently transcribed genes) and certain functions are represented by single genes (e.g., *cbbM*, which encodes form II RubisCO). Under aerobic conditions, upregulated genes include a large cluster putatively encoding proteins associated with iron acquisition (Tbd0705-0725, which account for more than 40% of the top 50 upregulated genes), a cytochrome cbb_3

1 oxidase (Tbd0638-0643), multiple chaperones (including Tbd1537-1539), and form I RubisCO
2 (*cbbS* and *cbbL*; Tbd2623 and 2624). Under aerobic conditions, as under denitrifying conditions,
3 gene clusters with less obvious functional associations are also included among the most
4 upregulated genes (e.g., Tbd2355-2358; Tbd2592-2594; Tbd2777-2778; Table 2) and certain
5 functions are represented by single genes (e.g.,Tbd1365, a putative *dsrC* presumably associated
6 with sulfur-compound oxidation).

7 The observation that a relatively small number of functional categories account for the
8 majority of the most differentially expressed genes is apparent in Fig. 1, which plots $\log_2$
9 probability of differential expression vs. $\log_2$ fold differential expression for all ORFs identified
10 in the genome. The color coding in Fig. 1 corresponds to the major categories listed in Tables 1
11 and 2, namely, denitrification, sulfur-compound oxidation, CO_2 fixation via RubisCO (forms I
12 and II), iron acquisition, cytochrome *cbb*$_3$ oxidase, and chaperones and stress proteins; all genes
13 not falling within these categories in Tables 1 and 2, and all genes not included in Tables 1 and
14 2, are gray in Fig. 1.

15 **Denitrification.** Although it is not surprising that genes associated with denitrification
16 (*nar*, *nir*, *nor* genes) were among the most upregulated genes under denitrifying conditions,
17 subtler trends in expression of these genes were more novel. Most notably, relative expression
18 levels (i.e., degree of upregulation under denitrifying conditions) tended to decrease in the order
19 *nar* > *nir* > *nor* > *nos* (Fig. 2). With the exception of a few genes (primarily associated with
20 transcriptional regulators, such as *narXL* and Tbd0078-0079), fold upregulation for
21 denitrification genes fell in the following ranges: *nar* – 54- to 95-fold, *nir*- 10- to 21-fold, *nor*- 4-
22 to 10-fold, and *nos*- 0.5- to 0.9-fold. This trend was both a function of generally decreasing
23 absolute expression levels under denitrifying conditions (except for the structural genes *nirS*,

1 *norCB*, and *nosZ*) and increased expression of *nos* genes (especially *nosZ*) under aerobic
2 conditions (Fig. 2).

3 To our knowledge, this is the most complete data set for differential aerobic/denitrifying
4 expression across the complement of denitrification genes; previous transcriptional studies have
5 focused primarily on structural genes or on gene clusters associated with only one of the four
6 denitrification enzymes. In a general sense, the microarray results for *T. denitrificans* are
7 consistent with the well-documented transcriptional activation of denitrification genes as a
8 function of low O_2 tension and the presence of a nitrogen oxide (NO_3^-, NO_2^-, NO, N_2O) (e.g.,
9 reviewed in ref. 36). With respect to *T. denitrificans* specifically, the microarray results are
10 generally consistent with greatly increased NAR and NIR enzyme activities (in crude extracts)
11 that were observed to accompany the transition from aerobic to denitrifying conditions in
12 continuous culture (18). Furthermore, the decreasing trend in upregulation shown in Fig. 2 could
13 be consistent with induction of each reductase component by its cognate substrate, as one might
14 expect the concentration pattern of $[NO_3^-] > [NO_2^-] > [NO] > [N_2O]$ in a denitrifying cell
15 (although this conception clearly oversimplifies the regulation of denitrification).

16 However, for *nos* genes in particular, the results for *T. denitrificans* appear to deviate
17 from findings for other denitrifying species for which data are available, namely, *Pseudomonas*
18 *stutzeri*, *Paracoccus denitrificans*, and *Paracoccus pantotrophus* (formerly *Thiosphaera*
19 *pantotropha*). For example, whereas expression of *nos* genes (including *nosD*) in *T.*
20 *denitrificans* was comparable under aerobic and denitrifying conditions, the amount of *nosD*
21 transcripts in *P. stutzeri* (revealed by Northern blot analysis) increased steadily and dramatically
22 during the first hour following a shift from aerobic to denitrifying conditions in continuous
23 culture (15). In another continuous culture study of *P. stutzeri* (20), NosZ levels were at least

1 10-fold greater for cells under denitrifying conditions than for cells under fully aerobic
2 conditions (in the presence of nitrate). A continuous culture study of *Paracoccus denitrificans*
3 revealed more than a 10-fold increase in the amount *nosZ* transcripts during the first hour
4 following transition from aerobic to denitrifying conditions (3); this temporal trend was
5 qualitatively similar to those of other denitrification genes, however, *narH* and *nirS* transcript
6 copy numbers increased more (approximately 30- to 45-fold) (3). In continuous culture and
7 batch culture studies of *P. pantotrophus* (23), NosZ expression was 2- to ca. 20-fold greater
8 under denitrifying conditions than under aerobic conditions (in the presence or absence of
9 nitrate) and clearly decreased as a function of increasing oxygen concentration in continuous
10 culture. Differences in experimental approach preclude a direct comparison of the results of the
11 present study with those just cited for *P. stutzeri* and *P. denitrificans*; such differences in
12 experimental approach include the use of continuous cultures vs. batch cultures and
13 measurement after aerobic/anaerobic transitions vs. comparisons of cultures grown exclusively
14 under aerobic or denitrifying conditions. Acknowledging this caveat, the available data suggest
15 inconsistent trends for differential aerobic/denitrifying *nos* gene expression in *T. denitrificans*
16 compared to other species studied. In *P. stutzeri*, *P. denitrificans*, and *P. pantotrophus*, there
17 appears to be considerable upregulation of *nos* genes (at least *nosZ* and *nosD*) under denitrifying
18 conditions; this is clearly not the case for *T. denitrificans* (indeed, there is slight upregulation of
19 these genes under aerobic conditions; Fig. 2).

20 In the absence of additional experimental evidence, we cannot explain the anomalous
21 lack of differential transcription of *nosZ* and other *nos* genes in *T. denitrificans* under aerobic vs.
22 denitrifying conditions. Nonetheless, examination of promoter regions for some key genes
23 associated with denitrification did reveal possible clues. Specifically, these promoter regions in

1 *T. denitrificans* were examined with respect to potential FNR boxes (i.e., DNA-binding motifs
2 for FNR-like transcription factors). When compared to the canonical FNR box 5'-TTGAT-N_4-
3 ATCAA-3' described for *E. coli* (36), slightly degenerate sequences were found upstream of
4 *narK* (aTGA**c**. . . **A**TCtt, located 778 nt from the translational start site of Tbd1401), *nirS*
5 (TTGA**c**. . . **A**TCAA, located 76 nt from the translational start site of Tbd0077), *norC* (aTGA**c** . .
6 . **A**TCAA, located 163 nt from the translational start site of Tbd0562), and *nosZ* (TTGAg . . .
7 gTCAg, 1310 nt from the translational start site of Tbd1389). Two features shared by the *narK*,
8 *nirS*, and *norC* versions and not in the *nosZ* version are the **c** and **A** shown in bold face type
9 above. Also, upstream of *nosZ*, there were five additional sequences with 5'-ends that matched
10 either the canonical FNR box or the *narK*, *nirS*, or *norC* FNR boxes cited above, but these were
11 more degenerate on the 3' end (with only 0 to 2 bases agreeing with the canonical sequence).
12 Degeneracy at the 3' end of FNR boxes upstream of *nosZ* has been observed for some
13 denitrifying species (8, 36) but not for others (2). Considering that FNR boxes for positive
14 regulation of denitrification genes are typically centered at a distance of -41.5 nt from the
15 transcription start site (36), most of the putative FNR boxes just described for *T. denitrificans*
16 seem to be very far upstream. It is not clear whether any of these characteristics of putative FNR
17 boxes in *T. denitrificans* could explain the lack of *nosZ* upregulation under denitrifying
18 conditions.
19 **Sulfur-compound oxidation.** Although a diverse complement of more than 50 genes
20 associated with sulfur-compound oxidation has been described in *T. denitrificans* ATCC 25259
21 (5), those genes associated with activity under aerobic vs. denitrifying conditions have not been
22 elucidated to date. Many of the *T. denitrificans* genes associated with sulfur-compound oxidation
23 (5) were not found to be differentially expressed in this study (Tables 1 and 2). Among the

genes not appearing in Tables 1 and 2 were clusters of sulfur-compound oxidation genes that were very highly expressed under both aerobic and denitrifying conditions. These include *soxXYZA* (Tbd0567-0564), *dsrABEFHCMKLJOP* (Tbd2485-2474), and the genes encoding ATP sulfurylase and APS reductase (Tbd0874-0872). The expression levels of these genes were typically at or above the 95th percentile expression level observed across the genome. Indeed, many of these genes are likely to be constitutively expressed in *T. denitrificans*, as, in most cases, their expression levels were similarly high under Fe(II)-oxidizing, denitrifying conditions when no sulfur-containing electron donor was present (Beller et al., unpublished microarray data).

Differential expression was observed for certain genes associated with sulfur-compound oxidation; in some cases, the absolute expression levels of these genes when upregulated were also in the range of the 95th percentile expression level observed across the genome. Among the most differentially regulated genes putatively associated with sulfur-compound oxidation, all but one were upregulated under denitrifying rather than aerobic conditions (Tables 1 and 2; Fig. 1). These included two copies of sulfide:quinone oxidoreductase (*sqr*) that share 43% amino acid identity (Tbd1407 and Tbd2225; 55- and 6.5-fold upregulated under denitrifying conditions), a rhodanese-like domain protein (Tbd1650; 8.7-fold upregulated), and two putative copies of *dsrC* that share 88% amino acid identity (Tbd1408 and 2327; 14- and 5.7-fold upregulated). Another putative copy of *dsrC* (Tbd1365) was upregulated 6.9-fold aerobically. Another rhodanese copy (Tbd2399) was less upregulated aerobically (3.8-fold), but was included in a gene cluster that exhibited some stronger aerobic upregulation (Tbd2398-Tbd2401; Table 2).

Inasmuch as three *dsrC* copies were among the most differentially regulated genes, it is noteworthy that the *T. denitrificans* genome includes eight putative *dsrC* copies overall (5); the

1 phylogenetic relationships and genomic organization of these homologs have been presented
2 elsewhere (5). Only one copy, Tbd2480, is located in the large gene cluster
3 *dsrABEFHCMKLJOPNR* (Tbd2485-2472) and is constitutively expressed at a high level (5).
4 Although the exact function of DsrC is not known, it is almost certainly involved with sulfur-
5 compound oxidation; the associated *dsrAB* genes encode a siroheme-containing sulfite reductase
6 that has been proposed to catalyze the oxidation of certain inorganic sulfur species (e.g.,
7 hydrogen sulfide or sulfane-sulfur derived from thiosulfate) to sulfite (27, 33).
8 In light of the strong upregulation of Tbd1407 (*sqr*) and Tdb1408 (putative *dsrC*) under
9 denitrifying conditions (Table 1), the genomic location of these genes is noteworthy: they are
10 immediately downstream of the *narKK$_2$GHJI* cluster (Tbd1401-1406)(5), which encodes a
11 membrane-bound, dissimilatory nitrate reductase (and associated nitrate/nitrite transporters) (Fig.
12 3A). As there is not even a single intergenic base separating Tbd1406 and Tbd1407, it follows
13 that Tbd1407 is part of a polycistronic transcript including *nar* genes (probably *narKK$_2$GHJI*).
14 However, the intergenic region between Tbd1407 and Tbd1408 includes a putative ribosomal
15 binding site and FNR box (Fig. 3B). Thus, co-regulation rather than co-transcription of Tbd1407
16 and Tbd1408 is plausible, and indeed is suggested by the anomalously high expression of
17 Tbd1408 relative to Tbd1407 under aerobic conditions (Fig. 3A). To further investigate whether
18 Tbd1408 was transcribed independently of Tbd1407 and upstream *nar* genes, RT-qPCR studies
19 were conducted. These studies confirmed that, for the most part, Tbd1408 was transcribed
20 separately from Tbd1407 and upstream *nar* genes: under denitrifying or aerobic conditions, the
21 copy number of transcripts of Tbd1408 (*dsrC*) was at least 10-fold greater than the copy number
22 of transcripts including Tbd1407 and1408 (*sqr* and *dsrC*) (Fig. 3C). Although the RT-qPCR
23 studies were constrained by amplicon length and did not address transcripts extending upstream

beyond Tbd1407, semi-quantitative RT-PCR analyses (Fig. 3D) suggested that, at least under denitrifying conditions, the Tbd1407-1408 transcripts actually extended at least from Tbd1406 (*narI*) to Tbd1408 (lane 9). Overall, the microarray and RT-qPCR results suggest that the promoter(s) controlling the expression of *dsrC* (Tbd1408), while clearly effecting stronger activation under denitrifying than aerobic conditions, may be further enhanced by the presence of sulfur compounds under aerobic or denitrifying conditions.

Carbon dioxide fixation. The genome of *T. denitrificans* encodes both form I and form II ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO) enzymes for CO_2 fixation (11, 14). The microarray results show clearly that the structural genes encoding form I and II RubisCO were differentially expressed: *cbbM* (Tbd2638, which encodes form II) was upregulated 6-fold under denitrifying conditions (Table 1) whereas *cbbL* and *cbbS* (Tbd2624-2623, which encode the large and small subunits of form I) were upregulated 7.4- and 6.5-fold, respectively, under aerobic conditions (Table 2). The other *cbb* genes included in the form I and form II RubisCO gene clusters were also differentially expressed, albeit to a lesser extent than the structural genes. Thus, *cbbQ* and *cbbO* (Tbd2637 and Tbd2636) in the form II cluster were upregulated 2.7- and 3.5-fold under denitrifying conditions, whereas their homologs in the form I cluster (Tbd2622 and Tbd2621) were upregulated 5.5- and 2.6-fold under aerobic conditions.

These results are consistent with the biochemical characterization of form I and II RubisCO in *T. denitrificans* with respect to their relative affinity to CO_2 and O_2. Molecular oxygen competes with CO_2 for the active site of RubisCO and thereby decreases its efficiency for carbon fixation. The relative specificity of RubisCO enzymes for CO_2 and O_2 (the CO_2/O_2 specificity factor, or τ) was determined in *T. denitrificans* (14); form I was shown to have

considerably higher CO_2/O_2 specificity ($\tau = 46$) than form II ($\tau = 14$). Thus, expressing form I under aerobic conditions would tend to maximize the efficiency of CO_2 fixation.

We are not aware of any previous studies of differential expression of form I and II RubisCO under aerobic vs. denitrifying conditions. The most relevant studies are those that investigated expression of form I and II RubisCO under a variety of chemoautotrophic, chemoheterotrophic, photoautotrophic, and photoheterotrophic conditions in *Rhodobacter sphaeroides* and *Rhodobacter capsulatus* (recently reviewed in ref. 10). Differential transcription of forms I and II was observed in some of these studies. In the absence of more experimental data for *T. denitrificans*, these existing studies allow us only to speculate about regulatory systems that might be involved in differential transcription of form I and II RubisCO in *T. denitrificans*.

RegB/RegA is a global, two-component, redox-responsive regulatory system that appears to have a role in differential expression of form I and II RubisCO in *Rhodobacter* species (10 and references therein). For example, in work with *regA* mutants of *R. sphaeroides* grown under aerobic, chemoautotrophic conditions, Gibson et al. (12) indicated that RegA (PrrA) functioned as a strong activator of form II RubisCO genes but had no effect on, or acted as a mild repressor of, the form I genes. However, rocket electroimmunoassay studies of *R. sphaeroides* strain HR-CAC showed that approximately 2.5-fold more form I than form II RubisCO protein was expressed under aerobic chemolithoautotrophic conditions (25). RegA also influences the differential, redox-responsive transcription of other genes including those associated with photosynthesis, cytochrome cbb_3 oxidase, and Cu-containing nitrite reductase (*nirK*).

Several lines of evidence suggest that the RegB/RegA system is present in *T. denitrificans* and may contribute to transcriptional regulation of RubisCO genes: (1) genes

1 putatively encoding RegA and RegB have been identified in *T. denitrificans* (Tbd2690 and
2 Tbd2689, respectively), (2) possible RegA-binding sites are present upstream of *cbbM* and *cbbL*,
3 and (3) at least one putative RegA-binding site is present in the intergenic region upstream of the
4 aerobically upregulated *ccoN* gene (Tbd0643), which encodes a subunit of cytochrome cbb_3
5 oxidase and has been associated with RegB/RegA regulation in *R. capsulatus* (29, 30). BLASTP
6 searches for RegA in the *T. denitrificans* genome using the RegA (PrrA) sequence from *R.*
7 *sphaeroides* (GenBank YP_351562) revealed that Tbd2690 was the best match; the deduced
8 amino acid sequence of Tbd2690 shares 51% identity with the RegA sequence of *R. sphaeroides*.
9 Alignment of these (and other) RegA sequences showed that the *T. denitrificans* homolog also
10 includes the highly conserved helix-turn-helix DNA-binding motif described for a range of RegA
11 homologs (10). RegB is putatively encoded by Tbd2689 in *T. denitrificans* (26% sequence
12 identity with the RegB sequence of *R. sphaeroides*; GenBank YP_351564). Alignment of the
13 deduced amino acid sequence of Tbd2689 with known RegB sequences revealed that the *T.*
14 *denitrificans* homolog contains a highly conserved, redox-active cysteine residue that has been
15 shown to exert control over the activity of the sensor kinase in *R. sphaeroides* (31). Searches for
16 RegA-binding sites upstream of *cbbM* and *cbbL* in *T. denitrificans* revealed possible degenerate
17 sequences. Laguri et al. (22) described the following main features of RegA-binding sites
18 derived from studies of *R. sphaeroides* and *R. capsulatus*: (i) a palindromic 5'-
19 GCGNC…GNCGC-3' consensus, (ii) a central AT-rich section, and (iii) a variable number of
20 bases between the 5' and 3' palindromic regions (with an apparent total of 9 to 15 bases in the
21 binding site motif). Sequences conforming to these characteristics were found upstream of *cbbM*
22 (5'-GCGACAGCCGC-3') and *cbbL* (5'-GCGCCTCTTGTCGC-3'). Notably, both of these
23 putative RegA-binding sites were located at least 950 nt upstream of the translational start sites

1 of *cbbM* and *cbbL* and occurred in a complementary *cbbR* coding region (i.e., in a *cbbR* coding
2 region on the opposite strand from *cbbM* and *cbbL*). Since the RegA-binding consensus features
3 were based on only two bacterial species, it is possible that other RegA-binding sites occur
4 upstream of *cbbM* or *cbbL* but could not be detected because they diverge from *Rhodobacter*
5 motifs.

6 Transcriptional regulation of RubisCO genes is characteristically complex and is
7 controlled by more than just the RegB/RegA system. For example, there is undoubtedly also
8 some positive control of form I and II RubisCO expression by the LysR-type transcriptional
9 regulator CbbR (e.g., 21). Both the form I and form II operons in *T. denitrificans* are adjacent to
10 divergently transcribed *cbbR* genes (5), and multiple putative CbbR-binding sites were found in
11 upstream regions of both *cbbM* and *cbbL*. To illustrate, in the intergenic region between *cbbL*
12 and the upstream *cbbR* gene, there were putative, often overlapping CbbR-binding sites located
13 from 4 - 17 nt and 89 - 152 nt upstream from the translational start site. In the intergenic region
14 between *cbbM* and the upstream *cbbR* gene, there were putative CbbR-binding sites located from
15 4 - 24 nt and 64 - 127 nt upstream from the translational start site. The motif used to identify
16 putative CbbR-binding sites was T-N_{12}-A, which deviates from the T-N_{11}-A motif characteristic
17 of LysR-type transcriptional regulators but may be more applicable to CbbR-binding sites in
18 autotrophic bacteria (21). There is currently no evidence suggesting that CbbR influences
19 differential expression of RubisCO genes under aerobic vs. denitrifying conditions, and it is very
20 possible that as yet unidentified transcriptional regulators may influence expression of RubisCO
21 genes (9, 12).

22 **Cytochrome *cbb*$_3$ oxidase.** The gene cluster Tbd0643-0637, which includes genes
23 putatively encoding one of two cytochrome *cbb*$_3$ oxidases in *T. denitrificans* (5), was upregulated

1 under aerobic conditions (Table 2; Fig. 1). The first four genes in this cluster (Tbd0643-
2 Tbd0640) appear to be *ccoNOQP*, and the entire cluster is highly similar in terms of gene
3 sequence and organization to a cluster in the related β-proteobacterium *Azoarcus* sp. strain
4 EbN1. Throughout this cluster in *T. denitrificans*, genes were upregulated 3.9- to 13.5-fold
5 relative to denitrifying conditions. The highest upregulation was for *ccoN* (9.6-fold) and *ccoQ*
6 (13.5-fold).

7 The results for *T. denitrificans* are generally consistent with those from *ccoN*::*lacZ*
8 transcriptional fusion studies of *R. capsulatus* and *R. sphaeroides*, which showed greater
9 expression of *ccoN* under aerobic, and particularly microaerophilic, conditions compared to
10 anaerobic conditions (24, 29, 30). Studies with *regA* mutants of *R. capsulatus* suggest that RegA
11 activates cytochrome *cbb*$_3$ oxidase expression semiaerobically or aerobically but represses
12 expression anaerobically (29, 30). In contrast, FnrL apparently activates cytochrome *cbb*$_3$
13 oxidase expression semiaerobically or anaerobically in these two *Rhodobacter* species (24, 29).

14 The promoter region upstream of *ccoN* was examined for potential RegA- and FNR-
15 binding sites, as these transcription factors have been implicated in the regulation of cytochrome
16 *cbb*$_3$ oxidases in *Rhodobacter* species. We focused on the promoter region of *ccoN* because it is
17 the first gene in this cluster and its upstream intergenic region is nearly 500 nt long, whereas the
18 intergenic regions upstream of *ccoO*, *Q*, and *P* only range from 0 to 11 nt. A probable RegA-
19 binding site (5'-GCGACACGTTGGCGC-3') was identified upstream of *ccoN*; this putative
20 binding site was located much closer to the translational start site (ca. 280 nt upstream) than
21 those we have identified in promoter regions of *cbbL* and *cbbM* (discussed previously). The
22 most likely FNR-binding site identified in the *ccoN* promoter region was TTGAT . . . cTCgc,

1 which was notably degenerate at the 3' end and was located 374 nt upstream of the translational
2 start site.
3 **Chaperones and stress proteins.** A number of genes associated with protein folding
4 and turnover were upregulated under aerobic conditions (Table 2; Figure 1). These include the
5 genes encoding the molecular chaperones ClpB (Tbd0815; 9.6-fold upregulated), GroEL and
6 GroES (Tbd0091-0092; 9.1- and 5.1-fold upregulated, respectively), GrpE, DnaK, and DnaJ
7 (Tbd1537-1539; 5.1- to 9.2-fold upregulated) and IbpA (Tbd1370; 6.3-fold upregulated). Several
8 genes occurring in a cluster with GroEL and GroES were also aerobically upregulated, albeit to a
9 lesser extent (Tbd0094-0096; 2.1- to 2.6-fold). Other aerobically upregulated genes encoding
10 proteins that are putatively associated with protein folding and turnover include genes for HtpG
11 (Tbd1078; 18-fold upregulated) and Lon protease (Tbd1252; 11-fold upregulated) (Table 2).
12 Several of these genes have been found to be regulated in *Escherichia coli* by sigma 32, the heat
13 shock/stress alternative sigma factor (Tbd0345). Sigma 32-regulated genes include *clpB*, *grpE*,
14 *dnaJ/dnaK*, *ibpA*, *htpG*, and *lon* (35). In turn, several of the proteins encoded by these genes
15 regulate intracellular levels of sigma 32, as do GroEL and GroES (13).
16 GroEL, an essential chaperone, and DnaK have been shown to play a significant role in
17 the viability of *E. coli* (16). In *E. coli*, it has been demonstrated that about 250 proteins interact
18 with GroEL, of which several could also utilize DnaK for proper folding (19). In the current
19 study, both DnaK and GroEL were found to be significantly upregulated under aerobic
20 conditions along with form I RubisCO, which has been shown to be a substrate of GroEL (17).
21 **Iron acquisition.** A cluster of 21 genes (Tbd0705 – Tbd0725), many of which are
22 associated with Fe^{3+} uptake (5), includes the 16 most aerobically upregulated genes observed in
23 this study (Table 2). In fact, all 21 genes in the cluster are among the top 50 aerobically

1 upregulated genes (Table 2). The level of upregulation within the cluster varies widely, ranging

2 from 6.1-fold upregulation for Tbd0707 to 159-fold upregulation for Tbd0725. Aerobic

3 upregulation of iron transport genes in bacteria occurs in response to limited iron availability due

4 to the lower solubility of Fe(III) species compared to Fe(II) species (1). To illustrate for the

5 conditions used in this study, although the amounts of iron added to the aerobic and denitrifying

6 cultures were similar (10 and 7.5 μM, respectively), equilibrium geochemical modeling (26)

7 indicated that the amounts of dissolved iron under these two conditions differed dramatically.

8 Whereas all of the 7.5 μM iron would be present in solution under denitrifying conditions (~75%

9 as $FeHCO_3^+$), less than 0.7 μM would be soluble under aerobic conditions [>93% of the Fe

10 would be present as $Fe(OH)_3$ precipitate].

11 Genes found in the cluster include those that encode proteins involved in siderophore

12 biosynthesis and export (Tbd0716-0721), Fe^{3+}-siderophore uptake across the outer membrane

13 (Tbd0711 – 0713, Tbd0715, Tbd0722), iron storage and mobility (Tbd0705), and heme uptake

14 (Tbd0725). Systems involved in iron acquisition have been found to be regulated by the ferric

15 uptake regulator protein, Fur (Tbd1123), which acts as a repressor in the presence of Fe^{2+} and a

16 de-repressor in the absence of Fe^{2+} (1). Putative DNA binding sites allowing for Fur-dependent

17 regulation (1, 34) were identified upstream of two genes in the cluster (Tbd0725 and Tbd0715)

18 and overlapped with an *E. coli*-type sigma 70 promoter sequence for both genes (Fig. 4). The

19 most highly upregulated gene, Tbd0725, encodes a putative homolog of HemP, a Fur-regulated

20 protein associated with heme uptake in *Yersinia enterocolitica* (28). Although other genes

21 associated with heme uptake are not found in this gene cluster, they are found scattered

22 throughout the *T. denitrificans* genome. The second gene, Tbd0715, encodes a homolog of

23 PsuA, a Fur-regulated Fe^{3+}-siderophore outer-membrane receptor in *Vibrio parahaemolyticus*

found associated with siderophore biosynthesis genes similar to those occurring in this *T. denitrificans* cluster (32).

Sequences closely matching the *E. coli* sigma 70 “consensus” sequences were also found in promoter regions of several other aerobically upregulated genes (Table 3), including *bfd* (Tbd0705), *groES* (Tbd0092), *lon* (Tbd1252), *dnaK* (Tbd1538), and *ccoN* (Tbd0643). As discussed previously, these genes encode proteins involved in a variety of functions including iron uptake and storage, protein folding and turnover, and aerobic respiration via cytochrome cbb_3 oxidase. In contrast, promoter regions for anaerobically upregulated genes, including those involved in denitrification (*narK*, *nirS*, and *norC*) and carbon fixation (*cbbM*), did not have sequences that closely matched the *E. coli* consensus sequence. Interestingly, the promoter region for *nosZ*, which was slightly upregulated aerobically despite being associated with denitrification (Fig. 2), also contains sequences similar to the *E. coli* sigma 70 consensus sequences (Table 3).

RT-qPCR validation of microarray trends. Twelve genes were selected for analysis by RT-qPCR to confirm that differential expression indicated by the microarray data was supported by an independent method. The selected genes (listed in the legend for Fig. 5) cover a wide range of expression and include genes that were most upregulated under aerobic conditions and under denitrifying conditions. Overall, the RT-qPCR data and microarray data were very consistent (Fig. 5); the data were highly correlated ($r^2 = 0.95$) and had a slope that approached unity (1.085).

Concluding remarks. As one of the first whole-genome transcriptional studies of a chemolithotrophic bacterium, and one of the few studies addressing transcriptional analysis of genes associated with chemolithotrophic sulfur-compound oxidation, this study provides a

number of novel findings, including the following: (i) strong upregulation under denitrifying conditions of two copies of *sqr* (which is explained by genomic location adjacent to the *nar* gene cluster for only one *sqr* copy), (ii) a variety of expression behaviors for the eight *dsrC* copies (ranging from aerobic upregulation to anaerobic upregulation to constitutive expression at a high level), (iii) consistently high-level expression under aerobic and denitrifying conditions of several important gene clusters associated with sulfur-compound oxidation (including *soxXYZA*, *dsrABEFHCMKLJOP*, and the genes encoding ATP sulfurylase and APS reductase), (iv) differential expression of genes putatively encoding rhodanese (an enzyme function previously lacking direct evidence for its involvement in thiosulfate oxidation), and (v) differential expression of Form I and II RubisCO under aerobic vs. denitrifying conditions. Whereas this study provides some insight into the unusual ability of *T. denitrificans* to oxidize sulfur compounds under aerobic and denitrifying conditions, additional whole-genome transcriptional studies by our group will provide information on other unusual abilities of this bacterium, namely, catalysis of anaerobic, nitrate-dependent Fe(II) and U(IV) oxidation. Combining these microarray results with the use of a newly developed genetic system in *T. denitrificans* (T. Letain, S. Kane, T. Legler, H. Beller, E. Salazar, and P. Agron; unpublished data) will facilitate better understanding of the biochemical and genetic basis of the oxidative metabolism of this widespread but unusual bacterium.

ACKNOWLEDGMENTS

We thank NimbleGen Systems, Inc. and Feliza Bourguet (Lawrence Livermore National Laboratory) for their valuable experimental contributions, Donovan Kelly (University of Warwick, UK) for insightful comments on the manuscript, and Alex Beliaev (Pacific Northwest National Laboratory) for useful discussions about RNA handling.

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

REFERENCES

1. **Andrews, S. C., A. K. Robinson, and F. Rodriguez-Quiñones.** 2003. Bacterial iron homeostasis. FEMS Microbiol. Rev. **27**:215-237.

2. **Arai, H., M. Mizutani, and Y. Igarashi.** 2003. Transcriptional regulation of the *nos* genes for nitrous oxide reductase in *Pseudomonas aeruginosa*. Microbiology (UK) **149**:29-36.

3. **Baumann, B., M. Snozzi, A. J. B. Zehnder, and J. R. van der Meer.** 1996. Dynamics of denitrification activity of *Paracoccus denitrificans* in continuous culture during aerobic-anaerobic changes. J. Bacteriol. **178**: 4367-4374.

4. **Beller, H. R.** 2005. Anaerobic, nitrate-dependent oxidation of U(IV) oxide minerals by the chemolithoautotrophic bacterium *Thiobacillus denitrificans*. Appl. Environ. Microbiol. **71**:2170-2174.

5. **Beller, H. R., P. S. G. Chain, T. E. Letain, A. Chakicherla, F. W. Larimer, P. M. Richardson, M. A. Coleman, A. P. Wood, and D. P. Kelly.** 2006. The genome sequence of the obligately chemolithoautotrophic, facultatively anaerobic bacterium *Thiobacillus denitrificans*. J. Bacteriol. **188**:1473-1488.

6. **Benjamini, Y., and Y. Hochberg.** 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B **57**:289-300.

7. **Bolstad, B. M., R. A. Irizarry, M. Astrand, and T. P. Speed.** 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics **19**:185-193.

8. **Cuypers, H., J. Berghöfer, and W. G. Zumft.** 1995. Multiple *nosZ* promoters and anaerobic expression of *nos* genes necessary for *Pseudomonas stutzeri* nitrous oxide reductase and assembly of its copper centers. Biochim. Biophys. Acta **1264**:183-190.

9. **Dubbs, J. M., and F. R. Tabita.** 2003. Interactions of the cbb_{II} promoter-operator region with CbbR and RegA (PrrA) regulators indicate distinct mechanisms to control expression of the two cbb operons of *Rhodobacter sphaeroides*. J. Biol. Chem. **278**:16443-16450.

10. **Elsen, S., L. R. Swem, D. L. Swem, and C. E. Bauer.** 2004. RegB/RegA, a highly conserved redox-responding global two-component regulatory system. Microbiol. Mol. Biol. Rev. **68**:263-279.

11. **English, R. S., C. A. Williams, S. C. Lorbach, and J. M. Shively.** 1992. Two forms of ribulose 1,5-bisphosphate carboxylase/oxygenase from *Thiobacillus denitrificans*. FEMS Microbiol. Lett. **94**:111-119.

12. **Gibson, J. L., J. M. Dubbs, and F. R. Tabita.** 2002. Differential expression of the CO_2 fixation operons of *Rhodobacter sphaeroides* by the Prr/Reg two-component system during chemoautotrophic growth. J. Bacteriol. **184**:6654-6664.

13. **Guisbert, E., C. Herman, C. Z. Lu, and C. A. Gross.** 2004. A chaperone network controls the heat shock response in *E. coli*. Genes & Dev. **18**:2812-2821.

14. **Hernandez, J. M., S. H. Baker, S. C. Lorbach, J. M. Shively, and F. R. Tabita.** 1996. Deduced amino acid sequence, functional expression, and unique enzymatic properties of

1 the Form I and Form II ribulose bisphosphate carboxylase/oxygenase from the
2 chemoautotrophic bacterium *Thiobacillus denitrificans*. J. Bacteriol. **178**: 347-356.
3 15. **Honisch, U., and W. G. Zumft.** 2003. Operon structure and regulation of the *nos* gene
4 region of *Pseudomonas stutzeri*, encoding an ABC-type ATPase for maturation of nitrous
5 oxide reductase. J. Bacteriol. **185**: 1895-1902.
6 16. **Horwich, A. L., K. B. Low, W. A. Fenton, I. N. Hirshfield, and K. Furtak.** 1993.
7 Folding *in vivo* of bacterial cytoplasmic proteins: role of GroEL. Cell **74**:909–917.
8 17. **Houry, W. A., D. Frishman, C. Eckerskorn, F. Lottspeich, and F. U. Hartl.** 1999.
9 Identification of *in vivo* substrates of the chaperonin GroEL. Nature **402**:147-154.
10 18. **Justin, P., and D. P. Kelly.** 1978. Metabolic changes in *Thiobacillus denitrificans*
11 accompanying the transition from aerobic to anaerobic growth in continuous chemostat
12 culture. J. Gen. Microbiol. **107**:131-137.
13 19. **Kerner, M. J., D. J. Naylor, Y. Ishihama, T. Maier, H. C. Chang, A. P. Stines, C.**
14 **Georgopoulos, D. Frishman, M. Hayer-Hartl, M. Mann, and F. U. Hartl.** 2005.
15 Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. Cell
16 **122**:209-220.
17 20. **Körner, H., and W. G. Zumft.** 1989. Expression of denitrification enzymes in response
18 to the dissolved oxygen level and respiratory substrate in continuous culture of
19 *Pseudomonas stutzeri*. Appl. Environ. Microbiol. **55:**1670-1676.
20 21. **Kusian, B., and B. Bowien.** 1997. Organization and regulation of *cbb* CO_2 assimilation
21 genes in autotrophic bacteria. FEMS Microbiol. Rev. **21**:135-155.
22 22. **Laguri, C., M. K. Phillips-Jones, and M. P. Williamson.** 2003. Solution structure and
23 DNA binding of the effector domain from the global regulator PrrA (RegA) from

1 *Rhodobacter sphaeroides*: insights into DNA binding specificity. Nucleic Acids Res.
2 **31**:6778-6787.
3 23. **Moir, J. W. B., D. J. Richardson, and S. J. Ferguson.** 1995. The expression of redox
4 proteins of denitrification in *Thiosphaera pantotropha* grown with oxygen, nitrate, and
5 nitrous oxide as electron acceptors. Arch. Microbiol. **164:**43-49.
6 24. **Mouncey, N. J., and S. Kaplan**. 1998. Oxygen regulation of the *ccoN* gene encoding a
7 component of the cbb_3 oxidase in *Rhodobacter sphaeroides* $2.4.1^T$: involvement of the
8 FnrL protein. J. Bacteriol. **180**:2228-2231.
9 25. **Paoli, G. C., and F. R. Tabita.** 1998. Aerobic chemolithoautotrophic growth and
10 RubisCO function in *Rhodobacter capsulatus* and a spontaneous gain of function mutant
11 of *Rhodobacter sphaeroides*. Arch. Microbiol. **170**:8-17.
12 26. **Parkhurst, D.L., and C. A. J. Appelo.** 1999. User's Guide to PHREEQC (Version 2) –
13 A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and
14 Inverse Geochemical Calculations, U.S. Geological Survey Water-Resources
15 Investigations Report 99-4259, 311 p.
16 27. **Schedel, M. and H. G. Trüper.** 1979. Purification of *Thiobacillus-denitrificans*
17 siroheme sulfite reductase and investigation of some molecular and catalytic properties.
18 Biochim. Biophys. Acta **568**:454-466.
19 28. **Stojiljkovic, I., and K. Hantke.** 1992. Hemin uptake system of *Yersinia enterocolitica*:
20 similarities with other TonB-dependent systems in Gram-negative bacteria. EMBO J.
21 **11**:4359-4367.

1 29. **Swem, D. L., and C. E. Bauer.** 2002. Coordination of ubiquinol oxidase and cytochrome
2 *cbb*$_3$ oxidase expression by multiple regulators in *Rhodobacter capsulatus*. J. Bacteriol.
3 **184**:2815-2820.
4 30. **Swem, L. R., S. Elsen, T. H. Bird, D. L. Swem, H.-G. Koch, H. Myllykallio, F.
5 Daldal, and C. E. Bauer.** 2001. The RegB/RegA two-component regulatory system
6 controls synthesis of photosynthesis and respiratory electron transfer components in
7 *Rhodobacter capsulatus*. J. Mol. Biol. **309**:121-138.
8 31. **Swem, L. R., B. J. Kraft, D. L. Swem, A. T. Setterdahl, S. Masuda, D. B. Knaff, J.
9 M. Zaleski, and C. E. Bauer.** 2003. Signal transduction by the global regulator RegB is
10 mediated by a redox-active cysteine. EMBO J. **22**:4699-4708.
11 32. **Tanabe, T., T. Funahashi, H. Nakao, S. Miyoshi, S. Shinoda, and S. Yamamoto.**
12 2003. Identification and characterization of genes required for biosynthesis and transport
13 of the siderophore vibrioferrin in *Vibrio parahaemolyticus*. J. Bacteriol. **185**:6938-6949.
14 33. **Trüper, H. G.** 1994. Reverse siroheme sulfite reductase from *Thiobacillus denitrificans*.
15 Meth. Enzymol. **243**: 422-426.
16 34. **Wan, X., N. C. VerBerkmoes, L. A. McCue, D. Stanek, H. Connelly, L. J. Hauser, L.
17 Wu, X. Liu, T. Yan, A. Leaphart, R. L. Hettich, J. Zhou, and D. K. Thompson.** 2004.
18 Transcriptomic and proteomic characterization of the Fur modulon in the metal-reducing
19 bacterium *Shewanella oneidensis*. J. Bacteriol. **186**:8385-8400.
20 35. **Zhao, K., M. Liu, and R. R. Burgess.** 2005. The global transcriptional response of
21 *Escherichia coli* to induced σ^{32} protein involves σ^{32} regulon activation followed by
22 inactivation and degradation of σ^{32} *in vivo*. J. Biol. Chem. **280**:17758-17768.

1 36. **Zumft, W. G.** 1997. Cell biology and molecular basis of denitrification. Microbiol. Mol.
2 Biol. Rev. **61**:533-616.

TABLE 1. Top 50 ORFs upregulated under denitrifying conditions.

Tbd ORF	Fold upreg-ulation	p-value[a]	Category[b]	Annotation[c]
1402	95	7.5E-37	DN NO3	$NarK_2$ protein; Nitrate/proton antiporter
1404	94	2.4E-34	DN NO3	NarH protein, Nitrate reductase beta subunit
1403	89	1.8E-33	DN NO3	NarG protein, Nitrate reductase alpha subunit
1401	58	5.4E-56	DN NO3	NarK protein; Nitrate/proton symporter
1406	57	3.8E-52	DN NO3	NarI protein, Nitrate reductase gamma subunit
1407	55	3.4E-64	S	Sulfide:quinone oxidoreductase
1405	54	9.5E-52	DN NO3	NarJ protein, Nitrate reductase chaperonin
2200	26	6.5E-36		Fimbrial protein PilA
2774	25	2.0E-46	NULL	Null
0077	21	2.9E-40	DN NO2	NirS (cytochrome cd_1); Nitrite reductase
0073	20	1.4E-31	DN NO2	NirG protein
2012	18	1.4E-37	S	Sulfate thiol esterase (tentative)
0075	14	1.5E-39	DN NO2	NirF protein
0072	14	1.4E-38	DN NO2	Nitrite reductase heme biosynthesis protein NirH
1408	14	4.0E-32	S	Putative DsrC protein
0076	13	2.2E-33	DN NO2	Probable NirC protein
1835	13	1.4E-37		Blr3518 protein
1649	13	9.9E-33	S?	Null
0071	12	1.6E-37	DN NO2	Heme D1 biosynthesis protein NirJ
0070	12	2.2E-37	DN NO2	Probable NirN protein

2226	11	7.6E-32	S?	Null
2175	11	3.5E-35		Hypothetical protein
0074	11	8.9E-40	DN NO2	NirD protein
2688	10	2.4E-51		TonB-dependent receptor protein
0561	9.5	4.8E-22	DN NO	Nitric oxide reductase subunit B (NorB)
1836	8.9	1.4E-28		Putative RNA polymerase sigma factor
0576	8.9	2.6E-38		Putative long-chain fatty acid transport protein
0559	8.9	1.9E-35	DN NO	Null
0562	8.7	1.3E-37	DN NO	Nitric oxide reductase subunit C (NorC)
1650	8.7	5.3E-27	S	Rhodanese-like domain protein
0558	8.3	1.2E-32	DN NO	Probable NorQ protein
1837	8.2	2.5E-40		Hypothetical signal peptide protein
0557	7.1	5.8E-40	DN NO	Hypothetical protein
0554	7.0	3.8E-34	DN NO?	Null
2225	6.5	4.7E-31	S	Sulfide:quinone oxidoreductase
2638	6.0	1.3E-24	CO2	RubisCO form II protein; CbbM
1501	5.9	1.1E-10	NULL	Null
2327	5.7	2.0E-20	S	Putative DsrC protein
2317	5.6	3.6E-35		Rubrerythrin/nigerythrin-like protein.
0555	5.5	7.7E-25	DN NO	Possible NorD protein
1742	5.4	5.4E-21		F_1-ATP synthase, epsilon subunit
1367	4.6	5.4E-26		Predicted periplasmic or secreted lipoprotein
0556	4.6	4.0E-27	DN NO	Iron-sulfur cluster-binding protein

1809	4.4	3.8E-37		Plasmid-related protein
1500	4.4	7.7E-19		V-type H^+-translocating pyrophosphatase
1388	4.3	3.2E-12		Cystathionine beta synthase domain protein
1739	4.3	6.8E-26	TT	Tetrathionate reductase subunit B
1741	4.3	6.3E-27	TT	Tetrathionate reductase subunit A
1499	4.3	8.3E-25		Hypothetical protein
0907	4.0	3.7E-27		Bacterial neuraminidase repeat

[a] p-value adjusted by procedures to control the False Discovery Rate criterion defined by Benjamini and Hochberg (6).

[b] Category definitions: DN NO3 (denitrification, nitrate reductase), DN NO2 (denitrification, nitrite reductase), DN NO (denitrification, nitric oxide reductase), S (sulfur-compound oxidation), TT (tetrathionate reductase), CO2 (CO_2 fixation via RubisCO); NULL (indicates no good hits and no clues from context); ? (indicates no good hits but location and expression suggest association with adjacent genes of known function).

[c] Best attempt at annotation based on examination of best BLASTP matches and genomic context.

TABLE 2. Top 50 ORFs upregulated under aerobic conditions.

Tbd ORF	**Fold upreg-ulation**	**p-value**[a]	**Category**[b]	**Annotation**[c]
0725	159	2.3E-78	Fe	Putative hemin uptake protein HemP
0706	117	2.0E-63	Fe	Anion-specific porin
0722	111	7.9E-59	Fe	Possible PvuA protein; polyhydroxycarboxylate siderophore uptake
0705	73	4.2E-51	Fe	Bacterioferritin-associated ferredoxin
0721	61	3.4E-54	Fe	Possible PvsA protein; ferritin biosynthesis protein
0710	57	2.1E-57	Fe	Putative lipoprotein
0715	46	7.5E-67	Fe	Possible PsuA protein; TonB-dependent receptor
0718	43	1.2E-59	Fe	Possible PvsD protein; ferritin biosynthesis protein
0712	40	8.5E-59	Fe	MotA/TolQ/ExbB proton channel family
0724	37	3.4E-43	Fe?	Null
0720	31	7.2E-46	Fe	Possible PvsB protein; ferritin biosynthesis protein
0709	27	2.0E-48	Fe?	Putative signal peptide protein
0723	27	7.3E-50	Fe	Possible high-affinity Fe^{2+}/Pb^{2+} permease
0711	25	5.7E-55	Fe	ExbD/TolR proton channel family
0716	24	6.2E-64	Fe	Putative aldolase; ferritin biosynthesis protein
0717	23	8.6E-41	Fe	Possible PvsE protein; ferritin biosynthesis

				protein
1078	18	3.4E-50	Chap	Heat shock family protein HtpG
0719	14	3.8E-52	Fe	Possible PvsC protein; siderophore efflux protein
0641	14	3.1E-34	CytOx	*cbb*$_3$-type cytochrome oxidase subunit IV (CcoQ)
2777	13	3.3E-35		Hypothetical protein
2778	12	4.7E-33		Predicted outer membrane lipoprotein
1252	11	8.1E-44	Chap	Probable ATP-dependent Lon protease
0815	9.6	1.1E-36	Chap	ClpB ATPase dependent protease
0708	9.6	1.6E-39	Fe?	Possible ApbE protein
0643	9.6	1.7E-41	CytOx	*cbb*$_3$-type cytochrome oxidase subunit I (CcoN)
0713	9.4	1.6E-35	Fe	Possible TonB-like energy transducer
1538	9.2	1.3E-17	Chap	Chaperone protein DnaK
2593	9.2	1.0E-31		Putative transcriptional regulator
0091	9.1	2.7E-30	Chap	Chaperonin GroEL (HSP60 family)
0640	8.8	7.1E-39	CytOx	*cbb*$_3$-type cytochrome oxidase subunit III (CcoP)
1053	7.8	2.9E-35		Hypothetical protein
2624	7.4	9.2E-29	CO2	RubisCO form I protein; CbbL (large subunit)
0638	7.4	1.6E-24	CytOx	*cbb*$_3$ cytochrome oxidase maturation protein CcoH
1420	7.4	1.1E-40		ABC transporter phosphate-binding protein
0714	7.1	5.1E-37	Fe	Uncharacterized iron-regulated membrane protein
1585	7.0	7.8E-39		Radical SAM enzyme of unknown function
1365	6.9	2.8E-29	S	Putative DsrC protein

2357	6.8	3.2E-26		Hypothetical protein
2356	6.6	6.6E-25		Hypothetical protein
2623	6.5	1.0E-27	CO2	RubisCO form I protein; CbbS (small subunit)
2398	6.4	7.4E-36		Hypothetical protein
0194	6.4	7.3E-31	NULL	
2401	6.4	4.9E-15		Thioredoxin domain-containing protein
1370	6.3	3.1E-16	Chap	Chaperone protein IbpA (small heat shock protein)
2270	6.1	4.6E-36		Porcine attaching-effacing associated protein variant 1
2355	6.1	5.9E-28	NULL	
0707	6.1	2.9E-40	Fe?	Probable transmembrane protein
0043	5.9	5.5E-32		Hypothetical protein
1539	5.7	3.1E-34	Chap	Chaperone protein DnaJ
2592	5.7	5.9E-29		Putative membrane protein

[a] p-value adjusted by procedures to control the False Discovery Rate criterion defined by Benjamini and Hochberg (6).

[b] Category definitions: Fe (iron acquisition), Chap (chaperones and stress proteins), CytOx (cytochrome oxidase), CO2 (CO_2 fixation via RubisCO), S (sulfur-compound oxidation), ? (indicates no good hits but location and expression suggest association with adjacent genes of known function).

[c] Best attempt at annotation based on examination of best BLASTP matches and genomic context.

TABLE 3. *E. coli*-like sigma 70 consensus sequences identified in promoter regions of selected genes that were upregulated aerobically in *T. denitrificans*.

Tbd ORF	gene	-35	-10
0705	*bfd*	TTGACA	TAGAAT
0715	*psuA*	TTGACA	TATTAT
0725	*hemP*	TTGACA	TATCAT
0092	*groES*	TTGAAA	TATTAT
1252	*lon*	TTGAAA	GATACT
1538	*dnaK*	TTGAAA	CATATT
0643	*ccoN*	TTGACA	TATATT
1389	*nosZ*	TAGACA	TACATG

FIGURE LEGENDS

FIG 1. Plot of $\log_2$ probability of differential expression vs. $\log_2$ fold differential expression for all genes identified in the *T. denitrificans* genome. The color coding corresponds to the major categories listed in Tables 1 and 2, namely, denitrification, sulfur-compound oxidation, CO_2 fixation via RubisCO (forms I and II), iron acquisition, cytochrome *cbb*$_3$ oxidase, and chaperones and stress proteins; all genes not falling within these categories in Tables 1 and 2, and all genes not included in Tables 1 and 2, are colored gray.

FIG. 2. Histogram displaying fold upregulation (denitrifying vs. aerobic conditions) for genes associated with denitrification, including *nar* cluster genes (Tbd1399-1406), *nir* cluster genes (Tbd0070-0079), *nor* cluster genes (Tbd0555-0562), and *nos* cluster genes (Tbd1389-1397). Structural *nir*, *nor*, and *nos* genes are labeled. Absolute expression levels for these genes are shown below the histogram, with green representing the lowest levels of expression, black representing intermediate levels, and red representing the highest levels. The plot of absolute expression levels shows all nine replicates for each condition.

FIG. 3. (A) Histogram of absolute expression levels of *sqr* (Tbd1407), *dsrC* (Tbd1408), and adjacent *nar* genes under aerobic and denitrifying conditions. (B) Partial sequence of the intergenic region between *sqr* and *dsrC*; the putative ribosomal binding site (RBS) and FNR box are indicated by underlining. (C) RT-qPCR results for cells exposed to aerobic and denitrifying conditions showing the relative number of transcripts that include *dsrC* (primers Tbd1408F and Tbd1408R2) or *dsrC* plus the intergenic region between *sqr* and *dsrC* (primers Tbd1407-1408IG_F and Tbd1408R). Numbers of transcripts are normalized to the largest value (*dsrC*; denitrifying conditions). Error bars represent one standard deviation based upon triplicate

analyses. (D) Electropherogram of RT-PCR products from RNA extracts for cells exposed to aerobic (Lanes 2-5) and denitrifying (Lanes 6-9) conditions. Lane 1, Hi-Lo DNA marker 50 bp-10 kbp (Bionexus Inc., Oakland, CA); Lanes 2 and 6, Tbd1407F/Tbd1408R2; Lanes 3 and 7, Tbd1407F/Tbd1407R; Lanes 4 and 8, Tbd1408F/Tbd1408R2; Lanes 5 and 9, Tbd1406F/Tbd1408R2. All bands represent cDNA amplicons of the expected length and are consistent with PCR reactions using genomic DNA as the template (not shown). No bands were visible for negative controls lacking reverse transcriptase (not shown).

FIG. 4. Nucleotide sequences in the promoter regions of the *hemP* (A) and *psuA* (B) genes. The putative -35 and -10 promoter sequences as well as the putative Fur box sequences for both genes are indicated.

FIG. 5. Correlation between aerobic fold upregulation as determined by RT-qPCR vs. microarray analysis for 12 genes: *narG* (Tbd1403), *nirS* (Tbd0077), *norB* (Tbd0561), *nosZ* (Tbd1389), *cbbS* (Tbd2623), *cbbL* (Tbd2624), *cbbM* (Tbd2638), *sqr* (Tbd1407), *ccoN* (Tbd0643), *dsrC* (Tbd1408), *dsrC* (Tbd1365), and *pvuA* (Tbd0722).

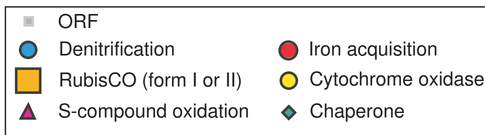
DENITRIFYING AEROBIC

200
150
100
50

Log$_2$ probability of differential expression

-10 -5 0 5 10

Log$_2$ fold differential expression



Fold upregulation (vs. aerobic)

100
80
60
40
20
0

nirS
norB
norC
nosZ

nar
nir
nor
nos

Denitrifying
Aerobic

A

Log$_2$ microarray intensity

15
12
9
6

Denitrifying
Aerobic

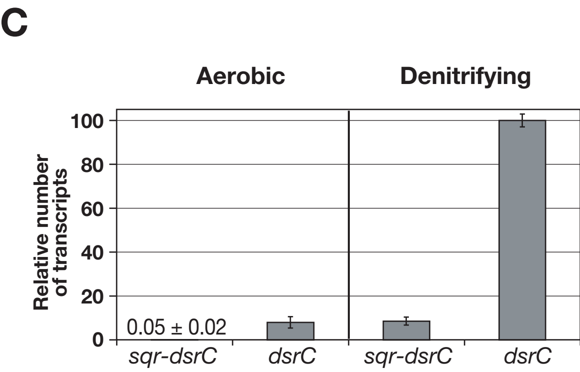
narK *narK$_2$* *narG* *narH* *narJ* *narI* *sqr* *dsrC*

B

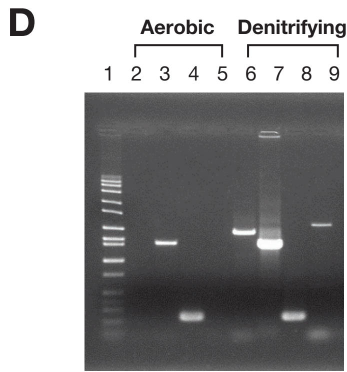
N$_{52}$-TTGAAATACATCTTTACGAGTCCATTCCGAATCCACCCACCTGAAGGAAAGCGAACATG

Putative FNR box **Putative RBS** **dsrC start**

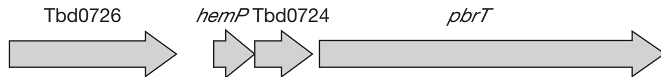
C

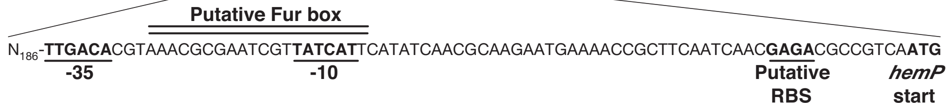


D

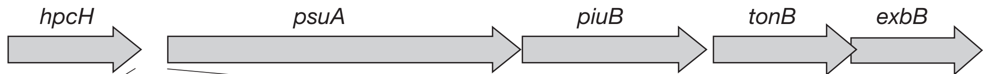


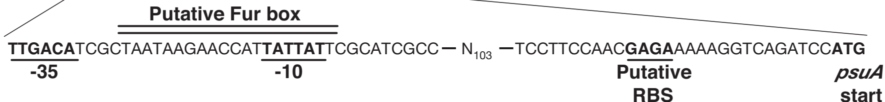
A

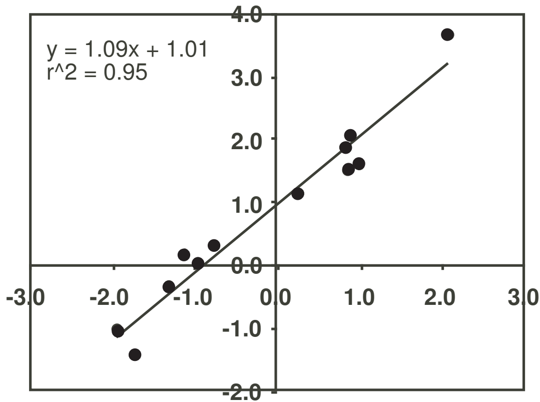




B






y = 1.09x + 1.01
r^2 = 0.95
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
RT-qPCR data (log fold difference)
Microarray data (log fold difference)