Title: A transcriptional map of respiratory versatility in the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*

Running title: Analysis of respiratory pathways in *P. aerophilum*

Aaron E. Cozen§, Matthew T. Weirauch†, Katherine S. Pollard††, David L. Bernick†, Joshua M. Stuart†, and Todd M. Lowe†*

§Department of Ocean Sciences, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

†Department of Biomolecular Engineering, University of California Santa Cruz 1156 High Street, Santa Cruz, CA 95064, USA

††Department of Statistics, University of California Davis One Shields Ave., Davis, CA 95616 USA

*author to whom correspondence should be addressed

Todd M. Lowe
telephone: (831) 459-1511
fax: (831) 459-4829
e-mail: lowe@soe.ucsc.edu
Contents

1 Introduction

2 Table 1 Genomic and respiratory characteristics of selected *Pyrobaculum* species

3 Materials and methods

4 Results & Discussion

5 References

6 Figure legends

Separately attached documents:

Table 2. Proposed functions and predicted orthologs for differentially expressed genes

Figure 1. Substrate usage

Figure 2. Aerobic cytochrome oxidases

Figure 3. Nitrate reductase and denitrification genes

Figure 4. Multicopper oxidase cluster

Figure 5. Molybdopterin oxidoreductases

Figure 6. Iron-induced genes

Figure 7. Iron-repressed genes

Figure 8. Summary schematic

Supplementary File S1. Gene expression heatmaps
Supplementary Table S1. *P. aerophilum* microarray features removed from analysis

Supplementary Table S2A. *P. aerophilum* microarray log2 ratios

Supplementary Table S2B. *P. aerophilum* microarray mean log2 ratios

Supplementary Table S2C. *P. aerophilum* microarray statistics

Supplementary Table S3. *Pyrobaculum* clade orthologs to *P. aerophilum* genes
Cozen et al. Analysis of respiratory pathways in *P. aerophilum*

**Abstract**

Hyperthermophilic crenarchaea in the genus *Pyrobaculum* are notable for respiratory versatility, but relatively little is known about the genetics or regulation of crenarchaeal respiratory pathways. We compared global transcriptional patterns in *Pyrobaculum aerophilum* cultures with oxygen, nitrate, arsenate and ferric iron as terminal electron acceptors using microarrays to identify genes and transcriptional patterns that differentiate these pathways. We also compared genome sequences for four closely-related species with diverse respiratory characteristics (*P. arsenaticum*, *P. calidifontis*, *P. islandicum*, and *Thermoproteus neutrophilus*) to identify conserved genes and potential regulatory motifs. Based on specific patterns of co-expression, gene clusters emerged that are associated with aerobic respiration, nitrate respiration, arsenate respiration, and anoxia. Functional predictions based on these patterns include separate cytochrome oxidases for aerobic growth and oxygen scavenging, a nitric-oxide-responsive transcriptional regulator, a multicopper oxidase involved in denitrification, and an archaeal arsenate respiratory reductase. A full-length ortholog of the predicted arsenate reductase appears to be required for growth of *P. calidifontis* on arsenate. We were unable to identify specific genes for iron respiration, but *P. aerophilum* exhibited a repressive transcriptional response to iron remarkably similar to those controlled by the ferric uptake regulator in bacteria. Together these findings present large number of functional and regulatory predictions for further study. The complete gene expression data set can be viewed in genomic context with the Archaeal Genome Browser at archaea.ucsc.edu.
Introduction

With the exception of the euryarchaeal Halobacteriales, many archaeal model organisms are either obligate aerobes (e.g. Sulfolobus spp.) or obligate anaerobes (e.g. the Thermococcales) with limited respiratory flexibility. As a consequence, relatively little is known about the genetics, regulation, and enzymology of archaeal respiratory pathways compared to those of bacteria. Species in the genus *Pyrobaculum* are notable both for their respiratory versatility, exemplified by *Pyrobaculum aerophilum*, and for diversity in respiratory capabilities between closely-related species (6, 28, 29, 31, 65), presenting a potential model system for studying respiratory pathways in the crenarchaea.

*Pyrobaculum* species grow optimally at 90-100°C and are common constituents of microbial communities in neutral pH geothermal springs and shallow marine hydrothermal vents (39, 42, 61). *Pyrobaculum aerophilum* is a facultative chemoautotroph that can use oxygen, nitrate, nitrite, arsenate, selenate, selenite, soluble ferric iron citrate, and insoluble ferric iron oxide as respiratory electron acceptors. The complete genome sequence of *P. aerophilum* was published in 2002 (23). The genome sequences of four additional *Pyrobaculum* species (*P. arsenaticum, P. calidifontis, P. islandicum*, and *Thermoproteus neutrophilus* [to be reclassified as a true *Pyrobaculum]*) have recently been completed (T.M. Lowe *et al.*, unpublished data), providing a valuable resource for comparing gene content to respiratory characteristics within a closely related group of crenarchaea Table 1.
A number of *Pyrobaculum* gene products have been characterized by heterologous expression or purification of native proteins, helping to identify genes for aerobic growth, oxidative stress, and denitrification (1, 4, 5, 17, 27, 44, 66). Although *Pyrobaculum* species are among the best-studied archaeal iron respirers, and are the only known archaeal arsenate respirers, the genes required for archaeal arsenate and iron respiration have not been identified (20, 21, 29, 31).

We used whole-genome DNA microarrays to compare gene expression in *P. aerophilum* cultures with oxygen, nitrate, arsenate, or ferric iron-citrate as terminal electron acceptors. The results show distinct transcriptional responses to these substrates. These support previous observations related to aerobic respiration, and suggest new candidate gene functions and regulatory relationships for denitrification and arsenate respiration. Genes for iron respiration remain enigmatic. However, expression patterns associated with exposure to millimolar iron suggest a transcriptionally regulated homeostatic response not previously described in the Archaea.
Table 1. Genomic and respiratory characteristics of selected *Pyrobaculum* species.

<table>
<thead>
<tr>
<th>species</th>
<th>genome Mbp</th>
<th>genome GC%</th>
<th>oxygen tolerance</th>
<th>terminal electron acceptors supporting growth</th>
<th>terminal electron acceptors tested: no growth</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. aerophilum</em></td>
<td>2.2</td>
<td>51%</td>
<td>facultative microaerobe</td>
<td>O$_2$, NO$_3^-$, NO$_2^-$, As(V), Se(VI), Se(IV), S$_2$O$_3^{2-}$§, Fe(III)-citrate, FeOOH</td>
<td>S$^+$, N$_2$O†</td>
<td>(2, 20, 21, 29, 31, 65)</td>
</tr>
<tr>
<td><em>P. arsenaticum</em></td>
<td>2.1</td>
<td>55%</td>
<td>strict anaerobe</td>
<td>S$^+$, S$_2$O$_3^{2-}$, As(V), Se(VI), Fe(III)-citrate, FeOOH</td>
<td>O$_2$, NO$_3^-$, Se(IV)</td>
<td>(21, 29)</td>
</tr>
<tr>
<td><em>P. calidifontis</em></td>
<td>2.0</td>
<td>57%</td>
<td>facultative aerobe</td>
<td>O$_2$, NO$_3^-$, As(V)*, Fe(III)-citrate, FeOOH</td>
<td>NO$_3^{2+}$††, S$^+$, S$_2$O$_3^{2-}$, SO$_3^{2-}$, SO$_4^{2-}$</td>
<td>(6, 21)</td>
</tr>
<tr>
<td><em>P. islandicum</em></td>
<td>1.8</td>
<td>50%</td>
<td>strict anaerobe</td>
<td>S$^+$, S$_2$O$_3^{2-}$, SO$_3^{2-}$, L-cystine, oxidized glutathione, As(V)§, Fe(III)-citrate, FeOOH</td>
<td>DL-lanthionine, fumarate, (CH$_3$)$_2$SO$_3$ S$_4$O$_6^{2-}$, SO$_4^{2-}$</td>
<td>(21, 28, 31)</td>
</tr>
<tr>
<td><em>Thermoproteus</em></td>
<td>1.8</td>
<td>60%</td>
<td>strict anaerobe</td>
<td>S$^+$</td>
<td>-</td>
<td>(22)</td>
</tr>
</tbody>
</table>

§ Weak growth; final densities ≤ 10$^7$ cells/ml

† *P. aerophilum* reduced NO$_3^-$ and NO$_2^-$ to N$_2$ during growth (2, 65), and N$_2$O reductase activity has been demonstrated in *P. aerophilum* cells (16), but N$_2$O provided as sole electron acceptor did not support growth (65).

†† *P. calidifontis* reduced NO$_3^-$ to N$_2$ during growth, but NO$_2^-$ as sole electron acceptor did not support growth (6).

* see Results & Discussion
**Materials and methods**

**Culture conditions for P. aerophilum**

*P. aerophilum* cultures derived from DSMZ strain 7523 were generously provided by Christopher House, Penn State University. Cultures of *P. aerophilum* were grown in media containing (per liter) 10g NaCl, 1.3g (NH4)2SO4, 0.28g KH2PO4, 0.25g MgSO4•7H2O, 1mg NaSeO4, 0.1mg Na2WO4•2H2O, 70mg CaCl2•2H2O, 0.5g yeast extract, 1mg (NH4)2Fe(SO4)2•6H2O, 0.5mg resazurin, and 10ml of a trace element solution. The trace element solution (DSM141) was prepared by combining the following components (per liter) in a 78.5μM solution of nitriloacetic acid adjusted to pH6.5: 3g MgSO4•7H2O, 0.5g MnSO4•H2O, 1g NaCl, 0.1g FeSO4•7H2O, 0.18g CoSO4•7H2O, 0.1g CaCl2•2H2O, 0.18g ZnSO4•7H2O, 0.01g CuSO4•5H2O, 0.02g KAl(SO4)2•12H2O, 0.01g H3BO3, 0.01g Na2MoO4•2H2O, 0.025g NiCl2•6H2O, and 0.0003g NaSeO3•5H2O. The complete media preparation was adjusted to pH 6.8 and sterilized by autoclaving.

All *P. aerophilum* cultures were incubated at 96°C, typically with 200ml of media in 0.5 liter wide mouth bottles under a gas headspace of N2, and shaking at 200rpm. For aerobic cultures, 5% (v/v) of atmospheric air was added to the headspace for a final concentration of approximately 1% O2. Three replicate respiratory induction experiments were performed in which cultures were grown with a 1% O2 gas headspace, followed by amendment with different respiratory substrates when the cultures consumed the initially supplied O2. No reducing agents were used. Depletion...
of O₂ in cultures was monitored by recording changes in the redox-indicator dye resazurin from pink to clear. After depletion of O₂ cultures were cooled to ambient temperature, diluted with a 25% volume of fresh growth media and dispensed in 200ml volumes to 0.5 liter bottles in an anaerobic chamber. Separate bottles were prepared for collection at each of three time points. Bottles received either 1% (v/v) additions from anoxic 1M stocks of NaNO₃, Na₂HAsO₄, or Fe(III)-citrate pH 6, 5% (v/v) of atmospheric air in the gas headspace, or no added electron acceptor in each induction experiment. These were re-equilibrated to growth temperature (96°C), and after 2.5 hours (T₁), 4.5 hours (T₂) or 7.5 hours (T₃) one bottle from each respiratory condition was collected for RNA preparation. Cells were collected for RNA analysis by centrifugation at 10,000g for 10 minutes. Cell pellets were frozen in liquid N₂ and stored at -80°C.

**Metabolite analyses**

Sub-samples for Fe(III) and Fe(II) analysis were collected directly into an equal volume of 1M HCl and stored at -20°C. Sub-samples for analysis of other metabolites were filtered using 0.2µm costar spin-x filters (Corning) and stored at -20°C. NO₂⁻ was measured spectrophotometrically by diazotizing with sulfanilamide followed by coupling with N-(1-naphthyl)ethylenediamine dihydrochloride using a Lachat Instruments Quickchem FIA+ autoanalyzer. Total concentrations of NO₃⁻ plus NO₂⁻ were measured by pre-reduction of NO₃⁻ on a cadmium column, and NO₃⁻ concentrations were calculated as the difference between reduced and untreated
samples (67). As(III) and total inorganic As were measured by hydride generation followed by inductively-coupled plasma optical emission spectrometry (ICP-OES) using a Perkin Elmer Optima 4300. Selective hydride generation from As(III) was achieved by diluting samples 1:100 in a 1M citrate-citric acid buffer pH 5.3 before mixing in equal volumes with a 1% NaBH$_4$, 0.5% NaOH solution to generate hydrides (7). Total inorganic As was measured by diluting samples 1:100 in a 5% solution of KI in 3.8M HCl to reduce As(V) to As(III) prior to hydride generation and ICP-OES. Fe(II) was measured using the ferrozine assay, and Fe(III) was measured indirectly as the difference between reduced and untreated samples (46).

**Microarray production and hybridization**

Microarray probes for the *P. aerophilum* genome were prepared by polymerase chain reaction (PCR) amplification of *P. aerophilum* genomic DNA using primers picked by a custom Perl wrapper program (T. M. Lowe, unpublished) utilizing the Primer 3 oligonucleotide selection software (48). Probe features included 2726 annotated open reading frames (ORFs), 243 potential open reading frames (dubbed “alternative ORFs”) (S. Fitz-Gibbon, personal communication), and 894 intergenic regions over 60 nucleotides (nt) long. Probes for ORFs and alternative ORFs were designed as 250-400 nt tags covering the most unique portion of genes, and were biased toward the 3’ end of transcripts. Probes for intergenic regions were tiled as overlapping segments of no more than 700 nt each. PCR-amplified probes were evaluated by gel electrophoresis. 93 loci (~2% of the total) with PCR probe
products of unexpected sizes or multiple bands were flagged and removed from analysis (supplementary Table S1). The majority of these flagged features are intergenic regions, repetitive gene families, and alternative ORFs. Probes were suspended in 3X SSC buffer, and printed in quadruplicate on GAPS II slides (Corning) which were stored in a desiccating chamber until use. A description of the microarray platform has been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO) database under the accession #GPL6755.

Total RNA was extracted from frozen cell pellets using a polytron tissue homogenizer and Tri reagent (Sigma-Aldrich). RNA samples were treated with DNase I (Ambion) to remove any residual DNA, re-extracted with Tri-reagent, and normalized to 1µg/µl before proceeding with preparation of cDNA for microarray analysis. Microarray analyses were performed for all 45 RNA samples generated in the three replicate induction experiments, using a pool of all samples as a universal reference for all hybridizations. Aminoallyl-labeled cDNAs were prepared from 5µg of sample RNA using 5µg of random hexamers (3µg/µl) and the Superscript III labeling kit (Invitrogen). These were subsequently labeled with cy3-NHS ester or cy5-NHS ester (GE Healthcare). Spotted microarray slides were hydrated for three minutes in a humidified chamber, snap-dried at 100°C, covalently linked with 250mJoules in a UV crosslinker (Stratagene), and incubated at 42°C in a solution of 5X SSC, 1% BSA and 0.1%SDS for 45 minutes prior to hybridization. For each hybridization Cy3-labeled cDNA derived from 5ug of experimental RNA and Cy5-
labeled cDNA derived from 5µg reference pool RNA were combined with 1mg/ml
herring sperm DNA, 0.38mg/ml yeast tRNA, 4X SSC and 0.5%SDS in a total volume
of 55µl, and hybridized for 14 to 18 hours at 60˚C in sealed hybridization chambers.
After hybridization, slides were washed for two minutes each in 0.05%SDS plus 0.1X
SSC, and 0.05X SSC, and dried by centrifugation. Slides were scanned at 5µm
resolution immediately after washing using a Genepix 4000B microarray scanner
(Molecular Devices).

Microarray Data Analysis

Microarray features were analyzed using Genepix 6 software (Molecular Devices). Median pixel intensities for each spotted feature were normalized using the
loess function of the open-source Bioconductor package “marray”, so that the
normalized global and print-tip median log2(Ch1/Ch2) ratio of each array was zero
(25). For each individual array feature the median intensity values of four replicate
spots were used for subsequent analyses. An intra-experimental normalization was
performed before statistical and clustering analyses by subtracting the average log2
ratio at each locus within the experiment from the log2 ratio representing each
condition and time point. Statistical analysis of differential expression was performed
using Edge Software and the SAMr package for R (57, 58, 63). A false discovery rate
(FDR) of 3% from Edge analysis was used as an upper limit for calling differentially
expressed genes statistically significant. Clustering of gene expression profiles and
production of heat maps was performed using Genesis software (59). Normalized
log2 ratios for all loci represented in the microarrays and results of statistical analyses are compiled in supplementary Tables S2A, S2B and S2C.

Northern analysis

RNA samples from each of the three replicate respiratory induction experiments were combined in equal amounts to generate 15 samples representing each condition and time point. 10 µg of total RNA from each of these samples were denatured for 45 minutes at 50°C with glyoxal loading buffer (Ambion) and resolved by electrophoresis in 1% agarose gels buffered with BPTE buffer (51). Size-separated RNAs were blotted onto Hybond N+ nylon membranes (GE healthcare) using capillary transfer. PCR products and primers identical to those used to generate microarray probes were used for single-stranded DNA probe preparation. Single-stranded DNA probes were prepared by linear PCR with Strip-EZ PCR reagents (Ambion) and body labeled using α-³²P ATP. Hybridizations were carried out at 42°C using Ultrahyb buffer (Ambion). Membranes were stripped of probes and prepared for re-hybridization using Strip EZ PCR reagents and protocols.

Computational genomics

Gene features and organization were evaluated using the Archaeal Genome Browser (53). Orthology was inferred by identifying reciprocal best BlastP hits (cutoff E=1x10⁻⁵) between genomes (3). Predicted orthologs for P. aerophilum genes in the other sequenced Pyrobaculum genomes are listed in supplementary Table S3.
Putative twin-arginine translocation signals were identified using TatP (9). The promoter regions of co-regulated genes were screened for possible regulatory motifs using BioProspector (32). The \textit{P. aerophilum} genome was screened for additional occurrences of identified motifs using position-specific scoring matrices (PSSM). Basal regulatory elements (BRE) and TATA boxes were identified using PSSMs (P.P. Chan and T.M. Lowe, unpublished data). Motif logo figures were generated using Weblogo (13).

\textbf{Culture conditions for \textit{P. arsenaticum} and \textit{P. calidifontis}}

\textit{P. arsenaticum} strain DSM 13514 was grown at 95°C in DSM390 medium reduced with 0.01% Na$_2$S, and maintained with 20mM Na$_2$S$_2$O$_3$ as terminal electron acceptor. 10mM NaNO$_2$ and 10mM NaNO$_3$ were tested as alternative terminal electron acceptors.

\textit{P. calidifontis} strain JCM 11548 was grown at 95°C in TY medium (10g tryptone, 2g yeast extract per liter) without Na$_2$S$_2$O$_3$, and maintained with 10mM NaNO$_3$ as terminal electron acceptor. 10mM Na$_3$HAsO$_4$ was tested as an alternative terminal electron acceptor. Solid media was prepared using 1.2% gelrite and 5 mM MgSO$_4$ as solidifying agents.

\textbf{Sequence analysis of the \textit{P. calidifontis} Pcal\_1601 locus}

\textit{P. calidifontis} cells were lysed using SNET lysis buffer, and genomic DNA was extracted using the phenol:chloroform:isoamyl alcohol method (51). Genomic
DNA was treated with RNaseA (Sigma) and re-extracted before analysis. A 1kb segment of the Pcal_1601 locus (chromosomal position 1489399-1490424) was PCR amplified using the forward primer 5’-GAGAGCTACACAGAGATCTTCG, the reverse primer 5’-GATAAGCACGAAGTTCTCAGG, and Phusion high fidelity polymerase (Finnzymes).

**Results and discussion**

Initial microarray experiments comparing gene expression patterns of *P. aerophilum* cultured to similar cell densities with nitrate, arsenate, or ferric citrate as terminal electron acceptors showed differences in gene expression of two-fold or greater at approximately one third of all loci (data not shown). We observed that growth rates and maximum densities varied with different terminal electron acceptors, which is consistent with other reports (20, 21), and attributed these global changes in gene expression to differences in growth phase as well as responses related to respiratory metabolism.

To focus on genes that are specifically affected by changes in terminal electron acceptors, we analyzed gene expression in five parallel time courses with cultures shifted from aerobic growth to either nitrate, arsenate, ferric iron-citrate, or additional O2. A group that received no additional terminal electron acceptor was also included. Gene expression and substrate usage were measured at three time points (2.5, 4.5, and 7.5 hours). This interval was designed to allow sufficient time for
changes in respiratory substrate usage while minimizing differences in growth phase between treatment groups.

Substrate usage and growth

Changes in media oxygenation, monitored using the redox-indicator dye resazurin, showed that cultures that received no additional electron acceptor and those amended with 1% O₂ depleted available oxygen at different stages in the time course. Cultures that received no terminal electron acceptor became anoxic between the first sampling time point (T₁) and the second (T₂). Cultures amended with 1% O₂ became anoxic between the second time point (T₂) and the third (T₃).

Analysis of culture media showed that cultures amended with nitrate, arsenate, or ferric citrate reduced each of these substrates during the time course. Cultures shifted to NO₃⁻ evolved approximately 1mM NO₂⁻ by T₃, measured spectrophotometrically (Figure 1A & 1B). Decreases in the total concentration of NO₃⁻ plus NO₂⁻ varied from 0mM to 0.5mM between replicate induction experiments, suggesting low and variable reduction of NO₂⁻ to gaseous NO, N₂O, or N₂ (Figure 1A). Concentrations of NO₃⁻ and NO₂⁻ in the cultures that were induced with O₂, As(V), Fe(III)-citrate or no electron acceptor were below 0.04mM, and no significant changes were detected over the time course (data not shown). Cultures shifted to As(V) reduced approximately 1mM As(V) to As(III) by T₃, as measured by ICP-OES (Figure 1C & 1D). Cultures shifted to Fe(III)-citrate evolved approximately 1mM Fe(II) by T₂ and 3.5mM Fe(II) by T₃, as measured by the ferrozine assay (Figure 1E).
With the exception of the Fe(III)-treated group, optical densities varied by less than 20% between all treatment groups by the end of the time course (Figure 1F).

Optical densities of cultures amended with O₂ or NO₃⁻ increased more than those amended with As(V) or those that received no additional electron acceptor. The growth media of the Fe(III)-treated cultures darkened substantially over the time course compared to un-inoculated Fe(III)-amended media, confounding the use the optical absorbance as a proxy for cell density in this group. However, total RNA yields and relative intensities for 16S and 23S rRNA bands were similar for cultures shifted to Fe(III) and those shifted to As(V), suggesting similar growth for these two groups (data not shown).

Identification of differentially expressed genes

Statistical analysis using Edge identified 543 features (14% of the total) with significant differences in expression (FDR <3%) between respiratory treatments. Hierarchical clustering of expression profiles from this group revealed clusters that are the focus of subsequent sections. Edge also identified 1013 loci (~27% of the total) with significant (FDR<1%) changes in gene expression over the time course, irrespective of respiratory treatment, suggesting that changes in growth phase between sampling time points affected gene expression at a large proportion of loci. The median range in expression values among loci in the respiratory response group was similar to that in the temporal response group (2.2-fold and 2.1-fold, respectively) and only slightly greater than loci in neither group (1.6-fold). However,
the majority of the most dynamic loci were highly ranked in the respiratory response group, and these showed a substantially greater range in expression (median 9.3-fold range for the top 50 respiratory response loci, compared to 3.4-fold for the top 50 temporal response genes, and 2.6-fold for the 50 most dynamic loci in neither group). Gene expression profiles for the complete set of loci in each group and the genome wide data set (supplementary File S1) can be viewed interactively using free genesis software (http://genome.tugraz.at/genesisserver/genesisserver_description.shtml) (59).

The ranked lists of differentially expressed loci reported by SAM were highly correlated with the lists reported by Edge ($R^2=0.92$ and $R^2=0.98$ for respiratory treatment and time course analyses, respectively). Results from statistical analyses and normalized log2 expression ratios from microarray analyses are compiled in supplementary Table S2A, S2B, & S2C. Microarray data produced from this study is deposited in the NCBI GEO database under the accession # GSE11366. All expression data and views of key respiration loci are also readily accessible in the Archaeal Genome Browser at archaea.ucsc.edu (53).

**Opposing regulation of aerobic cytochrome oxidases**

*P. aerophilum* grows micro-aerobically under gas headspace $O_2$ concentrations up to 5% (v/v) (65). The *P. aerophilum* genome contains two closely-spaced clusters of genes for aerobic cytochrome oxidase complexes (Figure 2A), described hereafter using ‘Sox’ nomenclature originating from studies of another
A crenarchaeal aerobe, *Sulfolobus acidocaldarius* (33, 34). One group of cytochrome oxidase genes encodes components of a SoxB-type enzyme (PAE1355 & PAE1357). Conserved domains in a flanking set of genes (PAE1358-PAE1362) suggest that these may be involved in biosynthesis of the SoxB complex. A second group of cytochrome oxidase genes (PAE1337-PAE1340) encodes a SoxM-type complex.

*Pyrobaculum oguniense* and *Pyrobaculum calidifontis* are both capable of growth under atmospheric O$_2$ concentrations, and each have syntenic clusters of genes orthologous to the cytochrome oxidase complexes found in *P. aerophilum* ([Table 2](#)) (6, 49). The SoxB complex, SoxM complex, superoxide dismutase, and catalase genes of *P. aerophilum* show strong similarity ($\leq 85\%$, $\leq 67\%$, 88%, and 88% identity, respectively) to those of its more aerotolerant relatives *P. oguniense* and *P. calidifontis*, revealing no obvious clues to its comparatively low O$_2$ tolerance.

Subunits I (PAE1357) and II (PAE1355) of the *P. aerophilum* soxB-type cytochrome oxidase complex were up-regulated under all conditions in the respiratory time course experiment except the first two time points after amendment with O$_2$, and all three time points after amendment with nitrate. These O$_2$ and nitrate-amended samples showed expression levels up to 15-fold lower ([Figure 2B](#) & [2D](#)). A small alternative ORF upstream of PAE1355 (PAE1353) and directly downstream of a strong promoter signal showed an identical expression pattern, suggesting that this may represent a small, co-transcribed gene. Northern analysis of PAE1355 showed a band at approximately 600nt, consistent with co-transcription of PAE1353 (144nt) and PAE1355 (477nt) ([Figure 2D](#)), and a minor band at 2.3kb (not shown), consistent
with co-transcription of these genes with PAE1357 (1674nt). Downstream genes
including a senC homolog (PAE1358) and several hypothetical proteins (PAE1359,
PAE1360, PAE1369) were expressed in a similar pattern (Figure 2B).

The second group of cytochrome oxidase genes was expressed in a pattern
that was essentially the opposite of the PAE1353-1357 SoxB-type cytochrome
oxidase (Figure 2C). These genes, annotated as a possible respiratory chain protein
(PAE1336), a SoxM-type fusion of cytochrome oxidase subunits I and III
(PAE1337), a cytochrome oxidase subunit II (PAE1338), and a SoxI homolog
(PAE1340) were up-regulated up to 11-fold in oxic (the first two time points
following induction with oxygen) versus anoxic samples, and up-regulated up to 16-
fold after induction with nitrate (Figure 2C, 2E & 2F). Genes encoding subunits of a
pyruvate dehydrogenase complex (PAE2644, PAE2646, PAE2648, PAE2649), as
well as an adjacent hypothetical protein with a putative DNA binding domain
(PAE2655), were up-regulated in a very similar pattern, suggesting enhanced
pyruvate consumption associated with expression of the SoxM cluster of genes
(Figure 2C).

Differential expression of the SoxM-type I/III fusion protein (PAE1337) was
not detected in the microarrays, but the co-expression of adjacent genes PAE1336,
PAE1338 and PAE1340 prompted us to verify the expression pattern for PAE1337.
Northern analyses of PAE1337 and PAE1338 (Figure 2E & 2F) showed identical
banding patterns that agree with the microarray results for PAE1338, and a major
band at approximately 3.3kb, consistent with co-transcription of a polycistronic
mRNA for PAE1336, PAE1337, PAE1338 and PAE1340 (annotated sizes 195, 2403, 450 and 240 nucleotides, respectively). Northern analyses of twenty other key loci showed very good agreement with the expression patterns indicated by microarray results (see following sections).

Many facultatively aerobic bacteria possess more than one respiratory oxygen reductase for growth under a range of oxygen tensions (47). The regulation of two cytochrome oxidases in the facultative aerobe *Pyrobaculum oguniense* (45) suggests an analogous specialization in the *Pyrobaculum* genus. *P. oguniense* up-regulated a SoxM-type cytochrome oxidase under aerobic conditions, and up-regulated a SoxB-type cytochrome oxidase under anaerobic, thiosulfate respiring conditions (45). This suggested that the SoxM complex is the terminal reductase for aerobic growth, and the SoxB complex functions as an O$_2$ scavenging reductase with a possible regulatory role.

The opposing expression pattern of the two homologous cytochrome oxidase complexes in *P. aerophilum* corroborates the expression patterns observed in *P. oguniense*, and provides support for the distinct functions proposed by Nunoura and colleagues (45). In this study, *P. aerophilum* SoxB complex genes were up-regulated under sub-oxic conditions, consistent with a role as an O$_2$ scavenger or sensor. The *P. aerophilum* SoxM complex genes were expressed in a strikingly opposite pattern, consistent with a primary role in O$_2$ respiration (Figure 2). Because *P. oguniense* can grow under atmospheric O$_2$ concentrations, whereas *P. aerophilum* tolerates only microaerobic conditions, Nunoura *et al.* speculated that the SoxB complex might be
the primary terminal reductase for microaerobic growth of *P. aerophilum*. However, the expression of the SoxM-type cytochrome oxidase genes in cultures amended with O₂ instead indicates that the SoxM complex in *P. aerophilum* is the terminal reductase for growth under microaerobic conditions.

Given this conclusion, the up-regulation of the SoxM complex after induction with nitrate was unexpected. Heme-copper cytochrome oxidases for O₂ reduction and heme-Fe containing nitric oxide reductases are structurally and evolutionarily related, and aerobic cytochrome oxidases from several bacteria have been shown to reduce nitric oxide (24, 26, 52, 64). However, a direct role for this SoxM-type cytochrome oxidase complex in nitrate respiration or denitrification would be novel. Whether expression of SoxM complex genes is characteristic of steady state growth on nitrate or is a transitional effect induced by the shift from growth on oxygen to growth on nitrate remains to be determined.

**Genes for nitrate respiration and denitrification in *Pyrobaculum***

*P. aerophilum* can reduce nitrate to dinitrogen during growth, and reductase activities for each individual substrate [nitrate (NO₃⁻), nitrite (NO₂⁻), nitric oxide (NO), and nitrous oxide (N₂O)] in this stepwise reaction have been demonstrated in association with *P. aerophilum* cell membranes (2, 16, 65). The *P. aerophilum* genome contains genes for a narG-type nitrate reductase (PAE3611-PAE3613), a cd1-type nitrite reductase (PAE3598 & PAE3600) and a heme O containing nitric oxide reductase (PAE3603) (1, 17, 23) (Figure 3A). However, the *P. aerophilum*
genome lacks a recognizable homolog of *nosZ*, which encodes nitrous oxide reductase in bacteria, and an alternative nitrous oxide reductase gene has not been identified (70).

Orthologs for the nitrate reductase, nitrite reductase and nitric oxide reductase were found in the genomes of both *P. calidifontis* and *P. arsenaticum* (*Table S3*). *P. calidifontis* also has a *nosZ* gene (Pcal_1928) of the *Wolinella succinogenes* type, in which a canonical multicopper oxidase domain is fused to a C-terminal cytochrome c containing domain (55). *P. calidifontis* produced N$_2$ when grown anaerobically on nitrate, indicating a capacity for denitrification consistent with the complement of denitrification genes in the *P. calidifontis* genome (6). In contrast, despite a large complement of genes for nitrate respiration and denitrification, *P. arsenaticum* was reported to be incapable of growth using nitrate (29). Our attempts to cultivate *P. arsenaticum* with nitrate or nitrite as electron acceptors were also unsuccessful (not shown). Further investigation will be required to determine whether *P. arsenaticum* is missing genes essential for nitrate respiration and denitrification, or if it requires specific culture conditions for growth with N-oxyanions.

**Limited modulation of the nitrate reductase operon**

Despite a > 250-fold difference in nitrate concentrations between cultures shifted to growth with nitrate versus those shifted to other electron acceptors (~10mM compared to < 0.04mM), microarray analysis showed relatively little variation in expression of the nitrate reductase alpha subunit, and only moderate (≤ 2-fold) up-
regulation after induction with nitrate (PAE3611; **Figure 3B**). Northern analysis showed that the nitrate reductase alpha subunit (PAE3611) mRNA was expressed under all conditions in the respiratory induction experiment (**Figure 3D**). Northern analysis also showed that PAE3611 is part of an approximately 7kb transcript, consistent with a polycistronic operon that includes an iron-sulfur subunit (PAE3612) a cytochrome B containing membrane subunit (PAE3613) (1), and a downstream ORF of unknown function (PAE3614) (**Figure 3A**).

The expression of the nitrate reductase operon under all respiratory conditions contrasts with that of other respiratory complexes identified in this study, and may reflect an adaptive preference for nitrate as a respiratory substrate. However, previous studies have shown large (10-fold) differences in nitrate reductase activity in *P. aerophilum* cultures grown to late log phase with nitrate compared to cultures grown with ferric iron, and a distinct redox optimum for growth on nitrate, suggesting that steady-state growth under controlled redox conditions may produce larger differences in nitrate reductase expression (20, 21).

**Co-expression of denitrification genes**

Denitrification genes including a subunit of the cd1-type nitrite reductase (PAE3600; cytochrome d1 subunit), a hypothetical protein (PAE3602), and the nitric oxide reductase (PAE3603) were up-regulated up to 15-fold by the end of the time course after induction with nitrate, compared to aerobic cultures (**Figure 3C**). A putative transcriptional regulator with homology to MarR-family transcriptional
regulators (PAE2679) and an adjacent hypothetical protein with a hemerythrin
binding domain (PAE2680) were also strongly (up to 17-fold) up-regulated late in the
time course after induction with nitrate (Figure 3C). The cytochrome c subunit of the
cd1-type nitrite reductase (PAE3598) and a uroporphyrin-III C-methyltransferase
possibly involved in cytochrome heme synthesis (PAE3601) showed similar
expression patterns, but were less strongly up-regulated.

Northern analysis of the cytochrome d1 subunit of the nitrite reductase
(PAE3600) showed a major band at 1.6-1.7kb and a minor band at 2.2kb, consistent
with a monocistronic transcript of PAE3600 and co-transcription together with the
uroporphyrin-III C-methyltransferase PAE3601, respectively (Figure 3E). The
cytochrome c subunit of the nitrite reductase (PAE3598) and the nitric oxide
reductase (PAE3603) showed similar expression patterns (Figure 3F & 3G).

Northern analysis of the putative MarR-like transcriptional regulator PAE2679
showed a primary band at approximately 1.1kb, consistent with co-transcription of
PAE2679-PAE2680 (predicted sizes 526nt and 543nt respectively), and a pattern of
expression that resembles those of the nitrite and nitric oxide reductases (Figure
3H). The results of the microarrays and Northern analyses are consistent in showing
strong up-regulation of these genes at the third time point during growth on nitrate,
and more moderate up-regulation during growth on arsenate (Figure 3B-H).

Conserved promoter motif for denitrification genes
Inspection of the PAE3598-PAE3603 locus revealed a conserved promoter motif that may be involved in regulating transcription of these genes. The motif is palindromic (Figure 3I), which is typical for targets of helix-turn-helix transcription factors, and occurs between the cytochrome c and d1 components of the nitrite reductase (PAE3598 and PAE3600), in the promoter region of the nitric oxide reductase (PAE3603), and in the promoter region of a hypothetical protein (PAE3602) (Figure 3A). The palindromic halves of the motif show particularly strong conservation between the two nitrite reductase genes, which are divergently transcribed in each of the three Pyrobaculum genomes where they occur (Figure 3J).

A candidate nitric oxide-responsive transcriptional regulator

The co-expression of genes for putative MarR-like and hemerythrin-binding proteins (PAE2679 & PAE2680) and genes for denitrification (Figure 3C, 3E-H) suggests that these may be regulated by common factors. Up-regulation of these genes when nitrite accumulated to ~1mM in the culture media (Figure 1B) suggests a possible role for nitrite in their regulation, with relatively low sensitivity to concentrations less than 1mM. An alternative possible effector is nitric oxide. Bacterial transcriptional regulators of denitrification generally fall in the Crp-Fnr family, are responsive to both oxygen and nitric oxide, and can be sensitive to nanomolar concentrations of nitric oxide (69). In contrast, bacterial MarR-family transcriptional regulators typically regulate antibiotic resistance, virulence, and oxidative stress (60). Hemerythrins are di-iron containing proteins that can reversibly
bind oxygen or nitric oxide (43). These annotations suggest that the PAE2679-
PAE2680 operon encodes a transcriptional regulator that controls either stress
responses, or denitrification genes in response to nitric oxide. An ortholog for only
the hemerythrin-like PAE2860 was found in the denitrifier P. calidifontis, and
orthologs for both genes were found in P. arsenaticum (Table 2).

Up-regulation of a multi-copper oxidase during nitrate respiration
A putative multi-copper oxidase (PAE1888) was among the genes most
specifically up-regulated (up to 5-fold) by growth on nitrate, and was highly ranked
by statistical analyses (Figure 4A & supplementary Table S2C). Northern analysis
with a probe for PAE1888 showed a band consistent with the annotated size of
1434nt, and increasing expression during the time course on nitrate (Figure 4B). The
putative 53kDa enzyme encoded by PAE1888 contains a predicted N-terminal twin-
arginine translocation signal, a predicted binuclear Cu-3 type center similar to those
in Cu-containing nitrite reductases, and a predicted mononuclear C-terminal Cu-2
type center (36). The twin-arginine signal suggests that this enzyme is exported after
assembly to the outer membrane. No orthologs of this gene were found in the other
Pyrobaculum or Thermoproteales genomes (Table 2). The closest BlastP hits (~40% identity) were annotated as bilirubin oxidases or multicopper oxidases, and were from
bacteria that are notable for either thermophily (Thermus, Aquifex), catabolism of
complex organic substrates (Polaromonas naphthalenivorans, candidatus
Desulfococcus oleovorans) or for ammonia oxidation and denitrification (Nitrosomonas, Nitrosospira).

Computational scans identified an over-represented sequence motif (Figure 4C) that may co-regulate PAE1888, a hypothetical protein with a thioredoxin-like domain (PAE2750), and two downstream genes [a putative cytochrome c-type biogenesis protein (PAE2751), and an intergenic region (PAE.i1473)] that showed similar expression patterns (Figure 4A). The motif is located immediately upstream from the predicted TATA box for both PAE1888 and PAE2750, and is conserved in the promoter regions of genes orthologous to PAE2750 in the genomes P. calidifontis and P. arsenaticum (Figure 4D-E). A putative cytoplasmic siroheme type ferrodoxin-nitrite reductase (PAE2577), typically involved in assimilatory or detoxifying reduction of nitrite to ammonium (41, 54), also showed an expression pattern similar to PAE1888 (Figure 4A), but lacked this promoter motif.

Enzymes in the multicopper oxidase superfamily are diverse, and typically use O₂ as an electron acceptor, but also include bacterial copper-containing nitrite reductases encoded by nirK genes and nitrous oxide reductases encoded by nosZ genes (70). Based on the specific up-regulation of PAE1888 during nitrate respiration and the putative domain structure of the encoded protein, we propose that PAE1888 encodes a novel nitrite or nitrous oxide reductase. The similarity between PAE1888 and genes in Nitrosomonas species, which also reduce nitrous oxide and lack identified nitrous oxide reductase genes, provides potential support for a role in nitrous oxide reduction (56, 70). The co-regulated, putative thioredoxin (PAE2750)
and cytochrome c biosynthesis (PAE2751) genes may function in biosynthesis of co-factors or electron donors for this enzyme.

**A predicted arsenate respiratory reductase**

Bacterial arsenate respiratory reductases (ARR) are members of the molybdopterin oxidoreductase superfamily (50). Degenerate PCR primers that reliably amplify bacterial ARR genes do not detect them in *P. aerophilum* or *P. arsenaticum*, the only archaeal arsenate respirers identified to date (29, 35). We identified nine putative molybdopterin oxidoreductase genes in the *P. aerophilum* genome, four of which (PAE1265, PAE2002, PAE2839-PAE2840, and PAE2859) show limited similarity to functionally characterized genes. We predicted that one of these encodes an arsenate respiratory reductase similar to those found in bacteria.

BlastP searches of the *P. aerophilum* genome using the Mo-containing subunit of arsenate respiratory reductases from diverse bacteria (*Shewanella ANA-3, Bacillus arseniciselenatis, Chrysiogenes arsenatis, Sulfurospirillum barnesii*) as query sequences all returned PAE2859 as the strongest match. However, the PAE2859 operon was up-regulated under anoxic conditions only in the absence of nitrate, arsenate, or ferric iron (**Figure 5B & 5E**). In contrast, a group of genes encoding a putative 132kD molybdopterin oxidoreductase (PAE1265), a putative 23kD iron-sulfur-protein (PAE1263), and a possible 40.7kD membrane-anchoring subunit (PAE1264) were strongly up-regulated (up to 8-fold) in cultures induced with arsenate (**Figure 5A**). These cultures reduced approximately 1mM arsenate to
arsenite (Figure 1C & 1D). Northern analysis with probes for the Mo-containing subunit PAE1265 and the membrane subunit PAE1264 showed transcripts of approximately 5.2kb and identical patterns of up-regulation, consistent with transcription of an operon containing all three of these co-regulated genes (Figure 5C & 5D). Northern analysis also showed increasing expression in cultures induced with ferric iron, and up-regulation at the last time point after induction with nitrate.

The predicted active site domain of PAE1265 is conserved in bacterial tetrathionate reductases and arsenate respiratory reductases (36). Both the iron-sulfur subunit and the molybdopterin-binding subunit contain predicted N-terminal twin-arginine translocation signals, suggesting that the gene products are translocated to the outer membrane after translation. Although there are orthologs of PAE1263 in several Pyrobaculum species, orthologs for all three components of the PAE1263-PAE1265 operon were found only in P. arsenaticum, the only other archaeon known to grow robustly on arsenate (Table 2) (29). Based on these data we propose that the PAE1263-1264-1265 operon encodes a membrane-bound arsenate respiratory reductase.

Few other genes were specifically up-regulated by arsenate (Figure 5A). These included a putative arsR homolog (PAE1592) which may be involved in regulation of arsenite-responsive genes, and a putative argininosuccinate synthase (PAE2884). Genes annotated as an arsenical pump-driving ATPase (PAE1427) and an arsenite permease (PAE2457) were not up-regulated in arsenate respiring P.
Peaerophilum cultures, even when arsenite accumulated to 1mM (supplementary Table S2B).

Recovery of a variant *P. calidifontis* strain after growth on arsenate

The *P. calidifontis* genome (Refseq NC_009073) contains orthologs for both PAE1263 and PAE1264 (Pcal_1599 & Pcal_1600, Table 2). An open reading frame adjacent to these orthologs (Pcal_1601a; currently annotated as a pseudogene) with strong primary sequence similarity to PAE1265, is truncated after 810 amino acids by a stop codon (TAG) at *P. calidifontis* chromosomal position 1489855 (Figure 5F).

Fourteen of fifteen sequencing reads used to assemble the *P. calidifontis* genome in this region unambiguously show a TAG stop at this position, indicating that this was not a sequencing error. However, one read (#1617717348) shows a GAG at this position (Glu; the consensus amino acid at the corresponding site in *P. aerophilum* and *P. arsenaticum*), which would produce a 1170aa ORF with 83% identity to the 1173aa PAE1265 protein (Pcal_1601b; Figure 5F). The stop codon at position 1489855 occurs approximately 500nt upstream of what would be a molybdopterin-binding domain in the full-length Pcal_1601, if translated. Although the *P. calidifontis* culture used for whole-genome sequencing was grown from a single colony, the anomalous sequencing record showing a GAG Glu codon at position 1489855 suggests that this culture may have contained a sub-population of cells with a variant Pcal_1601 locus encoding a full length, functional ortholog of PAE1265.
Cozen et al. Analysis of respiratory pathways in *P. aerophilum*

We re-sequenced the Pcal_1601 locus in our laboratory stocks of *P. calidifontis*, which were directly derived from cultures used for whole-genome sequencing. PCR amplification of genomic DNA confirmed the presence of the TAG stop codon, with no evidence for alternative bases at position 1489855. We tested the growth of these *P. calidifontis* cultures with arsenate to determine whether a full-length ortholog of the PAE1265 protein is required for arsenate respiration. *P. calidifontis* grew very slowly and to low densities during the first two passages with arsenate as the only electron acceptor. No growth was observed when arsenate was omitted. Growth during subsequent passages with arsenate was more rapid, and produced higher cell densities (not shown). After the sixth passage of *P. calidifontis* on arsenate, we extracted genomic DNA and sequenced the PCR-amplified Pcal_1601 locus again. This revealed a GAG Glu codon at position 1489855. No other differences relative to the reference sequence were observed in the 1kb PCR amplicon, suggesting that cultivation of *P. calidifontis* with arsenate as the only electron acceptor selected for a strain with a full length ortholog of the proposed arsenate respiratory reductase encoded by PAE1265. We have isolated clonal cultures of each strain from single colonies, and confirmed that only the strain with the full-length Pcal_1601 grows on arsenate. We are currently investigating the variant strains to describe these in more detail.

Up-regulation of a putative thiosulfate or polysulfide reductase during anoxia
Components of the putative membrane molybdopterin oxidoreductase encoded by PAE2859, PAE2860, and PAE2861, were specifically up-regulated in *P. aerophilum* following the onset of anoxia in cultures lacking an alternative respiratory electron acceptor. These included the second and third time points in cultures that received no added terminal electron acceptor, and the last time point in the cultures that received additions of O$_2$ (Figure 5B & 5E). The putative 91kD Mo-containing subunit encoded by PAE2859 contains a twin-arginine translocation signal and a MopB-thiosulfate-R-like domain that is conserved in thiosulfate, polysulfide and sulfur reductases. The putative 31kD membrane subunit encoded by PAE2861 contains a domain that is conserved in membrane subunits of polysulfide reductases. Orthologs of the PAE2859-2860-2861 complex were found in the *P. arsenaticum* and *P. calidifontis* genomes, whereas *P. islandicum* and *T. neutrophilus* were missing the membrane subunit (Table 2). Northern analysis of the molybdopterin containing subunit PAE2859 showed a single band of approximately 3.8kb, a size consistent with the co-transcription of these three genes, and possibly a small 3’ untranslated region (PAE.i1535; Figure 5E).

The domain structure and expression patterns of this operon suggest that it is a respiratory reductase of an electron acceptor such as thiosulfate or polysulfide that is not used by *P. aerophilum* in the presence of O$_2$, nitrate, arsenate or ferric iron. *P. aerophilum* is poisoned by the presence of elemental sulfur, and grows weakly using thiosulfate as an electron acceptor (29, 65). Growth on polysulfide has not been tested. The basal growth media contained 11mM sulfate and no other sulfur.
compounds aside from those present as trace constituents of the yeast extract used as
the carbon and energy source for growth. Physiological studies of *Pyrobaculum*
species have shown that their growth with different terminal electron acceptors is
optimized at different reducing potentials (21). This observation suggests that
reducing potential may directly modulate the regulation of respiratory reductases in
*Pyrobaculum*, and might explain the up-regulation of the PAE2859-PAE2861
molybdopterin oxidoreductase in the absence of its cognate substrate.

**Limited up-regulation during iron respiration**

Cultures shifted to ferric-citrate reduced approximately 3.5mM Fe(III) to
Fe(II) during the time course (**Figure 1E**). A large number of genes were moderately
up-regulated in these cultures (**Figure 6A**). However, we were unable to identify
genes that exhibited strong and specific up-regulation in response to iron, and positive
transcriptional responses to iron (**Figure 6A**) were generally weaker than the negative
responses (**Figure 7A**). A small cluster of weakly annotated genes (PAE.i1354,
PAE2526, PAE2527) showed a strong initial response to iron that declined over the
time course, and opposite trends in cultures amended with O₂ or no electron acceptor.
(**Figure 6A**). Other genes up-regulated by iron exhibited increasing trends over the
time course that were also apparent under other respiratory conditions. These
included a putative 40.9kD mono-heme c-type cytochrome (PAE0090), a putative
thiosulfate sulfurtransferase (PAE2582), and a number of loci adjacent to each of
these (**Figure 6A**). Genes including a succinate dehydrogenase complex (PAE0716-
PAE0725) (Figure 6B), exhibited positive responses that appeared to be more specific, but were also relatively modest.

Previous studies have shown that respiratory ferric reductase activity is regulated, and primarily associated with the cytoplasmic cell fraction in *P. aerophilum* and *P. islandicum* (20, 21). Specialized membrane proteins and c-type cytochromes are important in bacterial iron respiration, but were found to be lacking in iron-respiring *Pyrobaculum* cultures. We were similarly unable to identify a specific oxidoreductase for ferric iron respiration. It is possible that genes for respiratory iron reduction, like the genes in the nitrate reductase operon, were expressed under all conditions in our study with only modest differences between respiratory treatments. Alternatively, the lack of apparent specialization for ferric iron respiration may support the theory that respiratory iron reduction evolved early as a primitive process mediated by a variety of oxidoreductases and electron carriers rather than a single specific complex (47).

**Transcriptional repression by iron**

A large proportion of genes that were differentially expressed between respiratory conditions at a statistically significant level were down-regulated in cultures amended with ferric iron. A number of these have annotations suggesting roles in oxidative stress response or iron uptake and homeostasis. These included genes for superoxide dismutase (PAE0274), Mn-catalase (PAE2696), a putative antioxidant protein (PAE0877), a putative ABC transporter for Fe(III) (PAE2699),
and a gene annotated as a bacterioferritin co-migratory protein (PAE2732) (Figure 7A-D). Genes for putative GTP cyclohydrolases (PAE1588-PAE1589, PAE2984-PAE2986), which are involved in folate and flavin biosynthesis, were also repressed in cultures amended with iron (Figure 7A, 7F).

The Mn-catalase (PAE2696) was the top-ranked gene in statistical analyses for differential expression between respiratory conditions (supplementary Table S2C), and was very strongly down-regulated by iron (up to 40-fold compared to cultures induced with O2) (Figure 7A, 7C). Examination of the catalase gene and an adjacent cluster of iron-repressed genes (PAE2697-PAE2702) revealed a strikingly conserved palindromic promoter motif (Figure 7H). Three nearly identical versions of this motif are located directly over predicted basal regulatory elements (BRE) for PAE2696, PAE2699 and PAE2700 (Figure 7H-J). A genome-wide search for similar motifs revealed a perfect match in the promoter region for a predicted metal transporter that was up-regulated at the end of the time course with nitrate and down-regulated by iron (PAE0600; Figure 3C & Figure 7E). No orthologs for PAE2697, PAE2699 or PAE2700 were found in the other Pyrobaculum genomes (Table 2). However, the promoter motif is largely conserved in regions upstream of the P. arsenaticum and P. calidifontis orthologs for PAE2696 (Figure 7J) and PAE0600 (not shown).

Northern analysis of the putative Fe(III) transporter gene PAE2699 showed a band at approximately 0.8kb, somewhat larger than the annotated size of 624nt, and a band at 1.3kb consistent with co-transcription of the downstream putative ABC
Cozen et al. Analysis of respiratory pathways in *P. aerophilum*

transporter gene (PAE2697) (Figure 7D). Northern analysis of the predicted metal transporter PAE0600 (Figure 7E) showed a very similar expression pattern, with strong up-regulation at the end of the induction with nitrate, when denitrification genes were induced, and down-regulation after induction with iron. Northern analysis also confirmed microarray data indicating that a ferritin-like gene (PAE2701) located between the PAE2696-PAE2700 cluster and another iron-repressed gene, PAE2702, was not down-regulated by iron (not shown).

Evidence for an archaean Ferric Uptake Regulator system

Iron concentrations in the *P. aerophilum* cultures amended with ferric-citrate exceeded the concentration in the basal growth media by over 1000-fold (~9mM compared to ~7µM). Iron is a co-factor for numerous enzymes and an essential nutrient for nearly all microorganisms. However, the co-occurrence of high intracellular Fe(II) and O_2 can generate reactive oxygen species, requiring organisms to tightly regulate uptake and storage of iron in the cell (8). In bacteria, many genes involved in iron homeostasis and oxidative stress are controlled by the ferric uptake regulator (Fur) protein, which binds to DNA and acts as a transcriptional repressor when complexed to Fe(II) (8, 30, 62).

The primary targets of transcriptional repression by bacterial Fur are remarkably similar to those found in this study to be repressed by iron in *P. aerophilum*: genes for oxidative stress response and for acquisition, transport and storage of iron. GTP cyclohydrolases, which were repressed by iron in *P. aerophilum,*
are involved in reductive iron assimilation in bacteria and may also be controlled by Fur (14, 68). Bacterial Fur systems are mechanistically complex, with post-transcriptional regulatory roles for small regulatory RNAs and for RNA chaperones that are similar to eukaryotic small nuclear ribonucleoproteins (Sm-RNP) and archaeal Sm-RNP-like proteins (37, 38, 40). Whether similar Fur-regulated systems operate in archaea has not been directly investigated, but proteins with Fur-like domains can be found in over a dozen archaeal genomes including all those in the *Pyrobaculum* clade.

The *E. coli* Fur binding site consensus sequence has been alternatively modeled as a ‘FFnR’ repeat of ‘GATAAT’ (where ‘F’ stands for the forward strand, ‘R’ for reverse, and ‘n’ stands for any nucleotide) (19), and a ‘FnR’ repeat of ‘AATGATAAT, which shows good agreement with the consensus sites of *Bacillus subtilis* and *P. aerogenosa* (12). The motif identified in the PAE2696-PAE2700 locus shows some consistency both models. The putative *P. aerophilum* motif consists of the sequence ‘ATTAAC’ repeating in the pattern ‘FnnnR’, which can be viewed in reverse as, ‘GTTAAT’ repeated ‘RnnnF’ (**Figure Fe_down H**). Whether this motif is the target of the *P. aerophilum* Fur homolog remains to be determined. Fur is auto-regulatory in many, but not all bacteria, repressing its own promoter in the presence of iron (15, 18). However, no match for the PAE2696-PAE2700 promoter motif was found in the promoter region of the *P. aerophilum* Fur homolog (PAE2309), and the transcript was not strongly down-regulated after induction with iron (supplementary
Table S2B, Figure 7G). Nevertheless, similarities with the bacterial Fur system suggest a potentially homologous regulation system in archaea.

A transcriptional map of crenarchaeal respiratory versatility

The results of this study show that numerous predicted respiratory genes and many others are modulated response to different respiratory conditions in the remarkably flexible *P. aerophilum*. These provide a map of gene regulation for different respiratory pathways in crenarchaea, and a number of functional and regulatory predictions for further study. Two cytochrome oxidases were expressed in opposing patterns, suggesting that a SoxB-type complex is for oxygen scavenging and a SoxM-type complex is for aerobic growth (Figure 8A). The nitrate reductase operon was expressed under all conditions examined, whereas genes for nitrite reductase and nitric oxide reductase were strongly up-regulated when nitrite accumulated to ~1mM. The nitrite and nitric oxide reductases share a conserved promoter motif, and were co-expressed with a putative hemerythrin-containing transcriptional regulator that may control nitric oxide responsive genes (Figure 8B). Expression patterns suggest that a multicopper oxidase serves a specific role in denitrification, possibly as a nitrous oxide reductase. A molybdopterin oxidoreductases was up-regulated in arsenate respiring *P. aerophilum* cultures, and is proposed to encode a crenarchaeal arsenate respiratory reductase (Figure 8C). A full-length ortholog of this complex appears to be required for growth of *P. calidifontis* on arsenate. *P. aerophilum* expressed a second molybdopterin oxidoreductase operon,
which may encode a thiosulfate or polysulfide reductase, in anoxic cultures lacking
an alternative electron acceptor (Figure 8D). Although a number of loci were
moderately up-regulated during ferric iron respiration, we were unable to identify
specific genes for this process.

_P. aerophilum_ cultures shifted to growth with ~9mM iron exhibited
transcriptional repression at many loci with annotations suggesting roles in iron
homeostasis and oxidative stress (Figure 8E). These responses are similar to those
controlled by the ferric uptake regulator in bacteria, and suggest either convergent
evolution, an ancient global regulatory system, or some combination of the two.
Evidence for an ancient Fur system would have interesting evolutionary and paleo-
geochemical implications. The bacterial Fur-regulatory system evolved to cope with
iron scarcity caused by the low solubility of Fe(III)-oxides in an oxidizing
environment, and the toxicity posed by intracellular Fe(II) and free O_2_. The presumed
divergence of the Bacteria and Archaea occurred well before the global oxidation of
the atmosphere and oceans 2.3-2.7 Gya, when soluble ferrous iron was plentiful and
free O_2_ was scarce. However, based on the phylogeny of cytochrome oxidases, it has
been proposed that the capacity for aerobic growth evolved in the last universal
ancestor of Bacteria and Archaea prior to the broader oxidation of the biosphere,
within localized environments enriched in O_2_ (10, 11). A conserved, vertically
transmitted Fur system in the Archaea might provide additional support for
evolutionary adaptation to oxidizing environments in the last universal ancestor, and
therefore an early emergence of free O_2_.

40
Acknowledgements

Support for this work was provided in part by a NSF Graduate Research Fellowship to A.E.C. and an Alfred P. Sloan Research Fellowship to T.M.L. The authors thank the Department of Energy’s Community Sequencing Program and the other members of the *Pyrobaculum* Sequencing Consortium (Chad Saltikov, Christopher House, Sorel Fitz-Gibbon) for making comparative genomics possible in *Pyrobaculum*. We thank Jeffrey H. Miller for providing funding for oligonucleotide primers and other materials used in construction of the DNA microarrays, Jeffrey Savas, Lily Shiu, and Sofie Salama for help constructing the arrays, and Rob Franks and Carolina Reyes for assistance with analytical chemistry. Finally, we thank Chad Saltikov and Jonathan Trent for helpful discussions of the manuscript.
References


Cozen et al_________________

Analysis of respiratory pathways in *P. aerophilum*


40. Moller, T., T. Franch, P. Hojrup, D. R. Keene, H. P. Bachinger, R. G.
41. Moreno-Vivian, C., P. Cabello, M. Martinez-Luque, R. Blasco, and F.
Castillo. 1999. Prokaryotic nitrate reduction: molecular properties and
181:6573-84.
42. Niederberger, T. D., R. S. Ronimus, and H. W. Morgan. 2007. The
microbial ecology of a high-temperature near-neutral spring situated in
the binuclear iron site of hemerythrin: spectroscopy and reactivity.
44. Nunoura, T., Y. Sako, T. Wakagi, and A. Uchida. 2005. Cytochrome aa3 in
facultatively aerobic and hyperthermophilic archaean Pyrobaculum
45. Nunoura, T., Y. Sako, T. Wakagi, and A. Uchida. 2003. Regulation of the
aerobic respiratory chain in the facultatively aerobic and hyperthermophilic
in Oxalate Extracts of Sediment. Soil Science Society of America Journal
51:938-941.
47. Richardson, D. J. 2000. Bacterial respiration: a flexible process for a
nov., a novel facultatively aerobic and hyperthermophilic archaean growing at
N.Y.
unprecedented nos gene cluster of Wolinella succinogenes encodes a novel
respiratory electron transfer pathway to cytochrome c nitrous oxide reductase.


Figure legends

Figure 1. Growth and substrate usage after shifting to growth with different terminal electron acceptors. Error bars in each panel represent standard deviations for three replicate experiments. Concentrations of NO$_3^-$ and [NO$_3^-$+NO$_2^-$] (A), and concentrations of NO$_2^-$ (B) in cultures shifted to growth on NO$_3^-$; Concentrations of As(V) (C) and As(III) (D) in cultures shifted to growth on As(V); Concentrations of Fe(III) and Fe(II) in cultures shifted to growth on Fe(III)-citrate (E). Optical density after shifting to different respiratory substrates, with results for Fe(III) amended cultures omitted (see text) (F).

Figure 2. Genomic organization (A) and transcriptional expression (B-F) of aerobic SoxM type and SoxB-type cytochrome oxidase genes in *P. aerophilum*. Red arrows in (A) indicate transcripts detected by northern analysis (panels D-F). Individual genes are shown in shaded boxes, with arrow marks indicating the predicted direction of transcription. Black hatch marks at the bottom of (A) indicate regions of primary sequence similarity in multiple genome alignments with other *Pyrobaculum* species, and show a contiguous region of similarity in the aerobe *P. calidifontis*. Hierarchically clustered microarray gene expression profiles for SoxB cytochrome oxidase genes (B) and for SoxM cytochrome oxidase genes (C) during time courses of growth on different terminal electron acceptors. Positive (red) and negative (green) changes in gene expression are shown on a log 2 scale, with a scale bar shown at the
top. Expression values for each time course represent the average log 2 ratio of three replicate respiratory induction experiments. Blue dendrograms show clustering relationships based on the Pearson correlation. Northern analysis of SoxB type cytochrome oxidase (D) and the SoxM type cytochrome oxidase genes (E-F), with lanes corresponding to conditions indicated above the heatmaps. Methylene blue stain of total RNA showing 23s and 16s ribosomal RNA bands (G) on a representative blot used for Northern analysis.

Figure 3. Genomic organization (A), transcriptional expression (B-H), and promoter region schematics (I-J) of genes for nitrate, nitrite, and nitric oxide reductases in P. aerophilum. Red arrows in (A) indicate transcripts detected by Northern analysis (panels D-H). Red bars in (A) indicate the location of the conserved promoter motif shown in (I-J). Black hatch marks at the bottom of (A) show contiguous regions of genomic similarity in the denitrifier P. calidifontis, and in P. arsenaticum. Hierarchically clustered microarray gene expression profiles for nitrate reductase (B), and for nitrite and nitric oxide reductases (C) during time courses of growth on different terminal electron acceptors. Northern analysis of the nitrate reductase large subunit (D), nitrite reductase subunits (E-F), nitric oxide reductase (G), and a co-expressed MarR-like transcriptional regulator (H). Conserved palindromic promoter motif found in the promoter regions of the nitrite reductase (PAE3598 & PAE3600), the nitric oxide reductase (PAE3603), and a hypothetical protein (PAE3602) (I), with nucleotide frequencies indicated by the height of motif logos, and bit scores shown in
Cozen et al. Analysis of respiratory pathways in P. aerophilum

the y-axis. Promoter region schematic of the cytochrome cd1 nitrite reductase showing the position of the conserved promoter motif (underlined by red bar) relative to the TATA box (green box), and conservation of the palindromic motif in P. aerophilum, P. arsenaticum, and P. calidifontis (J). Sequence identity in the multiple genome alignment is indicated by black dots.

**Figure 4.** Expression (A-B), and promoter region schematics (C-D) of a multicopper oxidase gene and other genes up-regulated during growth on nitrate. Microarray gene expression profiles for the multicopper oxidase and co-clustered genes (A). Northern analysis of the multicopper oxidase gene (B), Promoter motif sequence and repeating pattern (C) for a conserved motif in the promoter regions of the multicopper oxidase (PAE1888; D) and a co-expressed thioredoxin-like gene (PAE2750; E). Arrows inside red bars indicate the orientation of each motif, and green boxes show predicted TATA boxes. Conservation of the promoter motif in upstream of the thioredoxin-like gene PAE2750 and its orthologs in P. aerophilum, P. arsenaticum, and P. calidifontis (E).

**Figure 5.** Contrasting transcriptional regulation of two putative molybdopterin oxidoreductases (A-B). Microarray gene expression profiles for a molybdopterin oxidoreductase complex (PAE1263, PAE1264, PAE1265) and other genes up-regulated during growth on arsenate (A), and a molybdopterin oxidoreductase complex (PAE2859, PAE2860, PAE2861) up-regulated during anoxia (B). Northern
analysis of PAE1265 (C), PAE1264 (D), and PAE2859 (E). *P. calidifontis* genome alignment showing showing truncated (Pcal_1601a) and full-length (Pcal_1601b) orthologs for PAE1265 identified before, and after selection for growth of *P. calidifontis* on arsenate, respectively (F). Conserved protein domains predicted by Pfam are shown in magenta. Black hatch marks show primary sequence similarity to *P. aerophilum* and *P. arsenaticum* in multiple genome alignments.

**Figure 6.** Microarray gene expression profiles for genes up-regulated during growth on ferric citrate (A). Northern analysis of succinate dehydrogenase large subunit (sdhA; PAE0716) showing modest up-regulation during growth on ferric iron (B).

**Figure 7.** Microarray gene expression profiles (A), and Northern analysis for genes down-regulated during growth on ferric citrate (B-F). Northern analysis of the putative *P. aerophilum* ferric uptake regulator homolog (G), which was not down-regulated by iron. Palindromic promoter motif logo (H) and location (red bars; I) in the promoter regions of genes for Mn-catalase (PAE2696), a putative iron (III) transporter (PAE2699), and a putative ABC transporter (PAE2700). Promoter region schematic for the Mn-catalase gene (J) showing the location of the palindromic motif (red underline) relative to the predicted BRE (blue box) and TATA box (green box), and conservation of the motif in *P. arsenaticum* and *P. calidifontis*. Black dots in the multiple genome alignment indicate nucleotide identities, and dashes indicate gaps.
Figure 8. Summary schematic showing proposed roles for \textit{P. aerophilum} genes in aerobic respiration and O$_2$ scavenging (A), denitrification (B), arsenate respiration (C), polysulfide or thiosulfate respiration (D), and iron-mediated transcriptional repression (E).