UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Investigation of the Fur regulon in G. sulfurreducens

Permalink
https://escholarship.org/uc/item/2455c099

Author
Shieu, Wendy Aiween

Publication Date
2010

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Investigation of the Fur regulon in *G. sulfurreducens*

A thesis submitted in partial satisfaction of the requirements for the degree
Master of Science

in

Bioengineering

by

Wendy Aiween Shieu

Committee in charge:

        Bernhard Palsson, Chair
        Xiaohua Huang
        Kun Zhang
        Karsten Zengler

2010

The Thesis of Wendy Aiween Shieu is approved, and it is acceptable in quality

and form for publication on microfilm and electronically:

_____

_____

_____

_____

Chair

University of California, San Diego

2010

DEDICATION

First and foremost, I would like to dedicate my thesis to my family, especially my parents, for raising me and supporting me. Their support and love through the years has enabled me to be successful throughout my entire life.

I would like to thank Dr. Bernhard Palsson for allowing me the opportunity to research in the Systems Biology Research Group at UCSD.

I am incredibly indebted to Karsten Zengler and Yu Qiu for all of their support and guidance over the past few years. If it were not for them, it would not have been possible to do the work that I was able to accomplish for this project. I have greatly enjoyed working with and slowly getting to know them. I would like to thank them so much for being patient with me and sharing so much knowledge with me.

I would also like to thank Nathan Lewis for being a great TA back in the day (a little over two years ago?) and making me aware of the research opportunities at the System Biology Research Group.

I am also grateful toward various people throughout the lab—including Kenyon, Richard, Kat, Nicole, Elisa, Pep, Mallory, Vasily, BK, Young-Seoub, and more!—for all of their support, advice, friendship, and shared laughter.

I would also like to thank all of my friends, who in one way or another, made this project and this past year easier for me. I'd like to thank Bassem, Vu, Timothy, Andrew, and Jason for being homework and study group friends. I wouldn't have been able to spend so much time researching if it hadn't been for their help on homework and projects.

I'd like to further thank all of the people who came up with witty Fur-related puns that I was seriously contemplating using for the title of my thesis. These gems included: InFURing transcriptional regulation, The Fast and the FURious, FURocious regulation in *Geobacter*, FURreal: The *Geobacter* story, Letting the Fur Fly: A No Holds Barred Investigation into the World of *Geobacter*, and the FURbidden Kingdom. The majority of these potential titles were provided by Timothy and Lee so thank you two for your wittiness. I would further like to thank Lee for providing me with the food necessary for good research and thesis writing and for generally being there for me for the past few months.

Last, but not least, I'd like to thank the eclectic bunch of musical artists that prevented me from going insane by providing good thesis writing background music, including: Silversun Pickups, Vienna Teng, DeVotchka, Josh Vietti, Bond, Placebo, Air, Belle and Sebastian, and Radiohead.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT OF THE THESIS

Investigation of the Fur regulon in *G. sulfurreducens*

by

Wendy Aiween Shieu

Master of Science in Bioengineering

University of California, San Diego, 2010

Professor Bernhard Palsson, Chair

Iron is an essential element in bacteria as a result of being commonly used as a cofactor. In *G. sulfurreducens,* the assimilatory need for iron is especially high as a result of up to 111 coding sequences for c-type cytochromes, many of which have multiple heme-binding sites. It may also be necessary for *G. sulfurreducens* to carefully regulate its iron uptake and

transport as a result of its environmental iron concentrations constantly being in flux. Thus, the direct transcriptional role of a key iron-dependent regulator known as the ferric uptake regulator (Fur) was investigated in *G. sulfurreducens* through expression profiling of a variety of iron conditions via microarray analysis in combination with binding sites elucidated via chromatin immunoprecipitation coupled with microarrays (ChIP-chip). 224 genes were determined according to the ChIP-chip binding profile to be directly regulated by Fur, and 22 of these genes overlapped with differential expression observed in expression profiling. From this study, Fur was determined to be the primary regulator of genes involved in iron uptake and transport as well as a key regulatory of several other genes that play key roles in *G. sulfurreducens.* The implications of this study also provide a commentary on the methods generally used to determine the genome-wide role of transcriptional regulations as well as the complexity of transcriptional regulation.

1. BACKGROUND

1.1 *Geobacteracae*

The *Geobacteraceae* is a family of metal reducing δ-proteobacterium that is an important component of many subsurface environments [1]. The first *Geobacter* species isolated, *Geobacter metallireducens,* was initially recovered from freshwater sediments of the Potomac River. It was soon realized that *G. metallireducens* had the ability to couple the oxidation of organic compounds with the reduction of extracellular ferric iron [1]. The ability to transfer electrons to extracellular acceptors is a result of the metabolism found in the *Geobacteraceae.* Unlike many other organisms, the *Geobacteraceae* typically do not use sugar as their main energy source, but instead preferentially oxidize acetate and other carbon sources [9]. Electrons generated from anaerobic oxidation are transferred through direct contact, rather than through electron shuttles, to extracellular electron acceptors [9, 14]. Since the initial discovery of *G. metallireducens,* several other members of *Geobacteracae*, such as *G. sulfurreducens*, have been isolated from other sediments. Analysis of these subsurface environments have suggested that members of the *Geobacter*acae are the most predominant metal reducing microorganism in sediments where Fe^{3+} reduction is a critical process [2-4].

In addition to playing a significant part in subsurface environments, the *Geobacteraceae's* electron transferring capabilities have the potential to be used for a variety of biotechnological purposes, including bioremediation and

the conversion of organic matter into carbon-neutral energy. Members of the *Geobacteraceae* have shown much promise in bioremediation through the natural biodegradation of hydrocarbon contaminants and the immobilization of uranium in polluted aquifiers via the reduction of soluble U(VI) to insoluble U(IV) [5, 6]. It has also been demonstrated the *Geobacteraceae* can use anodes—rather than metals—as terminal electron acceptors, thereby making it possible to use *Geobacter* to harvest electricity from organic compounds in wastewaters or sediments (Figure 1.1) [7]. Recently, a strain of *G. sulfurreducens,* designated as KN400, was selected for an enhanced capacity for current production. Its current and power densities were 7.6 A/m^2 and 3.9 W/m^2 respectively, which were eight times higher than the densities of the wild-type strain [8].

Cathode reaction:
$2O_2 + 8H^+ + +8e^- \rightarrow 4H_2O$

$4H_2O$ $2O_2$

Cathode

Water

Sediment

$8H^+$ $8e^-$

~10cm

Anode

Anode reaction:
$C_2H_4O_2 + 2H_2O \rightarrow 2CO_2 + 8H^+ + 8e^-$

$\uparrow e^-$

Fermentation Organic matter

$2CO_2$ Acetate
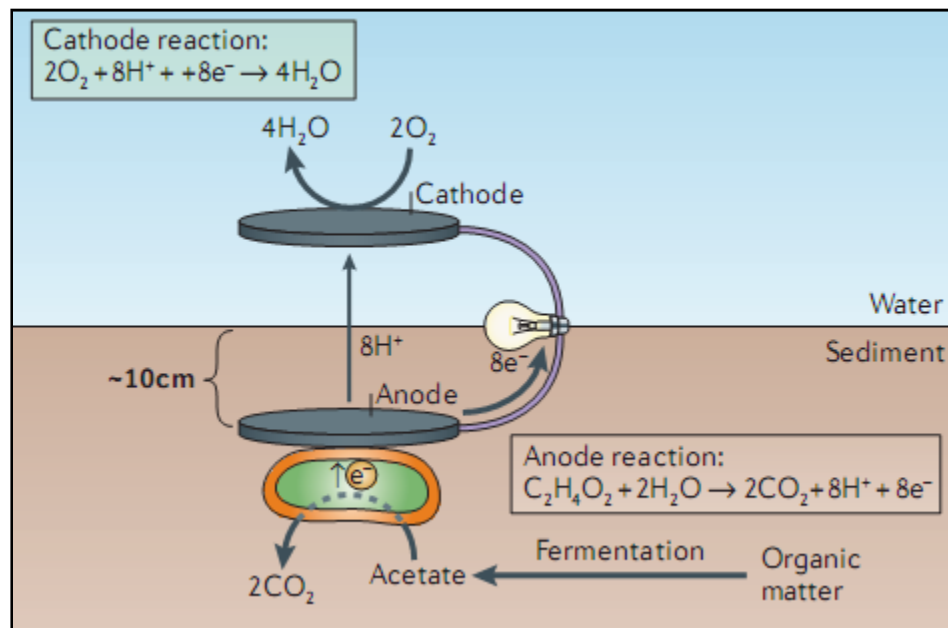
Figure 1.1: A schematic of how the *Geobacteracae* could potentially be used in a sediment microbial fuel cell [7].

G. sulfurreducens, a species of *Geobacter* originally isolated from petroleum-contaminated sediments, is often used as a model for *Geobacter* since it has a relatively well-defined genetic system [9, 10]. The publication of the complete genome sequence of *G. sulfurreducens* elucidated many interesting physiological characteristics, including an unprecedented number of genes encoding putative c-type cytochromes [9]. C-type cytochromes are monomeric proteins that contain at least one covalently bound heme group. Their heme groups allow them to perform a number of essential functions, including binding to O_2 as well as transferring and accumulating electrons [34].. In the *G. sulfurreducens* genome, 111 coding sequences contained one or more matches to the c-type cytochrome motif that binds to heme groups. Out of these 111 putative c-type cytochromes, 73 contained two or more heme groups, including one that contained 27 [9]. In comparison, *Escherichia coli's* genome only encodes for seven c-type cytochromes [15]. Furthermore, a proteomics study demonstrated that 91 out of 111 predicted c-type cytochromes (82%) found in *G. sulfurreducens* are produced to a detectable level under one or more culture conditions [12], thereby validating that the majority of the putative c-type cytochromes are being produced. It is hypothesized that the large number of cytochromes and heme groups help facilitate rapid electron transfer to extracellular electron acceptors by forming a continuous "electric wire" [35].

In addition to the large number of c-type cytochromes, there also appeared to be a large number of hypothetical genes that encoded for putative proteins of unknown function [9]. A better understanding of transcriptional regulation of the various c-type cytochromes and hypothetical proteins with unknown functions could give a better understanding of the physiology of *G. sulfurreducens* and other members of the *Geobacteraceae*, which could lead to a better understanding of how to best optimize their electron transferring abilities for biotechnological purposes.

1.2 Transcription regulatory network

As a result of advances in technology, sequencing of genomes of various microbial organisms has become easier and cheaper and thus more common in recent years. Despite the vast amount of sequence information available, there lacks proper annotation of genomic organizational structure and dynamics, which is necessary in order to completely elucidate under which conditions and for which processes a gene product is produced and consequently understand the fundamentals of gene regulation in an organism [16, 17].

Bacterial genomes are complex, and elucidating their structure and dynamics is a challenging task. Since exact annotation solely using *in silico* tools is not yet possible, much experimental work has been performed in order to outline the transcriptional unit architecture for a variety of microorganisms,

including *E. coli* [17] and *G. sulfurreducens* [18]. However, determining the architecture of the genome is not enough to properly describe the transcription regulatory network of a species of bacteria; once the architecture has been laid out, it is then necessary to determine the condition-dependent interactions of various regulatory proteins, such as sigma factors and transcription factors, in order to fully reconstruct the genome. Thus, after a genome is annotated with the transcription regulatory network, it should be possible to predict how an organism responds to a variety of stimuli.

1.3 Ferric uptake regulator (Fur)

Fur is a 147 amino acid, homodimeric protein, each with two metal binding sites: one for Zn^{2+} and the other for Fe^{2+}. The metal binding site for Zn^{2+} is thought to stabilize Fur's protein architecture whereas the Fe^{2+} site is thought to allow for the reversible association depending on intracellular iron conditions [40]. This reversible association allows for Fur to function as a transcriptional repressor that is activated through a conformational change that occurs when binding to Fe^{2+} in iron-replete conditions. The different configuration allows for the protein to bind to target DNA sequences—typically located in the promoter region and referred to as Fur boxes—and inhibit transcription by RNA polymerase [20, 21]. As a result of its iron-dependent behavior, Fur is usually implicated as the central regulator of iron homeostasis in most bacterial species, including *E. coli [*20]*, Helicobacter pylori* [51-52],

Neisseria gonorrhoeae [50], *Desulfovibrio vulgaris* Hildenborough [49], and *Shewanella oneidensis* [46-48]. This central role results from the fact that in iron replete conditions, Fur is in its active form, and it binds to the promoter of genes coding for iron-uptake systems, thereby repressing the expression of iron-uptake proteins [20-23, 29-31, 39-40].

The role of Fur in bacteria has typically been investigated because of its direct involvement in the regulation of iron uptake as a result of its activation being dependent on intracellular iron levels. Iron is an essential nutrient for bacteria since it is often an enzymatic cofactor and electron transport protein component. In aerobic environments, however, soluble iron (Fe^{2+}) is scarce since it forms insoluble hydroxides [20]. Furthermore, even when soluble iron is available, its uptake must be carefully regulated in order to prevent the formation of highly reactive hydroxyl radicals via Fenton reactions [22]. In anaerobic environments, Fe^{2+} is readily available and the formation of Fenton radicals are not likely. Thus, iron regulation and the role of Fur have primarily been studied in aerobic microbes while anaerobic microbes have received little attention.

Iron uptake by the *Geobacteraceae* has not been investigated in detail, but it could be of significant importance even though the *Geobacteraceae* live in anaerobic environments. As a result of their unprecedented number of c-type cytochromes as well as high numbers of iron-sulfur cluster containing proteins [23], the iron content of *G. sulfurreducens* is unusually high. The

Geobacteraceae as a whole may have high assimilatory requirements for iron that would require a fine-tuned mechanism for iron uptake. Furthermore, the concentration of available soluble iron may extensively vary within *Geobacter's* environment. Though the *Geobacteraceae* naturally produce a source of Fe^{2+} by way of Fe^{3+} reduction, much of this soluble Fe^{2+} forms solid minerals, such as iron sulfide and siderite, as a result of the abundance of sulfur produced [23-26]. Conversely, soluble iron levels may also be substantially high under sufficiently high Fe^{3+} reduction rates. As a result of the amount of soluble Fe^{2+} within its environment constantly being under flux, the *Geobacteraceae* may need to tightly regulate its uptake and assimilation of iron in response to environmental levels.

Although Fur is typically thought to be a repressor that becomes active as a result of the configuration change that occurs when it binds to Fe^{2+}, it has been hypothesized and documented that Fur can also function as a transcriptional activator [45-47] and can also be active when it is not bound to iron [21]. The iron-free activated form of Fur, apo-Fur, has been documented best in *H. pylori*, in which apo-Fur has been shown via microarray analysis to directly regulate the *pfr* gene, which codes for non-heme iron-containing ferritin that controls the intracellular amount of free ferrous iron by incorporating ferrous iron and stores ferric iron formed after oxidation [21, 30, 31]. This apo-Fur regulation seems to not be conserved across species, implying that despite the conservation of Fur proteins across bacterial species,

H. pylori Fur has unique features that allow for it to regulate transcription while not complexed to iron [43]. Furthermore, although Fur has typically been thought to have function primarily as a regulator of genes involved in iron uptake and assimilation, studies across multiple species of bacteria have proven it to be a global regulator that regulates genes involved in virulence, motility, and oxidative stress response [39, 40]. Thus, it can be seen that the regulation of Fur is not simply as straightforward as being the iron-dependent repressor in responsible for iron homeostasis. Thus, it could be of interest to determine the global regulatory role of Fur in *G. sulfurreducens* and how much of its regulation is a result of direct, rather than indirect, mechanisms.

1.4 Genomic profiling of transcription factors

Until recently, it was not possible to study the direct regulation of transcription factors on a genomic scale, and regulatory information had to be garnered through other means, usually microarray analysis. For example, in a previous, but unpublished, study of Fur in *G. sulfurreducens* performed by the Lovely group in the University of Massachusetts, Amherst*,* a mutant strain with the Fur gene deleted was created, and in order to characterize the function of Fur, comparisons were performed between the expression profiles of the wild-type and the mutant strain grown in iron-replete conditions as well as between the profiles of the wild-type strain cultured in iron-replete and iron-depleted conditions. Genes that were differentially expressed in both comparisons were

clustered together according to fold change in expression. By comparing the gene clusters to predicted operon organization, a list of operons predicted to be regulated by Fur was generated. To further determine whether or not genes and operons were under the direct control of Fur, their upstream regions were searched for putative Fur-box motifs [23, unpublished study]. While methods like these do shed some light onto the global role of the Fur regulon, they are limited in describing its role comprehensively as a result of their indirectness. It is nearly impossible to tell which genes are differentially expressed as a direct result of Fur, which genes are differentially expressed as a result of some other factor (e.g. growth limitation, cellular stress), and which genes are differentially expressed as a result of a regulatory cascade in which Fur is involved.

Comprehensive, genome-wide transcription factor binding maps elucidated by direct means have become possible as a result of the development of several different techniques. One such technique is chromatin immunoprecipitation coupled with microarray analysis, known commonly as ChIP-chip. Although there are many variations, generally the protocol for ChIP-chip begins with using formaldehyde to cross-link protein-DNA complexes *in vivo.* From there, the desired protein-DNA complexes are enriched by using a specific antibody for the protein of interest. Finally, the cross-links are reversed, and the DNA is purified and then hybridized onto a microarray (Figure 1.2). The data obtained from the microarray are then

normalized and manually curated so that genomewide binding can be identified [27].

Thus, although the role of Fur in *G. sulfurreducens* has previously been investigated [23, unpublished study], the use of ChIP-chip in combination with gene expression profiles should give an experimentally validated, global map of site-specific binding, thereby elucidating the binding profile of Fur in *G. sulfurreducens.* By comparing the binding profile to gene expression data, it should be possible to define the transcriptional regulation of Fur.
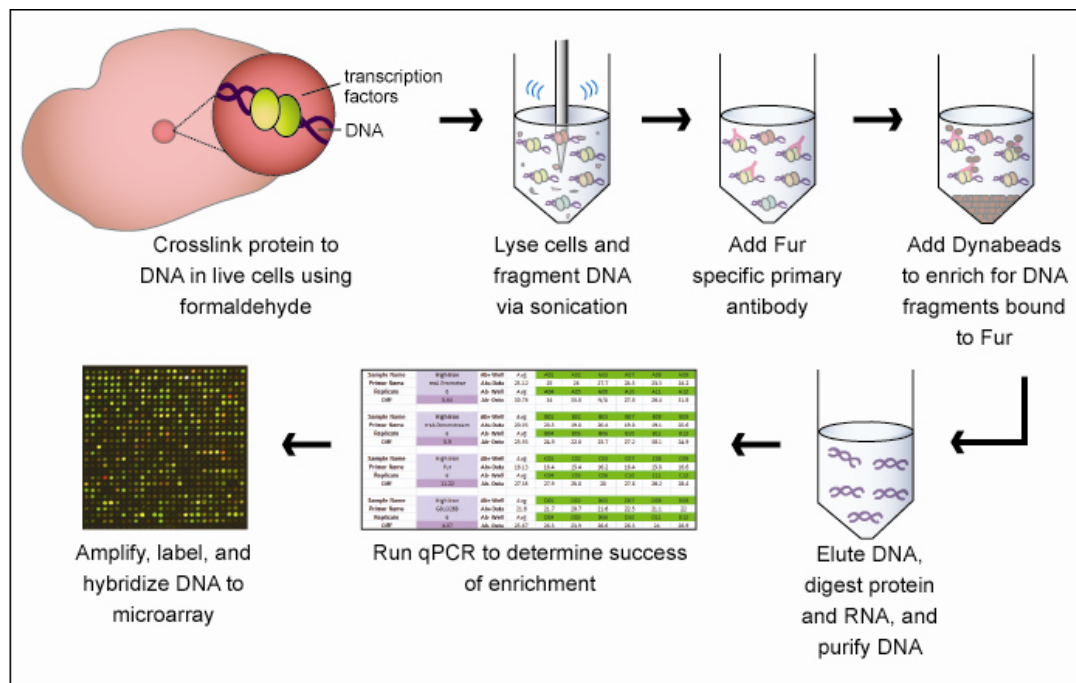
Figure 1.2: An illustration of the ChIP-chip protocol for Fur.

2. MATERIALS AND METHODS

2.1 *G. sulfurreducens* culturing conditions

2.1.1 "High" iron condition (trace mineral stock + ferric citrate): "High" iron conditions were simulated by growing anaerobic triplicate *G. sulfurreducens* cultures in 100 ml ferric citrate acetate medium (FCA) in 125 ml serum bottles at 30 °C. 5 mM acetate was used as th e electron donor and 55 mM ferric citrate was used as the electron acceptor. In addition to the iron source resulting from the reduction of ferric iron to ferrous iron, 1 mg of ferrous sulfate ($FeSO_4*7H_2O$), or 27.8 mM Fe^{2+}, was added to the media via a trace mineral stock. Cultures were harvested at mid-log phase, which was verified by using a previously described assay using ferrozine to determine the concentration of Fe^{3+} that had been reduced to Fe^{2+} [26]. The amount of iron determined to be reduced at mid-log phase had previously been established by making a growth curve of reduced iron concentration versus time. In this case, cultures were grown until 22.5-23.4 mM of iron had been reduced.

2.1.2 "Sufficient" iron condition (trace mineral stock only): G. sulfurreducens cultures were grown in triplicate under anaerobic conditions in 100 ml of fresh water acetate/fumarate medium (FWAF). 5 mM aceteate was used as the electron donor, and 27.5 mM fumurate was used as the electron acceptor. 1 mg of ferrous sulfate ($FeSO_4*7H_2O$), or 27.8 mM Fe^{2+}, was added to the media via a trace mineral stock. Cultures were grown until mid-log

phase, which was determined by a previously established growth curve of optical density versus time.

2.1.3 "Limited" iron condition (iron not added to trace mineral stock): G. sulfurreducens cultures were grown in triplicate under anaerobic conditions in 100 ml of "iron-limited" FWAF media. 5 mM aceteate was used as the electron donor, and 27.5 mM fumurate was used as the electron acceptor. To limit iron, iron was not added into the trace mineral stock, and normal fresh water acetate/fumarate cultures were passaged at least two times prior to experimentation. Cultures were grown until mid-log phase, which was determined by a previously established growth curve of optical density versus time.

2.1.4 Fur knock-out mutant condition: A mutant strain of *G. sulfurreducens* with the Fur gene deleted was created by the Lovely group in Ameherst. The mutant strain was grown under anaerobic conditions in 100 ml of "sufficient" iron FWAF media, with 5 mM aceteate used as the electron donor and 27.5 mM fumurate used as the electron acceptor. Cultures were grown in triplicate until mid-log phase as determined by a previously established growth curve of optical density versus time.

2.2 ChIP-chip experimental methods

2.2.1 Transformation of E. coli with Fur vector: Two primers were designed to allow for the PCR amplification of the Fur gene, the transformation of the Fur gene into *E. coli*, and the subsequent glutathione s-transferase (GST) purification of Fur protein from transformed *E. coli.* The N-term primer introduced the EcoRI restriction enzyme site and had the following sequence: GGAGAATTCATGAAGCGGGCAAAAAGC. The C-term primer contained the sequence for the NotI restriction enzyme site and had the following sequence: ATAGTTTAGCGGCCGCTCACTATTTGCGACATCGGCGG. The final PCR product created from the two primers was the sequence for Fur protein with a GST tag fused onto the N-terminus.

PCR was carried out according to the standard Stratagene protocol. The two primers were mixed with .25 mM dNTP mix, 100 ng DNA template, PfuUltra® II fusion HS DNA Polymerase (Stratagene), and 10X PfuUltra® II reaction buffer (Strategene) under the parameters detailed in Table 2.1. The PCR product's size was confirmed by running the amplified sample on 2% agarose gel (Invitrogen). The PCR product was then purified using QIAquick PCR Purification Kit (Qiagen) and then confirmed to be the sequence for the Fur-GST fusion protein by sequencing.

Table 2.1: PCR cycling parameters for amplification of sequence for Fur-GST fusion protein

Segment	Number of Cycles	Temperature	Time
1	1	95	2 min
2	3	95	20 sec
		Initial T_m	20 sec
		72	15 sec per kb
3	27	95	20 sec
		All T_m	20 sec
		72	15 sec per kb
4	1	72	3 min
		4	hold

Restriction enzymes EcoRI and NotI were used to digest cloning vector pGEX-6P-1 (GE Lifesciences) as well as the ends of the PCR product. 3-fold molar excess of PCR product was inserted into 50 ng of vector using 1 µl Quick T4 DNA Ligase (NEB) in 10 µl 2X Quick Ligation buffer (NEB), and the insertion was verified by sequencing. After verification, the vector was transformed into *E. coli* (One Shot® BL21(DE3) Chemically Competent *E. coli*, Invitrogen). Transformed *E. coli* was then plated onto LB agar plates coated with ampicillin and incubated overnight at 37 ℃. Colonies were picked and grown overnight at 37 ℃ in 4 ml LB containing ampicillin diluted 1:1000. Aliquots of cultures were preserved 1:1 in glycerol at -80 ℃. Plasmid DNA was purified from transformed *E. coli* using QIAprep Spin Miniprep Kit (Qiagen). Success of transformation was verified by sequencing.

2.2.2 Purification of Fur protein: Cells from preserved aliquots of successfully transformed *E. coli* were cultured overnight at 37 °C in 4 ml LB containing ampicillin diluted 1:1000. Cultures were then diluted 1:100 in 400 ml fresh LB medium containing ampicillin diluted 1:1000 and were grown at 37 °C until OD_{600} reached 0.5-0.6 absorbance units. Lactose analog isopropyl β-D-thiogalactoside was added to cultures at concentration of 0.5 mM in order to induce over-expression of the Fur-GST fusion protein, and cultures were incubated at room temperature for another three hours. Cultures were pelleted by centrifuging (7700 x g, ten minutes, 4 °C), and cell pellets were resuspended with 20 ml phosphate buffered solution (PBS). Cells were then sheared using Misonix Sonicator 3000 (power 2.5, 2x20 second, ten second interval). Triton X-100 solution was added to a final concentration of 2%, and cell solutions were gently mixed for 30 minutes to aid in solubilization of the fusion protein. Cell solutions were centrifuged (12,000 x g, ten minutes, 4 °C), and supernatants, which contained the Fur-GST fusion protein, were saved while cell debris pellets were discarded.

Supernatants were applied to GSTrap FF 1 ml columns (GE Lifesciences) in order to separate the Fur-GST fusion protein from solution. To cleave the Fur protein from the GST tag, the columns were then washed with PreScission cleavage buffer (50 mM Tris-Hcl, 150 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, pH 7.0) and then incubated overnight at 4 °C with 100 units of PreScission Protease (GE Lifesciences). Fur was eluted with elution buffer

(50 mM Tris-Hcl, 10 mM reduced glutathione, pH 8.0) and its concentration was quantified using Quant-IT Protein Assay Kit (Invitrogen).

2.2.3 *Production of specific Fur antibody:* Antibodies targeting Fur protein were produced by Young-Seoub Park. Purified Fur protein mixed with an adjuvant was injected into mice. Mice were immunized every two weeks, and serum was harvested after two months of inoculation. Fur antibody was purified from the serum using Protein G Chromatography Cartridges (Pierce) and then concentrated by centrifugation in microcon YM-3 tubes (Amicon). The purified antibody was tested via Western Blot against purified Fur protein as well as *G. sulfurreducens* whole cell lysate to validate antibody specificity for Fur. To produce whole cell lysate, *G. sulfurreducens* cultures were grown anaerobically in fresh water acetate/fumarate medium at 30 ℃, containing 5 mM acetate as the electron donor and 27.5 mM fumarate as the electron acceptor. Cells were harvested at mid-log phase (~24 hours) and pelleted. The pellet was then resuspended in 1X Nupage Sample buffer (Invitrogen) and boiled for 5 minutes.

2.2.4 *Chromatin immuoprecipitation*: After cultures had grown to mid-log phase, a modified version of the chromatin immunoprecipitation and amplification protocol developed by Cho, *et al.* [27] was followed in order to prepare ChIP-chip samples. 2.8 ml 37% formaldehyde was injected into 100

ml cultures and incubated for 25 minutes at room temperature in order to cross-link DNA-bound proteins to their respective DNA sequences. 4 ml of 2.5 M glycine was added and cultures were incubated for another 5 minutes to stop the cross-link reaction. Cells were then centrifuged (6000 rpm, 4 ℃, ten minutes), resuspended in 0.5 ml lysis buffer (10mM Tris-Hcl, 100mM NaCl, 1 mM EDTA), and incubated with 1 µl lysozyme (Epicentre) and 40 µl protease cocktail (50 mg protease inhibitor powder (Sigma), 250 µl DMSO, 750 µl H_2O) for 30 minutes at 37 ℃ on rotation. Immunoprecipita tion buffer (100 mM Tris-HCl ph 7.5, 200 mM NaCl, 2 mM EDTA, 2% Triton-X 100) was added and samples were incubated for another 30 minutes on ice. Samples were then sonicated using Misonix Sonicator 3000 (power 4.5, 4 x 20 seconds, ten second interval) and centrifuged (14,000 rpm, 4 ℃, 5 minutes). Cell debris pellets were discarded, and the supernatant of each sample was then divided into two separate 1 ml tubes. 20 µl containing 10 µg of Fur antibody was added to one of the 1 ml tubes for each sample. The other 1 ml tube for each sample remained the mock IP-DNA control, which did not have any antibody added to it. All samples were incubated overnight at 4 ℃.

Protein A Dynabeads (Invitrogen) were washed three times with Beads Wash Solution (125 mg bovine serum albumin (Sigma), 25 ml PBS), and then 50 µl Dynabeads were added to each sample. Samples were incubated with Dynabeads at 4 ℃ for six hours to facilitate the bindi ng of antibody-coupled chromatin to Dynabeads. Dynabeads were then pulled down with a MPC

magnet and washed multiple times. Dynabeads were then incubated overnight at 65 ℃ with elution buffer (50 mM Tris-HCl pH 8.0, 1% SDS, 1 mM EDTA) in order to elute antibody-Fur-DNA complexes. Dynabeads were then pulled down with the MPC magnet and supernatants were saved. Supernatant samples were incubated for two hours at 37 ℃ with 1 µl RNaseA (Qiagen) in 200 µl TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) in order to degrade RNA. Samples then were incubated another two hours at 55 ℃ with 4 µl protease K (20 mg/ml, New England BioLabs) in order to degrade proteins. Immunoprecipitated DNA (IP DNA) was then purified using QIAprep Spin Miniprep Kit (Qiagen).

2.2.5 Quantitative real-time PCR: To determine success of enrichment, qPCR was performed on IP DNA using iCycler real-time PCR detection system (Bio-Rad). Primers targeting promoter regions expected to be regulated by Fur were designed (Table 2.2). Primers for the promoter and downstream regions of rrsA, the gene which codes for 16S rRNA, were used as controls. 2x QuantiTect SYBR Green PCR master mix (Qiagen) was mixed with IP DNA, and qPCR was performed in duplicate for each primer. Amplification of IP DNA was performed if the average ΔC_t of IP DNA and mock-IP DNA samples was >3.

Table 2.2: Primers designed for qPCR on IP DNA. P = promoter, D = downstream.

Gene	Region	Primer sequence (5' → 3')	Gene product
GSU1379	P	*Forward:* AAGTGAATGTGCGACCGT *Reverse:* AATATCATCATAGAAGGCCCG	Ferric uptake regulator (Fur)
	D	*Forward:* GGCGAAGAGGGCGAAATC *Reverse:* ACGGAAAGACCTGCATGA	
GSU3268	P	*Forward:* TGAAGAAGGTGTTGGTGGTG *Reverse:* GCAAGCCGTAAGATCTCCTG	Ferrous iron transporter (feoA family)
	D	*Forward:* CCTACGCCTTCATGTGCTTC *Reverse:* GTAGACGATAAATGCCACCCC	

2.2.6 IP DNA Amplification: Random DNA amplification as described by Cho, et al. [27] was performed in order to uniformly amplify IP DNA for microarray hybridization. IP DNA from each sample was mixed with 5X Sequenase buffer (USB) and 1 µl of 40 µM Rand 9-Ns primer with sequence 5' TGGAAATCCGAGTGAGTNNNNNNNNN 3'. Mixture was heated to 94 ℃ and then cooled to 10 ℃. Round A mix, containing Sequen ase buffer (USB), dNTP, 0.75 µg bovine serum albumin (NEB), DTT (USB), 3.9 U Sequenase (USB), was added to each sample and PCR was performed (ramp 10 ℃ to 37 ℃ over eight minutes, hold at 37 ℃ for eight minutes, heat to 94 ℃ for two minutes, cool to 10 ℃). Sequenase dilution buffer (US B) and another 3.9 U Sequenase were added to each sample, and PCR was performed again (ramp 10 ℃ to 37 ℃ over eight minutes, hold for 37 ℃ fo r eight minutes, cool to 4 ℃). Samples were then diluted 1:4 with H$_2$O.

Diluted IP DNA samples were split into four separate samples, which were each mixed with 1 µl Rand univ primer (sequence: 5'

TGGAAATCCGAGTGAGT 3'), 10X *pfu* buffer (Stratagene), 5U *pfu* polymerase (Stratagene), and 0.25 mM dNTP mix. 27 cycles of PCR (94 ℃ for 30 seconds, 40 ℃ for 30 seconds, 50 ℃ for 30 second s, 72 ℃ for two minutes) were performed to amplify IP DNA. After amplification, IP DNA was purified (QIAprep Spin Miniprep Kit, Qiagen), and concentrations were determined (NanoDrop mass spectrophotometer, ThermoScientific).

2.2.7 Cyanine-Dye (Cy-Dye) labeling of IP DNA: If concentrations of IP DNA were at least 25 ng/µl, then Cy-Dye labeling was performed in-house according to a modified Nimblegen labeling protocol. 1 µg IP DNA was mixed with either 40 µl Cy5-Dye or Cy3-Dye nine-mers (TriLink Biotechnologies). Test samples were labeled with Cy5 while reference samples were labeled with Cy3. Samples were initially hybridized with Cy-Dye primers using PCR (98 ℃ for 10 minutes, 4 ℃ for 5 minutes, hold). 100 U Klenow (Roche) and 1 mM dNTPs were added to each sample, and samples were incubated for two hours at 37 ℃.

Reactions were stopped by adding EDTA, and labeled DNA was precipitated with 11.5 µl 5 M NaCl and 110 µl isopropyl alcohol. Samples were incubated for 10 minutes at room temperature in the dark, and then centrifuged at 14,000 rpm for 30 minutes at 4 ℃ to pellet. Pellets were washed with 80% ethanol and then dried. Samples were rehydrated with 21 µl H_2O and concentrations were determined (NanoDrop mass spectrophotometer,

ThermoScientific). Labeled IP DNA was then hybridized on to DNA microarrays.

2.2.8 Hybridization of labeled IP DNA to DNA microarray: 381,174 50-mer probes spaced 20 bp apart, with 30-bp overlap between two probes, were used to produce customized high density oligonucleotide microarrays (Roche Nimblegen) tiled with the entire *G. sulfurreducens* genome. Concentrated labeled IP DNA samples were hybridized onto tiled microarrays and scanned on an Axon GenePix 4000B scanner by the GeneChip™ Microarray Core at UCSD.

2.2.9 Validation of binding profile: To validate the binding profile for Fur elucidated from ChIP-chip, qPCR was performed at random with IP DNA on two binding sites determined from ChIP-chip data. Additionally, the ChIP-chip protocol was repeated using two conditions in which Fur should not be activated: wild-type strain cultured under iron-limited condition and Fur deletion mutant strain cultured under normal conditions. Under these conditions, it was expected that binding should not be observed. The qPCR experiments for the IP DNA samples from these ChIP-chip experiments were performed with primers from Table 2.2.

Table 2.1: Primers designed for qPCR based on Fur binding profile. P = promoter, D = downstream.

Gene	Region	Primer sequence (5' → 3')	Gene product
GSU2193	P	*Forward:* CCCTTTGCGGTGTGTCTATT *Reverse:* TGGTATCCTTTGCCCTACGA	Conserved hypothetical protein
	D	*Forward:* TGCTACCGACAATCAATGCC *Reverse:* TTTGCGGATCCGGACTTTCT	
GSU1401	P	*Forward:* TGGTCAAGTTCGACTTCCTG *Reverse:* AGCTTGTAGCTCTCCTCGT	DNA polymerase III
	D	*Forward:* TACTCCATCTACGCCGTCAT *Reverse:* TCGATCACCTCAAGCTCCTT	

2.3 Gene expression profiles experimental methods

2.3.1 RNA extraction and isolation from G. sulfurreducens: Gene expression profiles were generated for further characterization of Fur regulon. At harvest, cultures were rapidly transferred to a tube with 1/10th sample volume of stop solution (5% buffer equilibrated phenol in 100% ethanol) and then centrifuged (8000 x g, 10 minutes, 4 ℃).

Cell pellets were resuspended in 400 µl of 6.7% TE/Sucrose buffer and split into two tubes. To each tube, 3 µl of 20% SDS with 1 µl SUPERase-In^{TM} (Ambion), 1 µl lysozyme (Epicentre), and 1 µl proteinase K (Invitrogen) were added in order to inhibit RNase activity as well as lyse cells. Cell suspensions were incubated at 37 ℃ for ten minutes. 1 ml TRizol (Invitrogen) was added to each solution, and solutions were shaken and then incubated at room temperature for five minutes. Samples were centrifuged (14000 rpm, 15 minutes) and supernatants were transferred to new tubes while cell pellets were discarded. 200 µl chloroform was then added, and samples were

incubated at room temperature for another two minutes. Phase separation was induced by centrifugation (<12000 x g, 15 minutes, 4 °C), and clean upper phase containing RNA was preserved. Equal volume isopropyl alcohol was added, and samples were kept at -20 °C for 1 hour to o vernight to precipitate RNA. RNA was pelleted (13000 rpm, 30 minutes, 4 °C), washed with 75% EtOH, dried, and then resuspended in 50 µl RNase free water. Samples were incubated at 37 °C for 30 minutes with 2.5 µl DNase I (Qiagen) and 10 µl RPOD buffer (Qiagen) in order to digest any remaining DNA. RNA was purified using RNeasy Mini Kit (Qiagen) and quantified by NanoDrop mass spectrophotometer (ThemoScientific).

2.3.2 cDNA synthesis: If at least 10 µg RNA was extracted, synthesis of first strand cDNA was performed using a modified version of the standard Affymetrix protocol. 0.75 µg Random primer (Invitrogen) was mixed with 10 µg RNA and incubated at 70 °C for ten minutes and then 25 °C for ten minutes. Samples were then mixed with 5X 1st strand buffer (Invitrogen), .01 M DTT (Invitrogen), 0.5 mM dNTP (10 mM dATP, dCTP and dGtp, 6 mM dTTP, 4 mM aminoallyl-dUTP), 8 µg/ml actinomycin D, 30 U SUPERase-in™ (Ambion), and 1500 U SuperScript™ II (Invitrogen). Actinomycin D was added to remove antisense transcripts [28].

Reverse transcription was performed (25 °C for ten minu tes, 37 °C for 60 minutes, 42 °C for 60 minutes, 70 °C for ten minu tes). Samples were

incubated at 65 ℃ for 30 minutes with 20 µl of NaOH in order to hydrolyze cDNA. 20 µl of HCl was used to neutralize the samples. cDNA was purified (PCR Purification Kit, Qiagen), using phosphate wash (5 mM KPO_4, 80% EtOH) and elution buffers (4 mM KPO_4) instead of provided PE and EB buffers.

2.3.3: Amino-allyl labeling and microarray hybridization: Samples were then dried by Speed-Vac and resuspended in 4.5 µl sodium carbonate buffer. cDNA was then coupled to Cy3 monoreactive dye (Amersham) and incubated at room temperature in the dark for 1 hour. Labeled cDNA samples were then purified (PCR Purification Kit, Qiagen) and fragmented to ~30-500 bp by incubating samples for six minutes at 37 ℃ with 0.2 U R Nase-free DNase I (Epicentre) per µg of cDNA.

After fragmentation, DNase I was deactivated by incubating samples at 95 ℃ for ten minutes and results were verified on 2% agarose gel (Invitrogen). Samples were precipitated by incubating with 10 µl of 3 M sodium acetate and 110 µl isopropyl alcohol for ten minutes and then centrifuging at 14000 rpm for 30 minutes at 4 ℃. Pellets were washed with 70% EtOH, dried and rehydrated in 13 µl H_2O. Concentrations were quantified (NanoDrop mass spectrophotometer, ThermoScientific) prior to microarray hybridization. If labeled samples contained at least 1.5 µg cDNA in 11.5 µl volume, then hybridization onto tiled microarrays was performed by the GeneChip[TM] Microarray Core at UCSD as previously described above.

2.4 Data analysis

2.4.1 ChIP-chip analysis: Each ChIP-chip experiment was performed in triplicate and each IP DNA sample had a corresponding mock-IP DNA sample, which did not have Fur-specific antibodies added during the ChIP-chip procedure. Forward strand and reverse strand data were considered as separate data sets. IP/mock-IP ratios were computed and binding regions for each data set were determined by finding peaks using an algorithm in NimbleScan software. Data was further refined by manual curation. A region was considered a binding region if a peak was present in six out of six data sets. A gene was considered controlled by Fur if the binding site was found between the -35 and -10 region of the promoter since that is the region in which active Fur normally binds [29]. Binding data was extended using operon predictions from the Database of prokaryotic OpeRons (DOOR). Data was integrated and compared to expression profile data.

2.4.2 Gene expression analysis: Each microarray experiment was also performed in triplicate. The final expression data was normalized according to all nine data sets. Fold changes, $\log_2$ difference values, and P-values were calculated. Genes were considered differentially expressed if they had a fold change of $\geq$ 2 and a P-value $\leq$ .05. Differentially expressed genes were analyzed and compared to ChIP-chip data.

3. RESULTS

3.1 Validation of IP DNA enrichment via qPCR

Triplicate IP samples were tested two times with each target gene, totaling six replicates, during qPCR verification. Average ΔC_t was calculated by taking the average of the difference in the number of cycles for IP DNA and mock-IP DNA. $\Delta\Delta C_t$ was calculated by taking the difference of promoter region average ΔC_t and downstream region average ΔC_t (Table 3.1). $\Delta\Delta C_t$ for *fur* and *feoA* was > 3 and so amplification of IP DNA was performed.

Table 3.1: qPCR results on IP DNA from "high" iron culture condition.

Gene	Promoter region average ΔC_t	Downstream region average ΔC_t	$\Delta\Delta C_t$
rrSA	5.64	5.9	**-0.26**
GSU1379	11.22	4.49	**6.73**
GSU3271	11.07	4.07	**7**

3.2 ChIP-chip data

3.2.1 Genes found to be directly controlled by Fur: In total, 148 Fur binding sites were identified by ChIP-chip. If a binding site appeared to correspond to genes on the forward and reverse strand, it was assigned to both strands. The extent of Fur's transcriptional control was further investigated by using the Database of prokaryotic OpeRons operon prediction for *G. sulfurreducens.* From ChIP-chip data combined with operon predictions from DOOR, it was determined that Fur controls 224 genes (Table A1). Genes were classified according to their product's cellular function (Table 3.2). All

together, genes were classified into 16 different cellular functions. The categories with the most genes identified to be under the control of Fur were again cell envelope, energy metabolism, transport and binding, and mobile and extrachromosomal element with 20, 17, 17, and 16 genes, respectively.

Table 3.2: Classification of genes with known function according to gene product's cellular function.

Cellular function	Number of Fur-controlled genes (w/ operon prediction)
Amino acid biosynthesis	5
Biosynthesis of cofactors, prosthetic groups and carriers	3
Cell envelope	20
Cellular processes	14
Central intermediary metabolism	3
DNA metabolism	1
Energy metabolism	17
Fatty acid phospholipied metabolism	6
Mobile and extrachromosomal element	16
Protein fate	7
Protein synthesis	9
Purines, pyrimidines, nucleosides, and nucleotides	2
Regulatory functions	11
Signal transduction	4
Transcription	3
Transport and binding proteins	17
Unknown function	26
Hypothetical proteins	60
TOTAL	**224**

3.2.2: Validation of binding sites: ChIP-chip binding sites were validated by running ChIP-chip on conditions in which Fur should be inactive: 1) iron-limited media and 2) Fur deletion mutant strain cultured in iron-sufficient media.

It must be noted that it was difficult to amplify the Fur deletion mutant samples and so less only half of the optimal concentration was hybridized onto microarrays. As expected, no binding was observed in these validation samples. Sites that could have been argued to be binding sites were not consistent across all profiles and so they were regarded as noise (Figure 3.1).
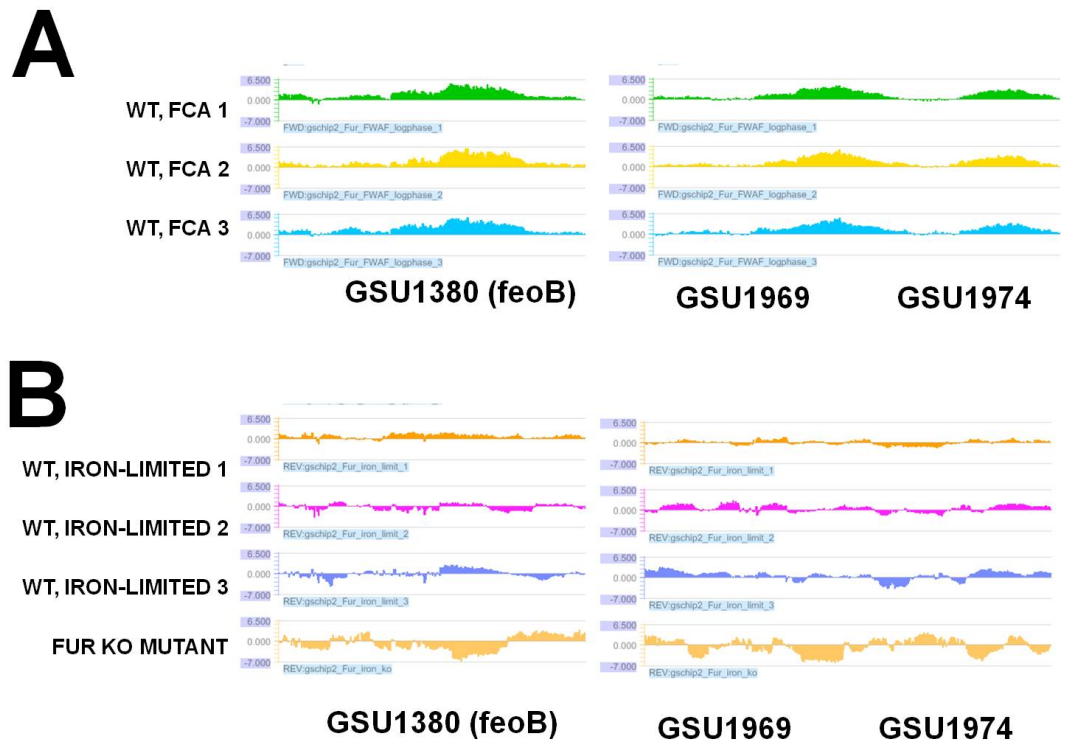
Figure 3.1: Example of SignalMap binding profile for GSU1380, GSU1969, and GSU1974 for A) WT in FCA media samples and B) WT in iron-limited media and *fur* knock-out mutant samples. Consistent binding peaks are visible in A) and not in B). WT = wild-type strain, FCA = ferric citrate acetate.

3.3 Gene expression data

3.3.1 Gene expression profiles of several different growth conditions:

Comparisons were made between gene expression of wild-type *G.*

sulfurreducens cultures in triplicate across several different conditions: 1) "sufficient" iron/"high" iron (FWAF vs. FCA); 2) "limited" iron/"high" iron (Fe-limited FWAF vs. FCA); and 3) "limited" iron/"sufficient" iron (Fe-limited FWAF vs. FWAF). A summary of the iron concentrations in each condition can be found in Table 3.3

Differential gene expression was determined by taking the difference between average, absolute values of $\log_2$ values. If the |$\log_2$ difference| was $\geq$ 1 (fold change $\geq$ 2) and if the P-value was $\leq$ .05, then the gene was considered differentially expressed. Genes were considered highly differentially expressed if their |$\log_2$ difference| was $\geq$ 3 (fold change $\geq$ 8). A list of all genes found to be differentially expressed in at least one expression profile can be found in Tables A2

In the "sufficient"/"high" expression profile, 185 genes were found to be differentially expressed in the "sufficient" iron condition (87 downregulated, 95 upregulated). 25 upregulated genes and two downregulated genes had a fold change $\geq$ 8. When comparing "limited"/"sufficient" iron, 95 genes were found to be differentially expressed (55 upregulated, 40 downregulated) in the "limited" iron condition. 14 upregulated genes and seven downregulated genes had a fold change of $\geq$ 8. For the comparison of "limited"/"high", 184 genes were found to be differentially expressed (76 upregulated, downregulated) in the "limited" iron condition. 32 upregulated and 17 downregulated genes had a $\log_2$ difference $\geq$ 3. A summary of all data can be found in Table 3.4.

Genes were classified according to cellular function (Table 3.5). A large portion of the differentially expressed genes in each of the three conditions coded for hypothetical proteins: 49, 21, and 49, in "sufficient"/"high", "limited"/"high", and "limited"/"sufficient", respectively. The majority of differentially expressed genes with known function were in energy metabolism for all three conditions.

Table 3.3: Summary of iron concentrations in each iron condition. FCA = ferric citrate acetate, FWAF = fresh water acetate fumarate.

Iron condition	Terminal electron acceptor	Growth rate (1/hr)	Fe^{2+} concentration (mM)
"High"	Ferric citrate	.085	27.8 mM from trace mineral stock and 22.5-23.4 mM from Fe^{3+} reduction
"Sufficient"	Fumarate	.090	27.8 mM from trace mineral stock
"Limited"	Fumarate	.051	None added

Table 3.4: Summary of the number of genes that were differentially expressed under each iron condition. "Upregulated" and "downregulated" are in reference to the condition with less iron.

Differentially expressed	"sufficient" vs. "high"	"limited" vs. "sufficient"	"limited" vs. "high"
Downregulated, fold change $\geq$ 8	2	7	17
Downregulated, fold change < 8	85	48	91
Upregulated, fold change < 8	73	26	44
Upregulated, fold change $\geq$ 8	25	14	32
TOTAL NUMBER	**185**	**95**	**184**

Table 3.5: Classification of differentially expressed (fold change $\geq$ 2, $\log_2$ difference $\geq$1) genes by gene product's cellular function. Number in parenthesis by each cellular function represents the number of unique genes found to be differentially expressed in each condition.

Cellular function	"sufficient" vs. "high"	"limited" vs. "sufficient"	"limited" vs. "high"
Amino acid biosynthesis (3)	0	2	2
Biosynthesis of cofactors, prosthetic groups and carriers (7)	3	2	3
Cell envelope (9)	4	4	6
Cellular processes (8)	2	5	8
Central intermediary metabolism (1)	0	0	1
DNA metabolism (1)	1	0	1
Energy metabolism (82)	49	17	59
Fatty acid phospholipied metabolism (2)	1	1	1
Mobile and extrachromosomal element (5)	2	3	1
Protein fate (13)	11	2	3
Protein synthesis (3)	0	3	3
Regulatory functions (10)	3	7	9
Signal transduction (4)	3	1	1
Transcription (1)	0	1	1
Transport and binding proteins (23)	12	14	14
Hypothetical proteins (79)	49	21	49
Unknown function (28)	17	8	15

3.3.2 Cross-comparisons between differentially expressed genes across different profiles: Cross-comparisons to the 184 differentially expressed genes from the "limited"/"high" expression profile were performed with the differentially expressed genes of the other two expression profiles in order to determine both the effects of the terminal elector acceptor and iron levels on differential gene expression. The "limited"/"high" expression profile was used as the reference profile since, in addition to containing the most differentially expressed genes, it also had the most variables, with both a variable terminal electron acceptor (fumarate vs. ferric citrate) and the largest different in iron concentrations ("limited"/"high"). Summaries of data can be found in Tables 3.6 and 3.7.

In this cross-comparison, it was observed that 75 genes with "sufficient"/"high" whereas 76 genes overlapped with "limited"/"sufficient". 13 genes were differentially expressed in all three profiles, and 46 genes were only found to be differentially expressed in the reference profile (Figure 3.2). Genes with a fold change of ≥ 8 in reference to the "limited"/"high" profile were also looked at in each cross-comparison and for all the cross-comparisons, there were more highly upregulated genes than highly downregulated genes (Table 3.6). Genes found in the cross-comparisons were also classified according to cellular function (Table 3.7). For the two cross-comparisons, most genes with known function were classified into energy metabolism. A significant number of genes were also seen to overlap in transport and binding

proteins and regulatory functions. For differentially expressed genes that overlapped in all three profiles, there were more genes in the transport and binding proteins category than the energy metabolism category.
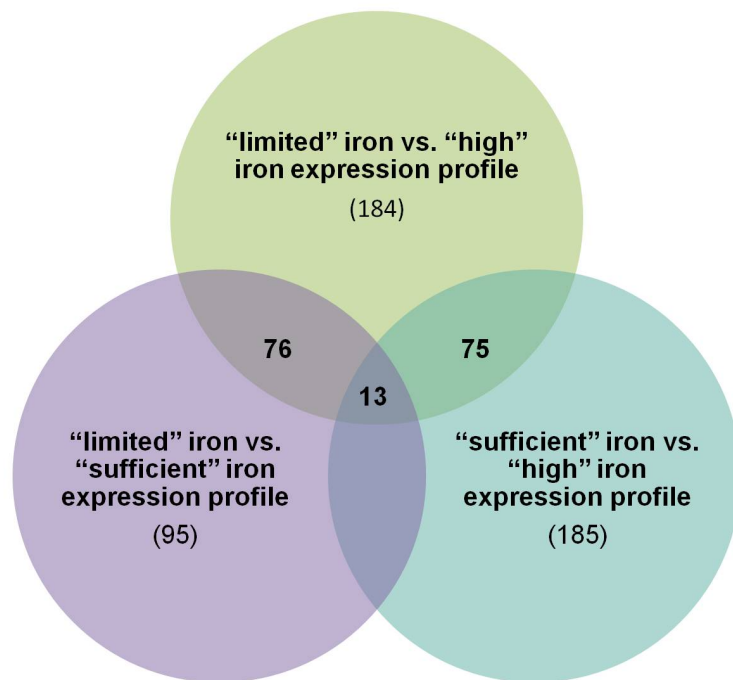
Figure 3.2: Summary of cross-comparisons of gene expression profiles. Numbers represent the number of genes in each expression profile and in each cross-comparison of expression profiles. 13 genes were found to be differentially expressed in all expression profiles.

Table 3.6: Summary of the number of genes found in cross-comparisons to "limited"/"high". A = "sufficient"/"high" expression profile, B = "limited"/"sufficient" expression profile, overlap = genes found in all three profiles, only in "limited" vs. "high" = genes not found in either cross-comparison. Fold change values are with respect to "limited"/"high" profile.

Differentially expressed	A	B	Overlap	Only in "limited" vs. "high"
Downregulated, fold change $\geq$ 8	10	4	0	3
Downregulated, fold change < 8	38	24	2	31
Upregulated, fold change < 8	8	24	0	12
Upregulated, fold change $\geq$ 8	19	24	11	0
TOTAL NUMBER	**75**	**76**	**13**	**46**

Table 3.7: Classification into cellular function of genes found to be differentially expressed in cross-comparisons with "limited"/"high" expression profile. A = "sufficient" iron vs. 'high' iron expression profile, B = "limited"/"sufficient" expression profile, overlap = genes found in all three profiles, only in "limited" vs. "high" = genes not found in either A or B. Number in parenthesis following cellular function represents number of unique genes.

Cellular function	A	B	Overlap	Only in "limited" vs. "high"
Amino acid biosynthesis (2)	0	1	0	1
Biosynthesis of cofactors, prosthetic groups and carriers (3)	1	1	0	1
Cell envelope (6)	1	4	0	1
Cellular processes (8)	2	5	0	1
Central intermediary metabolism (1)	0	0	0	1
DNA metabolism (1)	1	0	0	0
Energy metabolism (59)	26	16	2	19
Fatty acid phospholipied metabolism (1)	0	1	0	0
Mobile and extrachromosomal element (1)	1	0	0	0
Protein fate (3)	2	1	1	1
Protein synthesis (3)	0	3	0	0
Regulatory functions (9)	3	7	2	1
Signal transduction (1)	0	1	0	0
Transcription (1)	0	1	0	0
Transport and binding proteins (14)	4	13	3	0
Unknown function (15)	7	5	1	5
Hypothetical proteins (49)	23	16	4	15

3.4 Comparative analysis: ChIP-chip and gene expression profiles

3.4.1. Comparative analysis with a gene expression profile: The ChIP-chip binding profile was compared to the various gene expression profiles in order to further elucidate the regulation of Fur. There was very little overlap between the genes found in ChIP-chip compared to differentially expressed genes in any of the expression profiles (Figure 3.3). In summary, 32 out of the 224 genes with binding sites were found to be differentially expressed in at least one condition. Out of all the expression profiles, the "limited"/"high" profile had the most overlap with ChIP-chip binding data, with 27 out of 184 differentially expressed genes apparently containing Fur binding sites. However, the "limited"/"sufficient" profile contained the highest percentage of genes overlapping with ChIP-chip, with 22.1% (21 genes) overlapping.

For all three binding profiles, the number of upregulated genes with an apparent Fur binding site was greater than the number of downregulated genes. There were very few genes highly downregulated ($\log_2$ difference ≥ 3, fold change ≥ 8) that had an apparent Fur binding site (Table 3.8). When classifying overlapping genes with known function according to cellular function, it was seen that the cellular function with the most Fur-regulated genes was transport and binding proteins (Table 3.9). A list of all 32 genes can be seen in Table 3.9.

3.4.2. Comparative analysis with cross-comparisons: The ChIP-chip binding profile was also compared to cross-comparisons of expression profiles to see if there was a higher correlation to ChIP-chip data. Most of the cross-comparisons had similarly low levels of correlation, with only ~20-25% of differentially expressed genes in either cross-comparison found to have binding sites (Figure 3.4). However, when looking at the cluster of 13 genes found in all three expression profiles, it was seen that ten (76.9%) genes had Fur binding sites (Figure 3.4). These genes were highly upregulated (fold change $\geq$ 8, in reference to the "limited"/"high" expression profile) (Table 3.10). These ten genes were classified according to general function, and it was observed that there were four genes (40%) that coded for hypothetical proteins (Table 3.11). After further classifying the genes with known function, it was observed that the large majority were involved with either energy metabolism or transport and binding. A description of all ten of the overlapping genes with Fur binding sites can be found in Table 3.11.
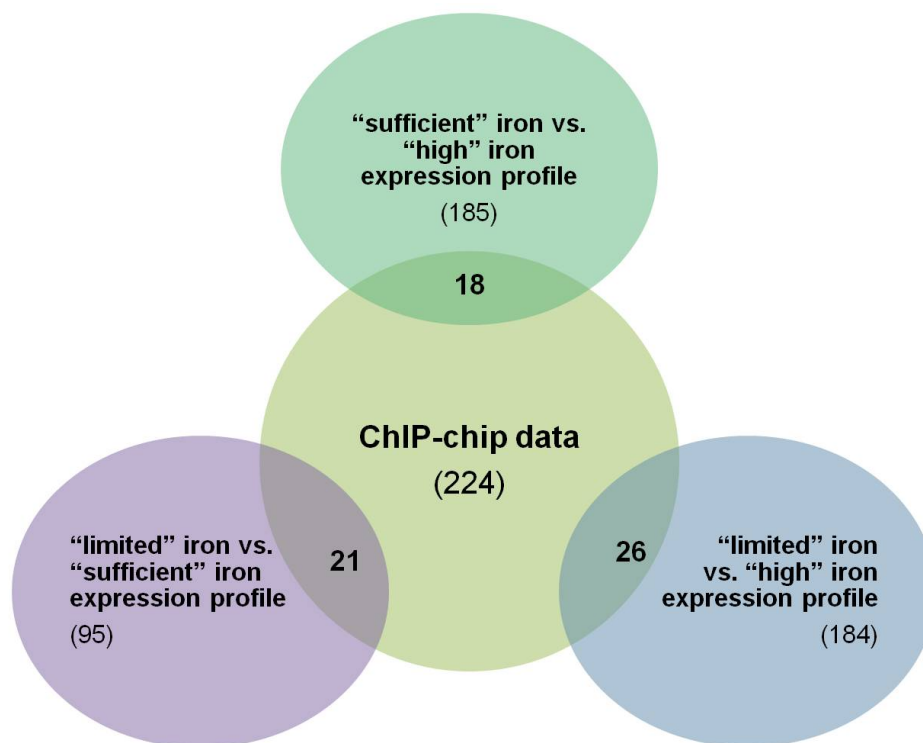
Total # of unique genes corresponding to ChIP-chip binding sites: 32

Figure 3.3: Summary of comparison of genes found to have Fur binding sites according to ChIP-chip data to differentially expressed genes from various expression profiles. In total, 31 genes with binding sites were found to be differentially expressed in at least one condition.

Table 3.8: Breakdown of genes found to be differentially expressed in expression profiles (Table A2) that also corresponded to ChIP-chip binding data (Table A1). Table is organized according to degree of differential expression.

# of differentially expressed genes overlapping ChIP-chip	"sufficient" vs. "high"	"limited" vs. "sufficient"	"limited" vs. "high"
Downregulated, fold change $\geq$ 8	0	2	2
Downregulated, fold change < 8	4	2	6
Upregulated, fold change < 8	3	10	1
Upregulated, fold change $\geq$ 8	11	7	18
TOTAL NUMBER	**18**	**21**	**27**

Table 3.9: Classification of the 32 genes overlapping in gene expression profiles and ChIP-chip binding profile with known function according to cellular function. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Main role	Gene	Gene name	Gene product	Expression profiles (1, 2, and/or 3)
Cell envelope (3)				
	GSU2940		rhodanese-related sulfurtransferase	2, 3
	GSU0832		lipoprotein, putative	2, 3
	GSU2133		lipoprotein, putative	2, 3
Cellular processes (2)				
	GSU0828		efflux pump, RND family, outer membrane protein	2, 3
	GSU0829		efflux pump, RND family, membrane fusion protein	2, 3
Energy metabolism (5)				
	GSU0341	nuoD	NADH dehydrogenase I, D subunit	2
	GSU2898	omcN	cytochrome c, 27-34 heme-binding sites	1, 3
	GSU2899		cytochrome c, 16-23 heme-binding sites	1, 3
	GSU0784	hybB	periplasmically oriented, membrane bound [NiFe]-hydrogenase integral membrane subunit	1
	GSU3274		cytochrome c, 1 heme-binding site	1, 2, 3
Hypothetical proteins (9)				
	GSU1060		conserved hypothetical protein	3
	GSU1381		conserved hypothetical protein	1, 2, 3
	GSU2200		conserved hypothetical protein	3
	GSU2897		conserved hypothetical protein	1
	GSU2938		conserved hypothetical protein	2, 3
	GSU3267		conserved hypothetical protein	1, 3
	GSU3273		conserved hypothetical protein	1, 2, 3
	GSU3272		hypothetical protein	1, 2, 3
	GSU3271		carbohydrate-selective porin OprB	1, 2, 3
Protein fate (1)				
	GSU2678		ATP-independent chaperone, alpha-crystallin/Hsp20 family	1, 2, 3
Regulatory functions (2)				
	GSU1379	fur	ferric uptake regulation protein Fur	1, 2, 3
	GSU0831		nitrogen regulatory protein P-II, putative	2, 3
Transport and binding proteins (7)				
	GSU2939		putative porin	2, 3

Table 3.9, cont.: Classification of the 32 genes overlapping in gene expression profiles and ChIP-chip binding profile with known function according to cellular function. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high.

Transport and binding proteins (7), cont.

GSU2982		TonB-dependent outer membrane receptor, putative	2, 3
GSU2482	kdpC	potassium-transporting ATPase, C subunit	1
GSU0830		efflux pump, RND family, inner membrane protein	2, 3
GSU1380	feoB-1	ferrous iron transport protein B	1, 3
GSU3268	feoB-2	ferrous iron transport protein B, putative	1, 2, 3
GSU3269		ferrous iron transport protein A, pseudo gene	1, 2, 3

Unknown function (3)

GSU1639		IPT/TIG domain protein	1
GSU1858		transcriptional regulator, putative	2
GSU3270		FeoA family protein	1, 2, 3
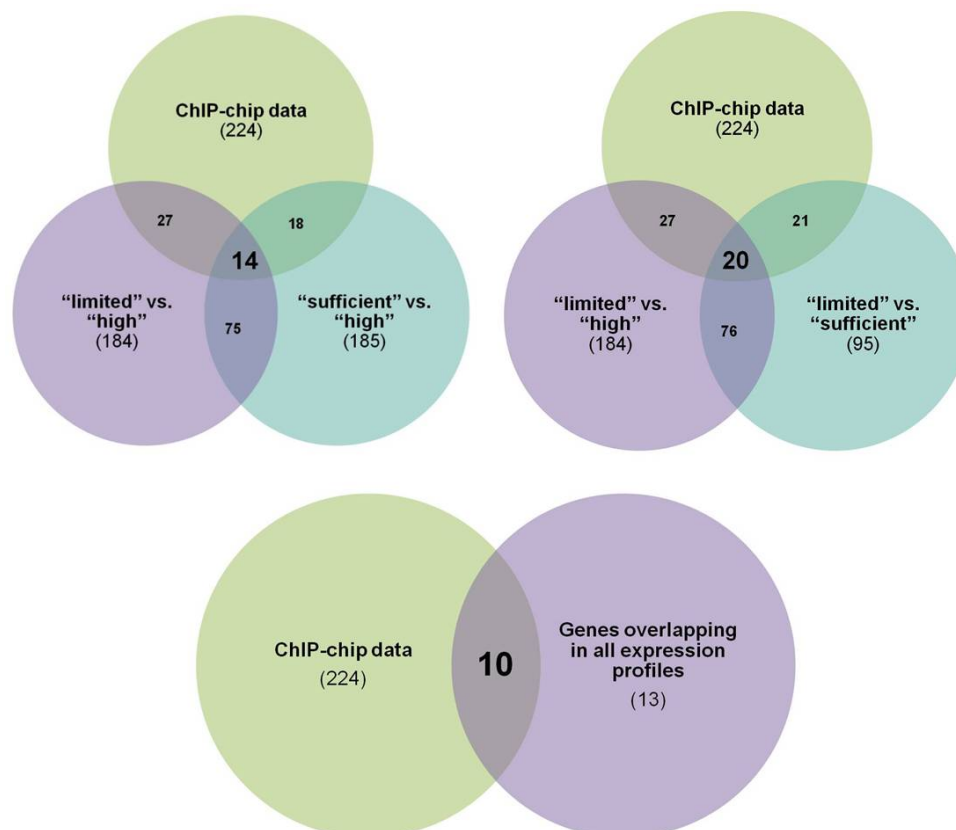
Figure 3.4: Summary of comparison of ChIP-chip binding data with differentially expressed genes found in cross-comparisons. The numbers represent the total number of genes found in each category.

Table 3.10: Comparison ChIP-chip binding data with differentially expressed genes found in cross-comparison of "limited"/"high" to other conditions. A = "sufficient"/"high" expression profile, B = "limited"/"sufficient" expression profile, overlap = genes found in all three profiles. Fold change values were in reference to "limited"/"high" differential expression values.

# of differentially expressed genes overlapping ChIP-chip	A	B	Overlap
Downregulated, fold change ≥ 8	0	2	0
Downregulated, fold change < 8	2	2	0
Upregulated, fold change < 8	0	0	0
Upregulated, fold change ≥ 8	12	16	10
TOTAL NUMBER	**14**	**20**	**10**

Table 3.11: Descriptions of all ten genes found in all expression profiles and ChIP-chip binding profiles.

Main role	Gene	Gene name	Strand	Door	Gene product
Energy metabolism (1)					
	GSU3274		-	30883	cytochrome c, 1 heme-binding site
Hypothetical protein (4)					
	GSU1381		+	30486	conserved hypothetical protein
	GSU3273		-	30883	conserved hypothetical protein
	GSU3272		-	30883	hypothetical protein
	GSU3271		-	GSU3271	carbohydrate-selective porin OprB
Protein fate (1)					
	GSU2678		+	GSU2678	ATP-independent chaperone, alpha-crystallin/Hsp20 family
Regulatory functions (1)					
	GSU1379	fur	+	GSU1379	ferric uptake regulation protein Fur
Transport and binding proteins (2)					
	GSU3268	feoB-2	-	30882	ferrous iron transport protein B, putative
	GSU3269		-	GSU3269	ferrous iron transport protein A authentic frameshift
Unknown function (1)					
	GSU3270		-	GSU3270	FeoA family protein

4. DISCUSSION

In the results portion, gene expression data from a variety of environmental iron conditions were compared to one another, Fur binding data were grouped according to functional category, and the two data sets were combined in order to characterize the Fur regulon. The following analysis and discussion serves to further elucidate the function of Fur and the overall relationship of iron on indirect and direct transcriptional regulation in *G. sulfurreducens.*

4.1 ChIP-chip binding profile

224 genes were found to be directly regulated by Fur according to the ChIP-chip binding profile and DOOR operon prediction, which were many more genes than expected. The genes found to have Fur binding sites spanned across a wide range of cellular functions, which suggests that Fur's role as a global regulator is not only mediated through indirect mechanisms. It is unlikely that the binding sites observed are false positives since the binding profile was based on six separate sets of normalized data, and the binding data was validated by performing ChiP-chip experiments under conditions in which it was expected that Fur would be inactive ("limited" iron condition and Fur knock-out mutant condition). As expected, it was observed there were no binding sites in these sets of data thereby confirming the Fur binding profile.

4.1.1 The role of Fur on various cellular functions: As mentioned before, Fur was found to have binding sites for genes in a wide variety of cellular functions. These binding sites were then extended with DOOR predictions in order to determine the extent of direct regulation by Fur. The majority of the genes found had ties to iron, such as their gene product having several heme binding sites or utilizing a high-iron containing protein. However, there were also many genes with no apparent ties to iron. Although it is difficult to elucidate reasons for Fur to regulate these non-iron related genes, it has been observed in previous studies that Fur does go beyond iron regulation and that it can regulate non-iron related genes, such as those for virulence factors, motility, metabolic pathways, and oxidative stress responses [29, 30, 39, 40]. It should also be noted that there were 60 genes that coded for hypothetical proteins and 26 genes that coded for proteins with unknown functions that were found to have a Fur binding site. The combination of the large number of genes coding for hypothetical proteins and genes of unknown function implies that there is still much that is not known about *G. sulfurreducens*, especially in terms of its relationship with iron. Most of the genes found in the ChIP-chip profile coded for proteins involved with cell envelope, energy metabolism, mobile and extrachromosomal elements, transport and binding, cellular processes, and regulatory functions. All the other cellular functions had $\leq$ ten genes found to be directly regulated by Fur.

Cell envelope (20 genes): The cell envelope in Gram-negative bacteria is composed of four major components: the inner cytoplasmic membrane, the periplasmic space, an outer membrane, and the polysaccharide surface layer. A wide range of genes coding for cell envelope proteins were found in the ChIP-chip data, including genes for glycosyl transferases (GSU1509-1510), lipoproteins (GSU0501, GSU0832, GSU2104, GSU2133), and several membrane proteins. The large number of cell envelope genes with Fur binding sites could be because the compositions of the cell membrane may need to be directly regulated and remodeled in order to allow for the optimal transfer of ferrous iron. A study of *Vibrio cholerae* in which cells were treated with polymyxin B—an antibiotic that disrupts the structure of the bacterial cell membrane—caused oxidative stress and an imbalance in iron homeostasis [54]. One of the implications of the study was that changes in the structure of the cell envelope, including genetic modifications, could cause downstream consequences in iron transport and uptake. Thus, it is possible that in *G. sulfurreducens*, in order for the cell to maintain optimal iron homeostasis, several key cell envelope genes are directly regulated by Fur.

Energy metabolism (17 genes): Several genes involved with energy metabolism, especially energy transport genes, are highly dependent on iron. As such, it is not surprising that a large number of energy metabolism genes were found to be directly regulated by Fur. For example, six genes found in the ChIP-chip data code for c-type cytochromes (GSU2724-2725, GSU2767,

GSU2898-2899, GSU3274), two of which are thought to have a large number of heme binding sites, with 16-23 and 27-34 heme binding sites, respectively. There was also a ferredoxin family protein (GSU0848), which is a protein that typically contains four iron-sulfur clusters, and an iron-sulfur cluster-binding protein (GSU0729) found to have Fur binding sites. A few of the other genes with Fur binding sites code for enzymes that utilize ferredoxin.

However, there were also a number of genes with no apparent direct ties to iron. However, maybe as a result of controlling the transcription of several iron-related electron transport proteins, Fur may also need to control other electron transport proteins in order to maintain homeostasis in the cell when the iron-related electron transport proteins are differentially expressed.

Transport and binding proteins (17 genes): Several iron transport-related proteins were found to be directly regulated by Fur. Uptake of iron, in both its Fe^{2+} and Fe^{3+} form, has typically been shown to be dependent on a variety of transport systems. Fe^{2+} can be thought to be the preferred form of iron since, unlike Fe^{3+}, it can be directly transported as a result of being relatively soluble at neutral pHs. For the transport of Fe^{2+}, two systems in the Feo family have been found in the *G. sulfurreducens*. The Feo system is typically comprised of three proteins: FeoA, a small, soluble SH3-domain protein; FeoB, a large protein with two "Gate" motifs which likely function as the Fe^{2+} permease; and FeoC, a small protein apparently functioning as an [Fe-S]-dependent transcriptional repressor [32]. In the anaerobic *G.*

sulfurreducens, two iron uptake systems belonging to the Feo family have been putatively identified: FeoB (GSU1380) and FeoB-2 (GSU3268). FeoB-2 was identified to be differentially expressed in each of the three conditions. In addition, genes coding for a FeoA protein (GSU3269) and a FeoA family protein with unknown function (GSU3270) were also identified.

The majority of Gram negative bacteria grow in aerobic conditions in which free ferrous iron is uncommon [32], necessitating systems which can uptake and transport Fe^{3+}. To facilitate this need, Gram negative bacteria typically use specific compounds, like siderophores and homophores, to procure iron by forming a complex with Fe^{3+} or heme from an extracellular source. The transport of the extracellular ferric-siderophore complex into the periplasmic space requires energy, which is not immediately available on the outer membrane. Thus, energy is typically provided by the proton motive force of the cytoplasmic membrane, which is coupled to the outer membrane by a complex of three proteins: TonB, ExbB, and ExbD. The TonB-ExbB-ExbD complex is then typically transported across the cytoplasmic membrane by a variety of ATP-binding cassette (ABC) transporters, which are the most common transport systems for a variety of nutrients [37]. Although *G. sulfurreducens* is an anaerobe and thus does not have a stringent need for a Fe^{3+} transport system and no genes coding for siderophore-like proteins have yet been identified in its genome, it could still be possible that *G. sulfurreducens* does transport Fe^{3+}. Several genes for ABC transporters

(GSU1097, GSU2009, GSU2187-2188, GSU2270) and TonB proteins (GSU2981-2982) were found to be directly regulated by Fur. It is possible that under some conditions, the environmental concentrations of Fe^{2+} are lower and/or the rate of Fe^{3+} reduction is too low to support growth, and *G. sulfurreducens* may need to meet its minimum iron requirements via the uptake and transport of Fe^{3+}. It has been proven that *G. sulfurreducens* can grow in oxic conditions [55], in which the concentrations of soluble Fe2+ are much lower, and it is possible that the uptake of Fe3+ is utilized under such conditions.

In addition to iron uptake and transport proteins, genes coding for Czc family heavy metal efflux proteins (GSU0828-0830) and another efflux protein for an unknown substrate were found in the Fur binding profile. The Czc family of proteins has been used for a variety of heavy metals. In *E. coli*, proteins with homology to the Czc system has been used for resistance of copper and silver ions, and in *Ralstonia eutrophus*, the Czc system is used to detoxify the cell of cobalt, zinc, and cadmium [38]. The regulation of these proteins could help *G. sulfurreducens* maintain homeostasis in its intracellular iron levels, which would be important considering the iron levels in its environment are constantly under flux. It is also possible that with the upregulation of iron transport, there could also be an increase of transport of other, more toxic heavy metals, necessitating heavy metal efflux proteins. As a note, GSU0828

and GSU0829 were categorized by NCBI as "cellular process" proteins whereas GSU0830 was classified as a "transport and binding" protein.

Mobile and extrachromosomal element (16 genes): The relatively large number of genes coding for mobile and extrachromosomal elements found to have Fur binding sites is surprising since there is not an apparent connection between intracellular iron concentration and transposons. Assuming that *G. sulfurreducens* is normally found in environments replete with iron, Fur could be used as a way to limit the effects of transposable elements. More study on the role of mobile elements in *G. sulfurreducens* would need to be performed in order to further elucidate a meaning on this result.

Cellular processes (14 genes): Fur was also found to directly control several genes coding for proteins involved with motility, including flagellar biosynthetic proteins (GSU3054-3056), and flagellar motor switch proteins (GSU0421-0422). A previous study using a Fur titration assay (FURTA) found a functional Fur box in front of *flhD* in *E. coli*, allowing for Fur to control flagellum assembly and chemotaxis. It was hypothesized that Fur would inhibit the transcription of these proteins so that the cells would be prevented from leaving iron-replete environments [39]. It is possible that Fur in *G. sulfurreducens* could have a similar function, only opposite in intention; these genes could typically be inhibited by Fur, but in "limited" iron conditions where Fur is inactive, these genes are no longer inhibited, which enhances the ability of the cell to travel to higher iron conditions. Fur has also been associated with

the control of flagellar biosynthesis and motility in *H. pylori,* except in positive regulation. It was hypothesized in that case that Fur was indirectly regulating the flagellar genes by interfering with the binding to the promoter of an acid response repressor, ArsR [52].

Regulatory functions (11 genes): Fur appears to regulate itself, which would consequently allow for the transcriptional regulation of Fur to be dependent on intracellular iron levels. Fur was also found to regulate another transcriptional regulator with a metal-binding domain (GSU2980). However, another transcriptional regulator (GSU1268), with no apparent ties to iron, was also found in the ChIP-chip binding profile. It should also be noted that binding sites were found for RNA polymerase sigma-32 factor (GSU0655) and a transcription elongation factor (GSU1586), which further emphasizes the potentially vast role that Fur has in transcriptional regulation in *G. sulfurreducens,* which further emphasizes the potentially vast role that Fur has in transcriptional regulation. Sigma-32, also known as RpoH, is a sigma factor that is typically active under heat shock conditions. The binding site of Fur to the promoter of RpoH potentially ties Fur's regulation to heat shock conditions. If Fur does directly control RpoH and other regulatory functions, then Fur has a much larger role in the regulation of *G. sulfurreducens* than originally thought.

4.1.2 Iron-dependent genes not found in ChIP-chip binding profile: As mentioned in the previous discussion, several of the genes found to be directly

regulated by Fur were somehow iron-dependent, such as using iron as a cofactor or being related to the transport of iron. At the same time, many seemingly iron dependent genes, such as the large majority of cytochromes, Fe-S cluster binding proteins, and metal ion efflux pumps, did not appear to have a Fur binding site according to the binding profile.

Notably, none of the ferritin (GSU1307) or ferritin-like (GSU0384, GSU0479, GSU1642, GSU2193, GSU2967, GSU3289, GSU3293) proteins were found to have a Fur binding site. Ferritin is a protein that stores iron and releases it in a controlled manner and is typically the major iron storage protein in an organism [36]. The lack of Fur binding sites in the genes coding for these ferritin proteins could be for several reasons. One is that proteins with ferritin-like domains are not necessarily involved in iron storage; ribonucleotide reductase contains a ferritin-like domain, for example [36]. Another possible reason is that *G. sulfurreducens* has another mechanism for iron storage other than ferritin. It has been previously hypothesized that the role of ferritin in anaerobic bacteria to defend against oxygen, but the mechanisms behind this proposed function are currently unknown [49]. It is also possible that these genes—and other iron-dependent genes not found to have Fur binding sites—are under the control of another transcriptional regulator, such as the iron-dependent IdeR (GSU1382), also known as DtxR. IdeR is known to regulate the expression of iron uptake proteins when activated by excess intracellular Fe2+. Studies have shown that Fur and IdeR

have similar tertiary structures despite having little conservation at the nucleotide level [35]. If another study were performed on the IdeR transcription factor, it would be interesting to compare its transcriptional regulation with Fur and to determine how the two iron-dependent transcription factors work together in tandem.

4.2 Gene expression profiles

Gene expression profiles were generated for further characterization of the Fur regulon. Comparisons were made between gene expression of wild-type *G. sulfurreducens* cultures in several different conditions with differing levels of extracellular iron ("high," "sufficient," or "low"). The gene expression profiles also served to elucidate changes in differential expression caused by changing iron levels. Prior to comparing the expression data with the ChIP-chip data, several observations were made about the expression data. Then, in order to better understand the role of iron in *G. sulfurreducens,* differentially expressed genes from the "limited"/"high" profile were compared to the differentially expressed genes from the other profiles in order to further define which changes in differential expression were likely to have been critically affected by changes in environmental iron concentration.

4.2.1 Expression data analysis: For the following section, please refer to Table 3.4 in the results for differences between the different conditions

("high," "sufficient," and "limited"). It was expected that growth in "limited" iron media versus growth in "high" iron media would have the most change in gene expression data since it contained the most variables: the terminal electron acceptor (fumarate versus ferric citrate) and the highest differences in iron concentration and growth rate ("limited" versus "high" iron). However, this profile actually has fewer differentially expressed genes compared to the "sufficient"/"high" profile, in which the major variable was the terminal electron acceptor. This result is likely to be at least partly caused by inconsistencies in the expression data for the "limited" condition compared to the data for the "high" and "sufficient" conditions. Inconsistent data would result in higher overall P-values for many genes found in the expression profile comparisons with the "limited" condition, resulting in these genes not being regarded as differentially expressed. The more inconsistent data for the "limited" condition is likely because the condition was not as well controlled as other conditions since it did not have a consistent amount of iron added to the media, resulting in inconsistent growth and stress across the different replicates, which would lead to inconsistent expression data. The "limited" condition had been established via three successive transfers of a "sufficient" condition culture into media in which no iron had been added. As such, there was a residual amount of iron from the original culture that had been transferred, and the amount of residual iron was impossible to control. Using an iron chelator had been considered, but it was unknown if the chelator could also affect

differential expression in *G. sulfurreducens* and so use of a chelator was not opted. Thus, it is possible that at such a residual level, the slightest differences in iron concentration could make a large difference in gene expression.

For all the profiles, energy metabolism was the cellular function with the most differentially expressed genes, with 49, 17, and 59 genes in the "sufficient"/"high," "limited"/"sufficient," and "limited"/"high" profiles, respectively. Transport and binding protein coding genes were the next highest in differential expression. These results imply that differing iron concentrations has a large influence on these cellular functions; however, the large number of differentially expressed genes, especially in energy metabolism, could also have been primarily caused by the critical changes in growth—as a result of iron limitation and/or difference in terminal electron acceptors—rather than as a direct result of iron concentrations.

4.2.2 Cross-comparison of differentially expressed genes from "limited"/"high" profile to "limited"/"sufficient": There appeared to be a high degree of overlap between the "limited"/"sufficient" profile and the "limited"/"high" profile, with 76 of the 95 (80%) differentially expressed genes from the "limited"/"sufficient" also being found to be differentially expressed in the "limited"/"high" profile. Since these genes were found in both profiles in which iron limitation was examined, it is likely that many of these genes were differentially expressed as a result of the differing iron concentrations. This

hypothesis is supported by the fact that genes coding for heavy metal efflux pumps from the Czc family (GSU0830, GSU2135, GSU2137, GSU3398, GSU3400), iron transport related proteins (GSU1338, GSU3268, GSU3271), and iron-dependent transcriptional regulators (GSU1379, GSU1382) were all found to be differentially regulated. Additionally, several genes coding for heme-bining c-type cytochromes (GSU0357, GSU0466, GSU1740, GSU2504, GSU2743, GSU2937, GSU3274) and subunits for an enzyme similar to Ech-hydrogenase (GSU0740-0742, GSU0745), which interacts with high iron-containing ferredoxin proteins, were all found to be downregulated in the limited iron condition; however, the differential expression of these genes, which are all involved with electron transport, could also be related to other factors, such as growth limitation. It should also be noted that there were several genes coding for hypothetical proteins (GSU0208, GSU0384, GSU0793, GSU0834, GSU0840, GSU1381, GSU1647, GSU2132-2132, GSU2742, GSU2780, GSU2936, GSU2938, GSU3271-3273), proteins with unknown function (GSU0434, GSU1338, GSU2571, GSU3270), and transport and binding proteins for unknown substrates (GSU0677-0678, GSU2664-GSU2665, GSU2982, GSU3399), which further emphasizes that there is still much that needs to be learned about *G. sulfurreducens*.

4.2.3 Cross-comparison of differentially expressed genes from "limited"/"high" profile to "sufficient"/"high": There was much less of an overlap

in differentially expressed genes in the cross-comparison between "sufficient"/"high" and "limited"/"high," with only 75 out of the 185 differentially (40.5%) expressed genes from the "sufficient"/"high" profile also being found in the "limited"/"high" profile. It would be expected that the two profiles would have had more of an overlap since both profiles had the same variable terminal electron acceptors (fumarate vs. ferric citrate) (please refer to Table 3.3 for differences between the "high," "sufficient," and "limited" conditions). These differences in the congruity of the amount of overlap between the different profiles could be for a number of reasons. One possible reason is that inconsistencies in the "limited" iron data could have resulted in many genes from the "limited"/'high" profile that would have overlapped with genes from the "sufficient"/'high" profile as not being regarded as being differentially expressed as a result of too high of a P-value caused by the inconsistencies in iron concentration in the "limited" condition as previously mentioned. Another possibility is that the iron limitation affected cellular function to the extent that many genes coding for proteins involved with electron transfer to fumarate could not be properly expressed. This hypothesis is given credence by the fact that 95 genes were found to be differentially expressed in the "limited"/"sufficient" profile, in which the only variable was iron levels in fumarate media, and 63 out of these 95 genes (66.3%) were found in the "limited"/"high" profile and not in the "sufficient"/"high profile. These 63 genes, many of which were involved in energy metabolism, could have severely

affected gene expression to the extent that completely different sets of genes were found to be differentially expressed in the "limited"/"high" profile and the "sufficient"/"high" profile.

It is likely that most of the 75 genes that were found to be differentially expressed in both profiles are a result of the difference in terminal electron acceptor rather than changes in iron concentrations since it is assumed that the differences in terminal electron acceptor is a more influential variable on gene expression than the dissimilar iron levels in the "sufficient"/"high" profile. However, it is also likely that at least some of the genes are related to the differing environmental levels of iron. In order to better determine which genes were differentially expressed as a result of the terminal electron acceptors and which genes were differentially expressed as a result of the differing iron levels, a cross-comparison of all three expression profiles was performed. Since the only apparently common variable between all three expression profiles was a difference in iron levels, it would be likely that genes that were differentially expressed in all three profiles are related to changes in iron levels.

4.2.4 Cross-comparison of all three expression profiles: 13 genes were found to overlap and be differentially expressed across all three expression profiles. 11 of these 13 genes were found to be highly differentially expressed (fold change $\geq$ 8, $\log_2$ difference $\geq$ 3) in at least one profile (Table 4.1). All 11

of these genes were found to be upregulated in relation to lower iron levels whereas the other two genes were found to be downregulated.

Although it is unlikely that this cluster of genes encompasses most of the genes that are important for iron response, this overlap across several conditions and the high differential expression implies that this cluster of 11 genes is critically related to iron response—rather than related to growth-related stress response, for example—since differing extracellular iron concentrations was the one variable that was consistent across all three expression profiles. This assumption is supported by the fact that several genes known to be related to iron were differentially expressed across all three profiles, each of which is studied in further detail in the following discussion.

Feo iron transporters: There were several genes from the Feo family of iron transporters that were found to be differentially expressed across all three profiles. For each of the expression profiles, both of the Feo proteins (FeoB-1 and FeoB-2) were upregulated in the condition with a lower concentration of iron.

The "limited"/"high" profile consistently had the highest difference in differential expression, with genes being highly upregulated in the "limited" condition. The "sufficient"/"high" profile, in which both conditions were assumed to have a replete amount of iron, was found to have highly upregulated Feo proteins in the "sufficient" iron condition. The relatively high upregulation of iron transport proteins in the "sufficient" iron condition

compared to the "high" iron condition could either be a result of changing iron needs in each condition or a result of the changing levels of environmental iron. It is possible that the proteins required for electron metabolism when fumarate is used a terminal electron acceptor have a relatively high need for iron compared to when ferric citrate is used as a terminal electron acceptor. For example, all the genes coding for the subunits of a [NiFe]-hydrogenase (GSU0782-GSU0785) were found to be highly upregulated in the "sufficient" condition, in which fumarate is the terminal electron acceptor, compared to the "high" condition, in which ferric citrate is the terminal electron acceptor (Table A2). [NiFe]-hydrogenases are enzymes that catalyze hydrogen oxidation, which can be coupled to the reduction of an electron acceptor like fumarate. Its small subunit contains three iron-sulfur clusters while its large subunit has a nickel-iron center [33]. The relatively high requirement for iron in order to meet the iron needs of [NiFe] hydrogenases and other iron-containing proteins that were preferentially upregulated as a result of fumarate being utilized as the terminal electron acceptor could have led to iron transporters being highly upregulated in the "sufficient" condition compared to the "high" condition.

However, the relatively high upregulation of Feo transporters in the "sufficient" condition compared to the "high" condition could also be a result of the differing iron concentrations between the two conditions. By the time cultures were harvested for gene expression profiling, the "high" condition had about twice the amount of ferrous iron concentration as the "sufficient"

condition. Thus, the "sufficient" condition may not actually be replete with iron, at least compared to the "high" condition, which would result in requiring a relatively high amount of Feo iron transporters in the "sufficient" condition compared to the "high" condition in order to maintain the homeostasis of iron in the cell.

Genes coding for Feo proteins were also upregulated in the "limited" condition compared to the "sufficient" condition, but this degree of upregulation was less than that between the "sufficient" condition and the "high" condition. Thus, assuming that the expression of Feo proteins is highly dependent on intracellular iron concentration, the expression data implies that differences in intracellular iron concentrations between the "limited" and "sufficient" conditions is not as large as the differences in intracellular iron concentrations between the "sufficient" and "high" conditions.

C-type cytochrome: Only one c-type cytochrome (GSU3274) was found in this cluster, and it was the only c-type cytochrome to be highly differential expressed (fold change $\geq$ 8) in any expression profile. Its expression pattern was similar to that of the Feo proteins, implying that the expression of this specific c-type cytochrome is directly related to intracellular iron levels. When looking at the expression of the other genes coding for cytochromes, it appears that the large majority of them were slightly downregulated in relation to decreasing iron levels (Table A2). Thus, it is possible that when iron becomes limited, the genes coding for many other cytochromes are slightly

downregulated as a result of their relatively high iron requirement, and in order to meet their electron transport needs, GSU3274 becomes highly upregulated since it has a lower requirement for iron as a result of only having one heme binding site. A table comparing the fold changes of GSU3274 and a selection of other genes coding for cytochromes whose differential expression values meet the P-value requirement is listed in Table 4.2. The role of this c-type cytochrome will be later discussed in Section 4.3.3.

Iron-dependent transcriptional regulators: Two key regulatory proteins were found to be upregulated in relation to decreasing iron levels: Fur (GSU1379) and the iron/manganese dependent transcriptional regulator (GSU1382: IdeR). Like Fur, IdeR (also known as DtxR) is a transcriptional repressor activated by Fe^{2+} [35]. As such, it is not surprising that the expression of these two Fe^{2+}-dependent regulators appear to be related to iron levels. It appears that these genes have the greatest amount of expression in the "limited" iron condition in which there should be a dearth in intracellular iron levels. This behavior is logical assuming that these genes auto-regulate themselves and that they are only active when they are complexed with iron. Under the "limited" condition, there is a lack of excess in intracellular iron levels, and thus even though these genes are being differentially expressed, their products are inactive as a result of not being bound to any excess iron. Under the "high" condition, when there should be a large amount of excess intracellular iron, IdeR and Fur complex to Fe^{2+} and become active, which

allows for them to auto-repress themselves and thereby decreases the expression of their respective genes. This autoregulatory behavior of Fur would allow for it to be extremely sensitive to changing intracellular iron levels and has been noted in other bacterial species, which will be later discussed in Section 4.4.

Another interesting observation is that the degree of differential expression in "sufficient"/"high" was higher than that of "limited"/"sufficient" for both genes (*fur* and *ideR*). Assuming that the expression of Fur and IdeR are directly related to the amount of intracellular iron, this result implies that the difference in intracellular iron levels is not as large in "limited"/"sufficient" as it is in "sufficient/"high", which further implies that there is a relatively low level of excess iron in the "sufficient" iron condition. This further implies that there is a relatively low level of intracellular iron in the "sufficient" iron condition compared to the "high" iron condition and that much of the differential expression observed in the "sufficient"/"high" condition could be a result of difference in iron condition instead of primarily being a result of the differences in terminal electron acceptors.

Hypothetical and miscellaneous proteins: In addition to iron-related proteins, there was differential expression of a fumarate hydratase, a heat shock protein, an arsenite efflux pump, and several hypothetical proteins in all three expression profiles. It is difficult to determine whether or not the differential expression of the genes is directly related to the differences in iron

levels or are a result of being indirectly related to multiple factors, such as growth limitation and changes in terminal electron acceptors, that appear in the three expression profiles. The differential expression of these genes and the significant number of highly differentially expressed hypothetical proteins further demonstrate that there is much that is unknown about the relationship of iron with *G. sulfurreducens*, and there will be further discussion of these genes and their roles in Section 4.3.3

Table 4.1: List of all 11 genes that are highly differentially expressed across all three expression profiles. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," and 3 = "limited"/"high."

Main role	Gene	Product	DOOR	Fold change		
				1	2	3
Energy metabolism						
	GSU3274	cytochrome c, 1 heme-binding site	30833	34.065	5.43	184.986
Hypothetical proteins						
	GSU1381	conserved hypothetical protein	30486	10.159	4.206	42.735
	GSU3273	conserved hypothetical protein	30883	34.254	5.44	186.351
	GSU3272	hypothetical protein	30883	44.416	4.986	221.504
	GSU3271	carbohydrate-selective porin OprB	GSU3271	84.345	3.174	267.746
Protein fate						
	GSU2678	ATP-independent chaperone, alpha-crystalline/Hsp20 family	GSU2678	7.685	11.288	86.752
Regulatory functions						
	GSU1379	ferric uptake regulation protein Fur	GSU1379	3.856	3.479	13.419
	GSU1382	iron/manganese-dependent transcriptional regulator	GSU1382	10.157	4.397	44.669
Transport and binding proteins						
	GSU3268	ferrous iron transport protein B, putative	30882	43.425	5.384	233.83
	GSU3269	ferrous iron transport protein A authentic frameshift	GSU3269	55.693	4.432	246.882
Unknown function						
	GSU3270	FeoA family protein	GSU3270	31.358	7.514	235.653

Table 4.2: Table of cytochromes. 1= "sufficient"/"high," 2 = "limited"/"high", 3 = "limited"/"sufficient."

Gene	DOOR	Gene product	Fold change		
			1	2	3
GSU3274	30883	cytochrome c, 1 heme-binding site	34.065	5.43	184.986
GSU2937	30809	cytochrome c, 5 heme-binding sites	-1.713	-8.704	-14.907
GSU2495	30717	cytochrome c, 26 heme-binding sites	-3.003	-1.560	-4.685

4.2.5: General conclusions: In general, these observations demonstrate several important notions about gene regulation. First, as can be seen from the "limited"/"high" expression profile, the results of an expression profile testing two different variables will not necessarily equal the results of two separate expression profiles testing each variable individually. Additionally, although an expression profile may show a gene being differentially expressed as a result of a specific variable, that variable may not be the only thing that is causing the observed effect. Conversely, there may be effects from a variable that remain unobserved as a result of other activity occurring within the cell that obscures the effects of the variable specifically being examined. Thus, from these conclusions, it would appear that trying to primarily use expression profiles to determine the effect of a variable, such as an active transcription factor versus an inactive one, would end up being highly inconclusive and convoluted in terms of trying to elucidate significance.

Nonetheless, there was important information garnered from the expression profile comparisons. The significantly different levels of differential expression across all three conditions ("high" vs. "sufficient" vs. "limited" iron) for iron transporters and iron-dependent transcriptional regulators implies that the differential expression of these genes highly differ depending on iron levels, which further implies that iron regulation is critical for *G. sulfurreducens*. It was also seen that many genes coding for proteins that appeared to be iron-dependent, either through function or iron content, did not appear to be

differentially regulated in the "limited" iron condition compared to other iron conditions like expected. This could imply that these proteins are so critical for cellular function that intracellular iron will preferentially be utilized for these proteins in order to maintain their homeostasis. It could also imply that the regulation of these proteins is dependent on something beyond iron concentrations. The changes in the differential expression of *ideR* in different levels of iron demonstrate its iron-dependency and give a potential direction for future studies of iron-related regulation in *G. sulfurreducens.* The relatively large number of hypothetical proteins and proteins with unknown functions that were highly differentially expressed under different levels of iron give another possible direction for future iron-related studies.

4.3 Comparing ChIP-chip data to gene expression data

ChIP-chip data was compared to gene expression in order to further characterize the Fur regulon. 32 out of the 224 genes found to be directly regulated by Fur corresponded to any of the expression profiles. It is unlikely that the limited overlap between the ChIP-chip data and the expression profile data pertains to setting the P-value $\leq$ .05 because even when P-values are not considered, there are only 39 differentially expressed genes that overlap with the 224 genes from the ChIP-chip data (data not shown). The "limited"/"high" expression profile had the largest number of genes that also overlapped in the ChIP-chip profile, with 27 genes. However, the "limited"/"sufficient" profile had

the highest percentage of overlap, with 22.1% (21 genes) overlapping with the ChIP-chip profile. The lower percentage of overlap in the "limited"/"high" profile most likely results from the profile testing two different variables (terminal electron acceptor and iron levels). Meanwhile, the "limited"/"sufficient" profile only varied in iron concentration.

The overlap was similarly poor when comparing genes differentially expressed in both the "limited"/"high" and the "sufficient"/"high" or the "limited"/"sufficient" profiles, with only ~23% of genes in each cross-comparison having Fur binding sites. However, when comparing the 13 differentially expressed genes found in all three profiles to the ChIP-chip data, it was found that ten of them had Fur binding sites.

The following analysis examines a number of questions about Fur, including:

1) Is Fur an activator, a repressor, or both?

2) What could be the reasons for the relatively low level of overlap between ChIP-chip and expression profile data?

3) What are the functions of the genes directly regulated by Fur and what are the possible implications of the role of Fur?

4.3.1 The transcriptional role of Fur: In general, the differentially expressed genes found to correspond with ChIP-chip data were found to be upregulated in lower iron levels (downregulated in higher/normal iron levels),

which supports the role of Fur as being primarily a transcriptional repressor when it is activated with Fe^{2+}. The correspondence of Fur to genes downregulated in lower iron levels initially lends credence to the hypothesis that active Fur could also be an activator of some genes. However, it appears that Fur activates "weakly" in that most of the genes that were found to be downregulated in lower iron conditions had fold changes < 8 whereas most of the genes found to be upregulated in lower iron conditions had fold changes of $\geq$ 8. It is possible that the binding affinity for Fur to the Fur box in the promoter region for the downregulated genes is relatively low, resulting in the gene not being highly repressed by Fur, especially if an activator or sigma factor is present. A previous study of Fur binding in *E. coli* using gel retardation assays demonstrated the different affinities of Fur for different arrangements of hexamer binding sequences [40], thereby showing that Fur does not necessarily strongly bind to all of its binding sites. A diagram clarifying the role of Fur as a "weak" and a "strong" transcriptional repressor can be seen in Figure 4.1.

The hypothesis that the Fur binding sites had different affinities for Fur protein is supported when looking at the correspondence of Fur binding sites to genes highly upregulated (fold change $\geq$ 8) in lower iron levels in each expression profile. There were 25, 14, and 32 highly upregulated genes in the "sufficient"/"high," "limited"/"sufficient," and "limited"/"high" expression profiles, respectively. Out of these highly upregulated genes, 11, seven, and 18 genes

in the "sufficient"/"high," "limited"/"sufficient," and "limited"/"high" expression profiles, respectively, corresponded with the ChIP-chip data, or in other words, ~50% of genes that were highly upregulated in lower iron conditions were found to be directly regulated by Fur according to the ChIP-chip data. Meanwhile, the percentage of other differentially expressed genes' correspondences with ChIP-chip data were often below 10% (Table 4.3). The highly upregulated genes found to correspond to ChIP-chip data may have high affinity Fur binding sites in their promoters, leading to the binding of multiple Fur proteins to their promoters and their differential expression being detectable in expression profiles.

When looking at genes that overlapped in all three expression profiles, it can be seen that ten out of 11 (90.9%) genes that were upregulated in relation to lower iron levels have Fur binding sites. The only upregulated gene in this cluster that did not have a Fur binding site was the iron-dependent transcriptional regulator, IdeR. From this information, it can be estimated that these ten genes are likely to have high affinity Fur binding sites, which would result in their differential expression being highly variable according to iron levels. It is also strongly possible that these genes are coregulated by Fur and IdeR, a hypothesis which is supported when looking at the degree of differential regulation of Fur compared to most of the other genes that overlapped in all three expression profiles (Table 4.1). The topic of coregulation by Fur and IdeR will be discussed in further detail in Section 4.6.3.

Table 4.3: Percentage of differentially expressed genes also corresponding to ChIP-chip data. "Upregulated" and "downregulated" are in reference to the condition with less iron.

Differentially expressed	"sufficient" vs. "high"	"limited" vs. "sufficient"	"limited" vs. "high"
Downregulated, fold change $\geq$ 8	0.0%	28.6%	11.8%
Downregulated, fold change < 8	4.7%	4.2%	8.2%
Upregulated, fold change < 8	4.1%	38.5%	2.2%
Upregulated, fold change $\geq$ 8	44.0%	50.0%	56.3%
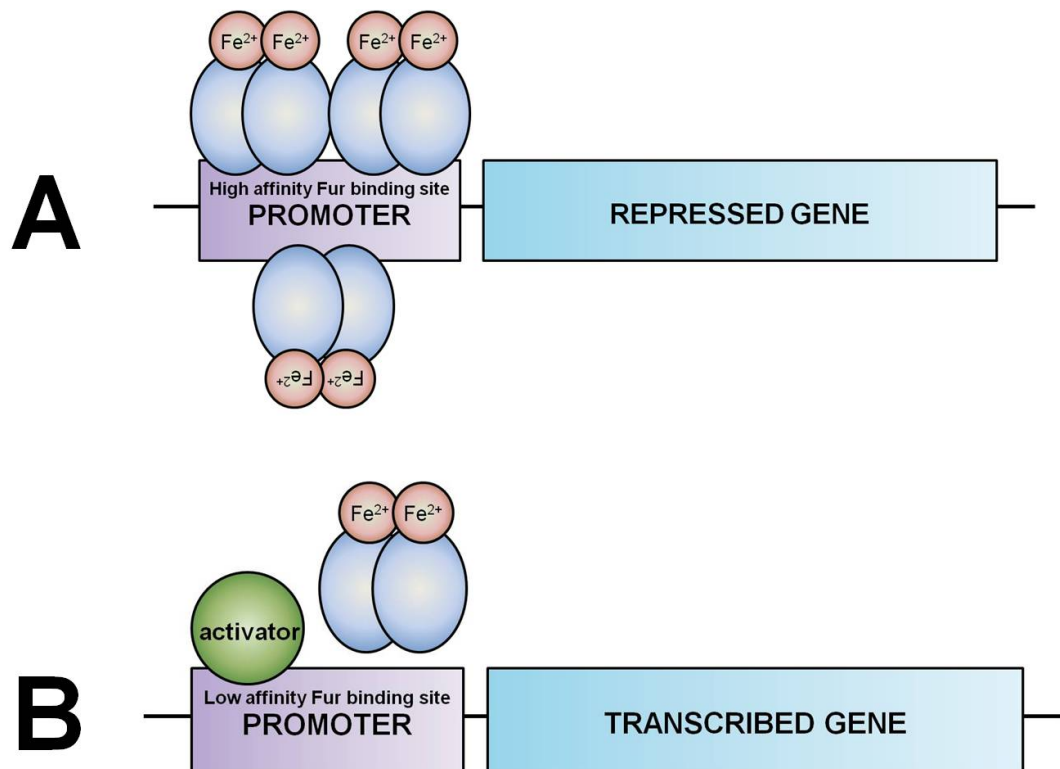
Figure 4.1: Diagram illustrating the transcriptional role of Fur for A) promoters with high affinity Fur binding sites, resulting in "strong" repression by Fur, and B) promoters with low affinity binding site and competition from an activator, resulting in little to no repression by Fur.

4.3.2 Possible reasons for the low level of overlap between data sets: This "low affinity" binding site hypothesis could also explain why there was not as much overlap as expected between the ChIP-chip data and the gene expression data. It could be possible that many of the binding sites found by the ChIP-chip data were lower affinity binding sites and that there was not an optimal concentration of activated Fur proteins to bind to these sites to result in differential expression, especially if there was interference of Fur activity by another transcriptional regulator. Furthermore, it is possible that many of the other genes found in the ChIP-chip data might be differentially expressed under a different iron condition, with a maximal amount of active Fur protein, or under a condition in which a particular transcription factor is not expressed. However, there could also be other possible reasons for the low level of overlap, including:

1) Noise: Noise either in the ChIP-chip data or the expression profile data could have caused some binding sites to be mistaken or caused a gene not to appear differentially expressed.

2) Operon prediction: Some of the DOOR operon predictions may be inaccurate.

These two reasons, however, do not seem like they would explain the large number of genes found in the ChIP-chip binding profile that did not correspond to gene expression data.

4.3.3 Genes overlapping in ChIP-chip and gene expression profiles: Only 23 out of 31 differentially expressed genes found to correspond to ChIP-chip data were upregulated with respect to the lower iron conditions in at least one expression profile. Assuming that active Fur functions only as a transcriptional regulator, the nine genes that were downregulated in the lower iron condition are likely not to have been strongly controlled by Fur in these expression profiles and thus they were not considered to be Fur-regulated in any of the iron conditions tested.

A few of the 23 genes demonstrated unexpected transcriptional behavior. One such gene was a cytochrome b containing subunit for a [NiFe]-hydrogenase (GSU0784). Although cytochrome b contains two heme-binding sites, it is difficult to determine if that is the reason that this gene was considered to have a Fur binding site. This specific subunit was found to be at the end of an operon for other subunits for this hydrogenase, but the Fur binding peak was only found for this specific gene. The other genes in this operon show a similar pattern of differential expression and so it does not appear that the Fur binding played an influence on the differential expression of this gene. However, it is also possible that the operon prediction for this gene is incorrect and that the gene is actually the start of another operon. A similar issue is seen with a potassium-transporting ATPase subunit (GSU2481) and a lipoprotein (GSU2133), which also have difficult to elucidate reasons for being iron-regulated and are also in the middle of an operon of other genes

with a similar pattern of differential expression. In any case, the determination of a Fur binding site in these genes may merit a reannotation of the predicted operon structure for *G. sulfurreducens.*

The 22 genes deemed as Fur-controlled cover a wide range of cellular functions, coding for proteins with roles in the following functions: cell envelope, cellular processes, energy metabolism, protein fate, regulatory functions, and transport and binding (Table 4.4). Many of the genes had direct ties to iron, such as the various Feo family of iron transporters (discussed in section 4.1.1 and section 4.2.4) and Fur. There were also genes coding for proteins with more indirect ties to iron, such as a heme-binding c-type cytochrome (discussed in section 4.2.4) and the Czc family of heavy metal efflux proteins (discussed in section 4.1.1). Other genes coded for proteins with little to no ties to iron. In addition to the genes mentioned in the previous paragraph, genes coding for a putative lipoprotein (GSU0832) and a putative nitrogen regulatory protein (GSU0831) were also found to be regulated by Fur. However, it can be seen that both of these genes are in the same operon as the genes coding for several efflux proteins from the Czc family (GSU0828-GSU0830) and show a similar pattern of regulation as the Czc family genes.

The Fur binding site on an alpha-crystallin heat shock protein (GSU2678) initially observed when examining the overlap between expression profiles confirms that this protein is iron and Fur-regulated. Alpha-crystallins bind to unfolding intermediates to protect them from irreversible aggregation.

Typically, they are recognized as heat shock proteins since they are induced by temperature shift [41]. The strong regulation of the alpha-crystallin protein could be critical in stress response resulting from iron limitation. Perhaps the stress of iron limitation is so high that protein structure integrity is threatened, necessitating a chaperone protein to reduce the damage caused. The strong regulation could also be further evidence of Fur's transcriptional regulation being related to heat shock response.

A single heme-binding c-type cytochrome (GSU3274) that was initially observed when examining the overlap between expression profiles was also found in the ChIP-chip profile, confirming that this cytochrome is iron and Fur-regulated. In poster recently presented by the Lovely group for a currently unpublished study [56], it was shown that the deletion of GSU3274 completely inhibited the ability of *G. sulfurreducens* to transfer electrons from electrodes. It was further proposed that this cytochrome acts as an intermediary between the outer cell surface and the inner membrane for electron transfer. If this proposed role is valid, then Fur could play a critical role in regulating electron transfer in *G. sulfurreducens.*

The relatively high degree in which the hypothetical proteins are differentially expressed and their direct regulation by Fur indicate that they play critical roles in cellular response to iron. Future investigations into the functions of these proteins could be of great interest in terms of elucidating the relationship of iron to *G. sulfurreducens.*

Table 4.4: List of all 22 genes that were considered Fur-regulated in at least one of the conditions tested. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," and 3 = "limited"/"high." Only the fold changes that had a P-value $\leq$ .05 are listed.

Main role	Gene	Gene name	Door	Gene product	Fold change		
					1	2	3
Cell envelope (2)							
	GSU0832		30370	lipoprotein, putative		21.61	22.32
	GSU2133		30645	lipoprotein, putative		18.44	23.58
Cellular processes (2)							
	GSU0828		30370	metal ion efflux pump, RND family, outer membrane protein		16.73	16.78
	GSU0829		30370	efflux pump, CzcB family, membrane fusion protein		20.31	20.80
Energy metabolism (2)							
	GSU0784		30359	periplasmically oriented, membrane bound [NiFe]-hydrogenase integral membrane subunit	34.90		
	GSU3274		30883	cytochrome c, 1 heme-binding site	34.07	5.43	184.99
Hypothetical proteins (6)							
	GSU1381		30486	conserved hypothetical protein	10.16	4.21	42.74
	GSU3267		30882	conserved hypothetical protein	26.25		110.04
	GSU3273		30883	conserved hypothetical protein	34.25	5.44	186.35
	GSU3272		30883	hypothetical protein	44.42	4.99	221.50
	GSU3271		GSU3271	carbohydrate-selective porin OprB	84.35	3.17	267.75
	GSU1060		GSU1060	conserved hypothetical protein			5.15
Protein fate (1)							
	GSU2678		GSU2678	ATP-independent chaperone, alpha-crystallin/Hsp20 family	7.69	11.29	86.75
Regulatory functions (2)							
	GSU1379	fur	GSU1379	ferric uptake regulation protein Fur	3.86	3.48	13.42
	GSU0831		30370	nitrogen regulatory protein P-II, putative		21.26	21.63
Transport and binding proteins (5)							
	GSU3268	feoB-2	30882	ferrous iron transport protein B, putative	43.43	5.38	233.83

Table 4.4, cont.: List of all 22 genes that were considered Fur-regulated in at least one of the conditions tested. 1= "sufficient"/"high," 2 = "limited"/"sufficient," and 3 = "limited"/"high." Only the fold changes that had a P-value $\leq$ .05 are listed.

Transport and binding proteins, cont. (5)

GSU3269		GSU3269	ferrous iron transport protein A	55.69	4.43	246.88
GSU0830		30370	efflux pump, CzcA family, inner membrane protein		17.33	17.18
GSU1380	feoB-1	30486	ferrous iron transport protein B	10.87		48.61
GSU2481	kdpB	30714	potassium-transporting ATPase, B subunit	2.32		

Unknown function (2)

GSU3270		GSU3270	FeoA family protein	31.36	7.51	235.65
GSU1639		GSU1639	transcriptional regulator, putative		3.11	

4.4 Overview of the Fur regulon in G. sulfurreducens

This study confirmed several characteristics about Fur in *G. sulfurreducens.* First of all, Fur was confirmed—via ChIP-chip experimentation of the "limited" iron condition—to be active only when bound to Fe^{2+}. Fur was also verified to have iron-responsive auto-regulation, allowing its regulatory behavior to be highly sensitive to changing iron levels. A diagram elucidating this auto-regulatory behavior can be seen in Figure 4.2 Fur was furthermore confirmed to be the primary transcriptional regulator for genes highly dependent on iron in *G. sulfurreducens*, such as iron transport proteins. Gene expression analysis was performed on three different iron conditions, and 11 genes were found to be upregulated in higher iron concentrations across all three expression profiles ("limited"/"sufficient," "sufficient"/"high," and "limited"/"high"). Ten out of 11 of these genes were found to be directly regulated by Fur according to ChIP-chip analysis. The only gene in this cluster

not found to be regulated by Fur was the IdeR transcriptional regulator. However, not all genes in this cluster had direct ties to iron, such as a gene coding for an alpha-crystallin protein, which is normally associated with heat shock. The regulation of these genes is likely to be critical for growth, which can be seen in the hindered growth of the Fur knock-out mutant in the "sufficient" iron condition. Fur was additionally found to have directly regulated 12 other genes according to ChIP-chip analysis. These 12 other genes have a wide range of functions and thus further demonstrate that Fur's regulation extends beyond iron uptake and transport. A diagram of the Fur regulon in *G. sulfurreducens* is illustrated in Figure 4.3.

In total, when comparing ChIP-chip data to gene expression data, 19 genes were found to be part of the Fur regulon. However, according to only the ChIP-chip data, which tested for the direct binding of Fur to a DNA sequence, there 148 binding sites for Fur and a total of 220 genes directly regulated by Fur. It is unlikely that the large number of Fur binding sites found by ChIP-chip is a result of the Fur antibody binding to other proteins besides Fur since no consistent binding sites were found when replicating the ChIP-chip experiments in "limited" iron or Fur knock-out mutant conditions. The lack of differential expression of the majority of the 220 genes in any of the expression profiles could be a result of a number of things, such as incorrect curation, incorrect operon prediction, or noise from the microarray data. However, it is also possible that not all Fur binding sites have an equal amount

of affinity for Fur protein. A lower affinity binding site in addition to the presence of another transcriptional regulator, such as an activator or a sigma factor, could result in genes not being recognized as differentially expressed under certain iron conditions. If this hypothesis is true, then it further demonstrates that it is difficult to gain a large amount of information about the direct regulation of a transcription factor from gene expression profiling.
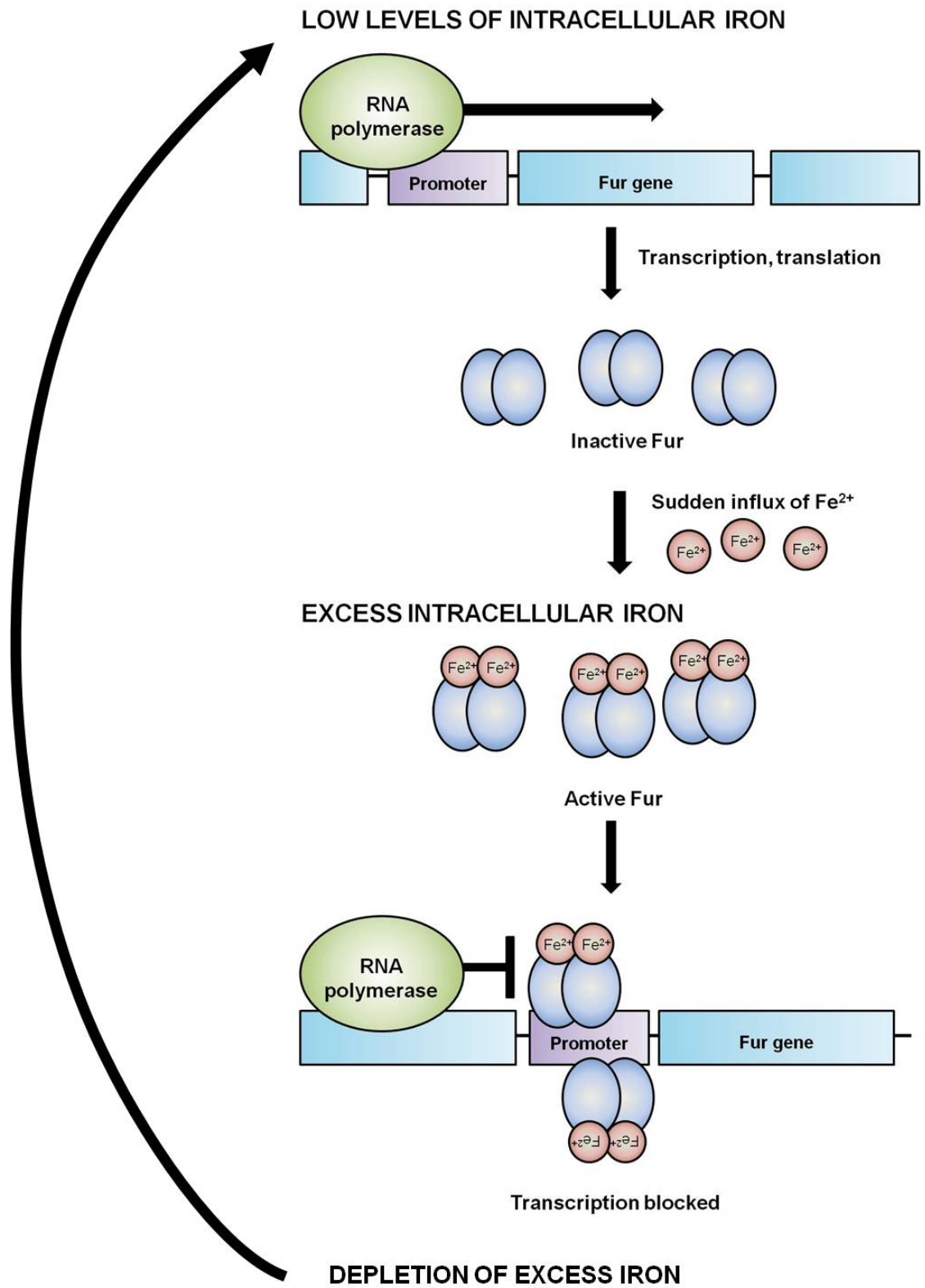
Figure 4.2: Illustration of Fur's auto-regulatory behavior.

INTRACELLULAR FE²⁺ CONCENTRATION

LOWER   HIGHER

De-repression   Repression

DIRECT REGULATION

HIGHLY REGULATED Feo family of proteins Fur Hypothetical proteins Cytochrome c (GSU3274) Alpha-crystallin heat shock protein	**OFTEN REGULATED** Heavy metal efflux proteins (GSU0828-GSU0830) Lipoproteins (GSU0832, GSU2133) Nitrogen regulatory protein (GSU0831)	**CONDITIONALLY REGULATED** Other 200 genes found in ChIP-chip

SECONDARY, INDIRECT REGULATION

Figure 4.3: Diagram of the Fur regulon in *G. sulfurreducens.*

4.5 Comparison to other Fur regulons

Since its initial discovery in 1981 [43], Fur has been studied in a variety of bacteria. With the advent of advances in microarray technology, investigations of Fur have become even more widespread. Thus, when looking at Fur regulons for comparative studies, there are dozens of options. Typically, these studies involve creating a Fur knock-out mutant and comparing its expression profile in iron-replete and iron-depleted conditions to the wild-type, and typically, these studies have focused on aerobic bacteria since iron is

limited in oxic conditions. However, there are a few studies are prominent, either through their methods or through the nature of the bacteria that they studied.

4.5.1 G. sulfurreducens: As mentioned before, the Fur regulon has been previously studied in *G. sulfurreducens* in an unpublished study performed by the Lovely group at the University of Massachusetts, Amherst [23]. This study created a Fur knock-out mutant, which was later used for this study, and analyzed gene expression profiles of the Fur knock-out mutant and wild-type strains of *G. sulfurreducens* in iron-replete and iron-limited conditions. In total, 50 differentially expressed genes were found to overlap in the Fur knock-out and the wild-type strain under iron-limited conditions, and for the sake of comparison to our study, these 50 genes could be considered as the Fur regulon.

10 out of 50 of the genes identified by the Amherst study were found in the Fur binding profile as determined by ChIP-chip in our study. One of these genes, GSU0915, which codes for a hypothetical protein, had not been identified as differentially expressed in any of the expression profiles as a result of too high of P-values. Another one of these genes, GSU2133, which does for a putative lipoprotein, had been disregarded as Fur-regulated in this study as a result of it being in the middle of an operon of genes with similar levels of differential expression. However, the correspondence of these genes

to the genes determined to be in the Fur regulon according to the Amherst study lends credence to the possibility that these genes are indeed Fur-regulated.

A search for a putative *G. sulfurreducens* putative Fur box motif was also performed in the Amherst study, using three different sources: 1) an *E. coli* motif [48], 2) a δ-*proteobacterial* consensus sequence [49], and 3) Fur boxes in *G. sulfurreducens* identified in genes that were co-regulated in response to the elimination of the regulator for stringent response, Rel_{Gsu} [50]. The correspondence of putative Fur boxes in front of genes found in the overlapping microarray data was relatively poor, with 2%, 35% (out of 20 putative Fur boxes in total), and ~15% of the putative Fur boxes identified from using the *E. coli*, δ-*proteobacterial*, and Rel_{Gsu} mutant, respectively, being found. The lack of success in identifying a *G. sulfurreducens* Fur box opens the door for future Fur binding motif studies using the data procured from this study.

Although there are many inconsistencies between the two studies on Fur in *G. sulfurreducens,* one especially notable inconsistency is considering the transcriptional regulation of IdeR and Fur. The Amherst study found that the *ideR* gene was differentially expressed in their microarray data sets whereas the *fur* gene was not, implying that IdeR is Fur-regulated and that Fur does not participate in auto-regulation. Our investigation of Fur found the exact

opposite. A future study on the transcriptional role of IdeR could shed some light on this contradiction.

4.5.2 Other metal-reducing anaerobes: Although the Fur regulon has been primarily study in aerobic bacteria, it has been previously studied in two other metal-reducing anaerobes: *Shewanella oneidensis* [51-53] and *Desulfovibro vulgaris* Hildenborough [54]. All four studies primarily utilized gene expression profiling via microarray analysis of Fur knock-out mutant and wild-type strains to characterize the regulation of Fur.

Shewanella oneidensis: S. oneidensis is a γ-proteobacterium often found in deep sea anaerobic environments, but can also subsist in soil habitats. In addition to having the ability to reduce a variety of electron acceptors, like fumarate and Fe^{3+}, *S. oneidensis* is similar to *G.* sulfurreducens in that it has been found to possess a high number of cytochromes [53]. These similarities in energy metabolism and high assimilatory needs for iron could result in similar regulation of iron uptake by Fur.

In the 2002 study of the role of Fur in *S. oneidensis* [51], the mutant and wild-type strains were grown in aerobic and anaerobic conditions and differentially expressed genes were determined. 11 genes reproducibly were highly upregulated (fold-change $\geq$ 3), in the Fur knock-out mutant compared to the wild-type. Most of these genes were involved with iron acquisition and utilization as well as energy metabolism, and none of these genes were found

to have a corresponding homolog within the Fur binding profile from this study according to a BLAST search.

The 2004 study of Fur in *S. oneidensis* [52] was almost exactly like the first study, except by that time, the complete genome of *S. oneidensis* had been sequenced. Gene expression data was coupled with a Fur box search on genes demonstrating a five-fold change in differential expression between wild-type and mutant strains, and it was suggested from this motif search that Fur directly controls the transcription of 39 genes, with 19 of these genes coding for hypothetical proteins. Like Fur for *G. sulfurreducens,* the Fur regulon determined in this study of *S. oneidensis* is thought to control genes in a wide range of cellular functions, including cellular processes, energy metabolism, and protein fate. The large majority of genes, however, were for transport and binding proteins, specifically those related to iron. A search for homologs of the proteins for *S. oneidensis* with BLAST did not turn up any *G. sulfurreducens* homologs with Fur binding sites.

Fur was again studied in *S. oneidensis* in 2008 [53], but this time in terms of its role in acid tolerance response in addition to iron uptake. As such, the wild-type and the Fur knock-out mutant strains were tested in iron-replete and iron-limited conditions and also tested in media buffered at pH 5.5 and pH 7. The knock-out mutant demonstrated much more sensitivity to the acidic condition compared to the wild-type. However, transcriptome analysis was only performed in iron-replete and iron-limited conditions, and this analysis did

not reveal any known acid resistance systems as regulated by Fur. It would have been perhaps more interesting if acidic versus non-acidic conditions in different iron conditions for the wild-type and knock-out strains had been tested. Although such a study would have been more difficult to analyze, especially with primarily using gene expression profiles, it could have potentially revealed the transcriptional role of Fur as being dependent on other factors other than intracellular iron levels and/or different sets of genes appearing to be regulated by Fur under different environmental conditions as a result of the presence or absence of other transcriptional regulators.

Desulfovibro vulgaris Hildenborough: D. vulgaris Hildenborough, like *G. sulfurreducens,* is an anaerobic, sulfate-reducing δ-proteobacterium known for its ability to reduce heavy metals. It also contains a high assimilatory requirement for iron since it contains a large number of iron-containing cytochromes, hydrogenases, and electron transport proteins [54]. In addition to metabolically similar to *G. sulfurreducens*, the Fur protein for *D. vulgaris* Hildenborough has a 50% identity with that for *G. sulfurreducens* according to a BLAST search. Furthermore, like *G. sulfurreducens,* even though *D. vulgaris* contains genes that are typically thought to be involved in siderophore uptake, no genes for siderophore production have been found. As a result of all of these similarities, the roles of the two Fur proteins in the two different species of bacteria could regulate homolog genes.

In the study of the role of Fur in *D. vulgaris* Hildenborough [54], gene expression profiles for a Fur-deletion mutant and wild-type grown in iron-replete and iron-limited conditions were compared. Analysis revealed that gene expression with *fur* deleted affected 12 functional categories. 13 genes were differentially expressed in both iron-replete and iron-limited condition in response to the *fur* deletion. A putative binding motif was then determined via computational analysis of the promoter regions of the *feoAB* operons. In total, the Fur regulon was thought to encompass 22 genes in total. In addition to this transcriptome analysis, the phenotype of the *fur* deletion mutant was investigated. The *fur* deletion mutant was seen to be more sensitive than the wild-type in the presence of nitrite, but was observed to not be any more growth-limited in iron-replete conditions than the wild-type.

Like *G. sulfurreducens*, proteins in the *Feo* family were found to be apparently regulated by Fur bound to Fe^{2+} and genes thought to be involved in iron storage did not appear to be Fur regulated. Additionally, genes coding for several ABC transporters and TonB-related proteins were also found to be a part of the Fur regulon in *D. vulgaris* Hildenbourough, despite it not having any identified siderophore genes. These types of genes were also found in *G. sulfurrences*' Fur binding profile; however, the *G. sulfurreducens* and *D. vulgaris* Hildenborough ABC transporters and TonB-related proteins did not appear to have any significant homology according to a BLAST search. In *G. sulfurreducens*, the genes did not appear differentially expressed whereas in *D.*

vulgaris Hildenborough, the genes were differentially expressed in the *fur* mutant regardless of iron transportation. The apparent regulation of these genes by Fur implies that ABC transporters and TonB-related proteins play a role in iron transport in these anaerobic bacteria. An investigation of these genes could illuminate new features in both species' iron uptake systems, possibly through the discovery of a modified mechanism involving these genes and Fe^{2+} uptake or through the discovery of a currently unbeknownst genes coding for siderophores. Furthermore, a BLAST search revealed that a hypothetical protein (DSU0304) was found to have a Fur-controlled homolog (50% identity) in *G. sulfurreducens* (GSU3273). In *G. sulfurreducens*, GSU3273 is found to be in a highly iron-dependent operon, containing another hypothetical protein and a c-type cytochrome with one binding site. The appearance of both homologs in Fur regulons supports the hypothesis that these hypothetical proteins exist and are somehow involved with iron response.

Interestingly, the growth of the two *D. vulgaris* Hildenborough strains was identical in iron-replete conditions, indicating that Fur deletion was not a limiting growth factor. There was also a lack of evidence for the auto-regulation of Fur. In addition, a BLAST search of the other hypothetical proteins that were predicted to be in the Fur regulon of *D. vulgaris* Hildenborough did not have a Fur-regulated *G. sulfurreducens* counterpart. These dissimilarities emphasize that although these two species of bacteria

are similar in many ways, there are still a large number of differences between the transcriptional regulation of their respective Fur proteins.

4.5.3 Other studies of interest: A few other studies are of interest as a result of their method of determining the Fur regulon of their respective species. Although they heavily relied on the use of gene expression profiling, they coupled their gene expression profiling with other global techniques to better investigate the direct function of Fur.

Neisseria gonorrhoreae: A recently published study investigating the role of Fur in *Neisseria gonorrhoeae* used three global strategies: gene expression profiling, computational analysis to determine a conserved Fur box sequence, and Fur titration assays (FURTA) to detect genomic regions able to bind Fur *in vivo* [50]. The gene expression profiling was performed by comparing wild-type *N. gonorrhoreae* grown under iron-depleted conditions to wild-type *N. gonorrhoreae* grown under iron-replete conditions for one, 2, three, and four hours, resulting in gene expression data for a variety of time points. Computational analysis was performed using Fur box seed sequences from organisms with Fur proteins highly similar to *N. gonorrhoreae's.* FURTA was performed by titrating Fur protein on a *N. gonorrhoreae* genomic library, and FURTA-positive clones were screened.

FURTA is a method of interest since, like ChIP-chip, it examines which genes that Fur directly controls. The study admitted that FURTA likely had a

limited sensitivity in detecting Fur binding. In the original FURTA study, the binding of a gene coding for a transferrin-binding protein (NGO1495) with two predicted intragenic Fur boxes had not been detected. A plasmid was then transformed with the sequence including both Fur boxes, and the transformation gave a weakly positive result by the FURTA. It was hypothesized that these Fur boxes had a lower affinity binding for Fur.

Out of the 28 FURTA-positive clones that were found, 24 had been found in *in silico* prediction, and the remaining four FURTA-positive clones were found in intragenic regions of the genome. Out of these 24 FURTA-positive clones, 14 were iron derepressed and ten were iron induced. The ten genes that were iron induced included genes coding for: a protein involved in oxidative phosphorylation (NGO1751), a transposase (NGO1317), an ATP binding protein (NGO2116), a Fe-S oxioreductase (NGO0904), an alcohol dehydrogenase (NGO0711), a putative phosphatase (NGO0076), two hypothetical proteins (NGO0432 and NGO1430), and *norB* (NGO1275), which is a putative protein involved with anaerobic growth. NorB is of interest since it had previously been shown that gonoccocal Fur indirectly activates NorB by preventing the binding of another repressor, whose binding region overlapped with a Fur box binding site located upstream of the promoter. It is possible that there could be Fur binding sites with a similar role in *G. sulfurreducens.*

Heilcobacter pylori: Fur has been widely studied in *H. pylori* as a result of metal ion-dependent regulators being involved in a regulatory cascade with

the acid resistance regulator (HP0166, *arsR*), which in turns controls many of the genes directly involved with gastric colonization, a critical aspect for *H. pylori* infection. As mentioned before, Fur has also been of interest in *H. pylori* since it is the only known bacteria with a Fur protein that can regulate genes when it is in its apo-Fur form. In the special apo-Fur case, Fe2+ acts as an inducer instead of a co-repressor [51].

One study of Fur in *H. pylori* is especially of interest because of its use of gene expression profiling coupled with ChIP-chip to determine direct regulation of transcription by Fur [52]. In that investigation of Fur, triplicate Fur-IP and *fur* knock-out mutant control-IP experiments (from both iron-replete and iron-chelated) cultures were performed. Iron-dependent enrichment was calculated by dividing IP values obtained after iron treatment by those obtained after chelation [(Fur-IP Fe+)/(Fur-IP Fe-)]. In total, 200 binding sites were defined as Fur targeted, and 119 of these loci were found in the promoter region. Included in these targets were genes that code for transcription factors, which parallels some of the binding sites found in the *G. sulfurreducens* binding profile. The *H. pylori* study similarly stated that if these sites were Fur-regulated, many genes that may be found to be deregulated by inactive Fur may be a result of the Fur regulation of transcription factors rather than a direct result of Fur.

Gene expression profiles comparing the *fur* knock-out mutant to the wild-type were created in order to compare the binding profile to differential

expression. The Fur regulon was defined as the 59 genes that were derepressed in the *fur* knock-out mutant were also found in the ChIP-chip binding profile. Out of these 59 genes, 34 genes, including *fur*, were repressed in iron-replete conditions whereas 25 genes were upregulated.

As in this study of Fur in *G. sulfurreducens*, there were a large number of binding sites that did not correspond to differential expression for *H. pylori*. The authors proposed that these binding sites —may represent fortuitous false positives [51]‖ or, as had been previously been proposed in a study of LexA in *E. coli* [53], these targets could be subject to cooperative regulation with other regulatory proteins or unusual DNA conformation. As speculated in this study of *G. sulfurreducens*, the binding of a transcription factor to an "unconventional [58]" site may affect transcription of adjacent genes, but only under specific conditions related to another regulatory factor.

4.6 Future directions

This study opened the door for future studies that could be of interest when determining the role of Fur and the relationship of iron uptake in *G. sulfurreducens.*

4.6.1 Determine the Fur binding motif: Determining the Fur binding motif for *G. sulfurreducens* could be of interest, especially since it could possibly validate the genes found in the ChIP-chip binding profile that were not differentially expressed. The Fur binding site has typically been described as a

consensus sequence ~19 bp long. In *E. coli,* this binding site had been typically described as a palindromic sequence composed of two 9-bp inverted repeats [20]. However, more recently, this binding site has been proposed to be a series of three adjacent hexamers, with the third hexamer not being an exact duplicate of the previous two (an F-F-F configuration), or as two hexamers separated by one base pair from a third hexamer in reverse orientation (an F-F-x-R configuration) [architecture paper]. In a 2003 study of the Fur binding motif in *E. coli*, Fur had been shown to bind to a wide range of architectures, with different degrees of affinity. The highest degree of affinity, however, was found with the F-F-x-R configuration. Nonetheless, the fact that Fur can bind to slightly different motifs with different affinities could make it difficult to ascertain a binding motif, especially none have been successfully been elucidated in the *Geobacter* species or, as far as it is known, in any other closely related bacterial species.

When attempting to determine a Fur binding motif, it would most likely be easiest to start with the ten genes which were both found in the ChIP-chip profile and had a high degree of differential expression in relation to iron concentration since these genes are likely to have high affinity Fur binding sites. For these genes, there were three Fur binding sites that had been determined, and these three sites could be used to find a general, initial motif. From there, the other 13 genes which were found in the Fur regulon, and in which there were another seven binding sites, could be analyzed in order to

further refine the motif. Then, if it were possible, it could be seen if this motif, or small variations of the motif, could be applied to the other genes found in the ChIP-chip profile.

4.6.2 Determine differentially regulated genes in other iron conditions: There could be interest in performing more expression profiles in order to determine the conditions in which specific genes with Fur binding sites are differentially regulated. Some conditions that could be looked at could include:

1) A condition, like in the *N. gonohorreae* study [55], in which a culture is harvested at various time points after ferrous iron has been titrated into the culture. In such a condition, there would be a relatively large number of active Fur protein, at least compared to a constant "high" iron culture, which may lead to more differential expression as a result of regulation by Fur.

2) The *fur* deletion mutant in the *D. vulgaris* Hildenborough study [54] demonstrated nitrite sensitivity. Since *D. vulgaris* Hildenborough and *G. sulfurreducens* are metabolically similar in many ways and have relative similarity in Fur protein, testing the effects of nitrite on Fur regulation could be of interest.

3) Heat shock or growth on electrode conditions were implicated in this study as being Fur-regulated and thus could also be of interest to study.

4.6.3: Determine the function of hypothetical proteins found to be highly iron-dependent: A large number of genes coding for hypothetical proteins were found in the ChIP-chip binding profile, including four which were highly differentially regulated in the "limited"/"high" iron expression profile. It is thus likely that these genes are somehow largely involved with iron response, and the determination of the function of these proteins could elucidate more about the relationship between iron and *G. sulfurreducens.*

4.6.4 Comparison study of IdeR: A study of IdeR's transcriptional role could be of interest when trying to elucidate the function of role of Fur in *G. sulfurreducens.* As mentioned before, IdeR is another transcriptional repressor that is activated by excess intracellular Fe^{2+}. In this study, it was shown that, like Fur, IdeR was differentially regulated across all three iron conditions that were tested. It would be interesting for the sake of comparison and in order to better elucidate the iron-dependent regulatory circuit in *G. sulfurreducens* to determine which genes IdeR directly controls.

Furthermore, it is possible that the two iron-dependent regulators could cooperatively regulate highly iron-dependent genes in some instances. The idea of cooperative regulation between the two iron-dependent regulators could help explain why the differential expression of the *fur* gene was relatively low compared to the majority of the other genes that were found to be differentially expressed across all three iron conditions in this study. In addition,

the cooperative behavior of Fur with that of another transcriptional factor has previously been documented. In *H. pylori*, Fur and the nickel-responsive regulator, NikR, have been found to have coordinated regulation of several genes via overlapping binding motifs within the promoters [51].

4.7 Conclusion

The critical role of Fur in *G. sulfurreducens* can be simply observed in the limitations in growth caused by knocking out the Fur gene *G. sulfurreducens*, and this study validated its importance. It was found that Fur may play a role in the transcriptional regulation of up to 224 genes. In terms of what was confirmed by both ChIP-chip and gene expression analyses, Fur was confirmed to be the primary transcriptional regulator for iron transport related genes and was also found to directly regulate other genes that span a wide range of cellular functions beyond iron transport, including genes coding for an alpha-crystallin protein and a lipoprotein. Furthermore, Fur's transcriptional role was validated as being an auto-regulated repressor that is activated when bound to Fe^{2+}.

The final results of the study bring up interesting issues beyond the role of Fur in *G. sulfurreducens.* The sophistication involved with transcriptional regulation was illustrated within this study. In addition to many apparent Fur binding sites no affecting the expression of genes, there was some evidence given of possible coordinated regulation of Fur with another iron-dependent

repressor, IdeR. Additionally, the hypothesis of different affinity binding sites opens up a world of complexity when considering the possibilities of competitive binding and multiple variables beyond a given variable, such as iron levels for Fur, that could affect the regulation of a single transcription factor. Furthermore, the lack of correspondence between ChIP-chip data and gene expression data, in addition to the inconsistencies between the unpublished Amherst study and this study, begs the question of which method—gene expression profiling or ChIP-chip—is more accurate in terms of trying to determine the genome-wide role of a transcriptional regulator. It has been commonplace for almost a decade to rely heavily on the use of gene expression profiling in terms of determining a regulon. Moreover, if transcriptional regulation is as complicated as was suggested by this study, then such complexity would be nearly impossible to ascertain using primarily expression profiling. As was seen in this study, it is difficult to pinpoint the causes of the differential expression that is observed in a gene expression profile even when trying to test for a specific variable. Nonetheless, ChIP-chip is not without its own issues, such as the need to manually curate data and the difficulty of optimizing experiments such that noise is kept to a minimum. It will be interesting to see if these issues arise in the future for others who attempt to employ both methods to define a regulon, and if so, how they resolve them. We may find in the future that many previous studies, which primarily

depended on expression profiling to determine the role of a transcriptional regulator, may be invalidated.

APPENDIX

Table A1: List of all genes controlled by Fur in *G. sulfurreducens* according to ChIP-chip data.

Gene	Gene name	Strand	Door	Gene product	Main role
GSU0008		+	30195	response receiver histidine kinase (REC, PAS, HisKA, HATPase_c)	Signal transduction
GSU0215	folD-1	+	GSU0215	5,10-methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase	Central intermediary metabolism
GSU0224		-	GSU0224	conserved hypothetical protein	Hypothetical proteins
GSU0226		+	30241	membrane protein, putative	Transport and binding proteins
GSU0288		-	GSU0288	conserved hypothetical protein	Unknown function
GSU0289		+	GSU0289	conserved hypothetical protein	Hypothetical proteins
GSU0311		-	GSU0311		No Data
GSU0312		+	30258	PilZ domain protein	No Data
GSU0313		+	30258	DnaJ domain protein	Unknown function
GSU0341	nuoD	+	30265	NADH dehydrogenase I, D subunit	Energy metabolism
GSU0420	fliL	+	30281	flagellar protein FliL	Cellular processes
GSU0421	fliM	+	30281	flagellar motor switch protein FliM	Cellular processes
GSU0422	fliN	+	30281	flagellar motor switch protein FliN	Cellular processes
GSU0499		-	GSU0499	zinc metallopeptidase, M23/M37 family	Unknown function
GSU0501		+	30298	lipoprotein, putative	Cell envelope
GSU0502		+	30298	lipoprotein, putative	Cell envelope
GSU0503	crcB	+	30298	predicted chromosome condensation membrane protein CrcB	Unknown function
GSU0555		+	30310	ISGsu7, transposase OrfA	Mobile and extrachromosomal element functions
GSU0569		+	30311	nicotinamidase-related cysteine hydrolase	Unknown function
GSU0655	rpoH	-	GSU0655	RNA polymerase sigma-32 factor	Transcription
GSU0656	ilvE	+	30329	branched-chain amino acid aminotransferase	Amino acid biosynthesis
GSU0657		+	30329	conserved hypothetical protein	No Data
GSU0680		+	GSU0680	DUF748 repeat protein	Hypothetical proteins
GSU0729		-	30349	iron-sulfur cluster-binding protein	Energy metabolism
GSU0731	lpxC	+	GSU0731	UDP-3-0-acyl N-acetylglucosamine deacetylase	Cell envelope
GSU0784	hybB	+	30359	periplasmically oriented, membrane bound [NiFe]-hydrogenase integral membrane subunit	Energy metabolism
GSU0822		-	GSU0822	sensor histidine kinase (Cache, HAMP, PAS, PAC, HisKA, HATPase_c)	Signal transduction
GSU0823		+	30369	PpiC-type peptidyl-prolyl cis-trans isomerase	Protein fate
GSU0824		+	30369	conserved hypothetical protein	Hypothetical proteins

Table A1, cont.: List of all genes controlled by Fur in *G. sulfurreducens* according to ChIP-chip data.

GSU0828		+	30370	efflux pump, RND family, outer membrane protein	Cellular processes
GSU0829		+	30370	efflux pump, RND family, membrane fusion protein	Cellular processes
GSU0830		+	30370	efflux pump, RND family, inner membrane protein	Transport and binding proteins
GSU0831		+	30370	nitrogen regulatory protein P-II, putative	Regulatory functions
GSU0832		+	30370	lipoprotein, putative	Cell envelope
GSU0847		-	30375	rubredoxin	Energy metabolism
GSU0848	frx-5	-	30375	ferredoxin family protein	Energy metabolism
GSU0850		+	30376	protein disulfide bond isomerase, DsbC/DsbG-like	No Data
GSU0851		+	30376	rhomboid-related membrane protein	Cell envelope
GSU0856	htpX	-	GSU0856	membrane-bound Zn-dependent protease HtpX	Protein fate
GSU0857		+		membrane protein, putative	Cell envelope
GSU0915		+	GSU0915	conserved hypothetical protein	No Data
GSU0938		+	GSU0938	conserved hypothetical protein	No Data
GSU1020		+	30407		No Data
GSU1025		-	GSU1025	conserved domain protein	Hypothetical proteins
GSU1026		-	30409	conserved hypothetical protein	Hypothetical proteins
GSU1060		-	GSU1060	conserved hypothetical protein	No Data
GSU1097	pstA	-	30426	phosphate ABC transporter, membrane protein	Transport and binding proteins
GSU1101		-	GSU1101	sensor histidine kinase (HAMP, PAS, HisKA, HATPase_c)	Regulatory functions
GSU1103		+	30427	AMP-forming acyl-CoA synthetase	Fatty acid and phospholipid metabolism
GSU1104		+	30427	conserved hypothetical protein	No Data
GSU1105	pepQ-1	+	30427	Xaa-Pro dipeptidase	Protein fate
GSU1116		+	GSU1116	GAF domain protein, putative	Hypothetical proteins
GSU1123		+	30430	predicted zinc-dependent hydrolase	Unknown function
GSU1130		+	30432	chromosome segregation ATPase SMC	Cellular processes
GSU1131		+	30432	conserved hypothetical protein	Hypothetical proteins
GSU1132	ftsY	+	30432	cell division protein FtsY	Protein fate
GSU1137		+	30434	metal-dependent phosphohydrolase, HD superfamily	Unknown function
GSU1148		+	GSU1148	sensor histidine kinase (HAMP, HisKA, HATPase_c)	Regulatory functions
GSU1212		-	30450	conserved hypothetical protein	Hypothetical proteins
GSU1213		-	30450	conserved hypothetical protein	Hypothetical proteins
GSU1214		+	30451	O-methyltransferase, family 2	Cellular processes
GSU1215	cydD	+	30451	cytochrome bd ubiquinol oxidase assembly ATP-binding protein CydD	Transport and binding proteins

Table A1, cont.: List of all genes controlled by Fur in *G. sulfurreducens* according to ChIP-chip data.

GSU1216	cydC	+	30451	cytochrome bd ubiquinol oxidase assembly ATP-binding protein CydC	Transport and binding proteins
GSU1217		+	30451	conserved hypothetical protein	No Data
GSU1230		+	30454	putative ABC transporter periplasmic protein DUF534	Hypothetical proteins
GSU1268		-	GSU1268	transcriptional regulator, LysR family	Regulatory functions
GSU1269		-	GSU1269	hypothetical protein	No Data
GSU1318		+	GSU1318	conserved hypothetical protein	No Data
GSU1367		+	30485	conserved hypothetical protein	No Data
GSU1368		+	30485	conserved hypothetical protein	No Data
GSU1377		-	GSU1377	enoyl-CoA hydratase/isomerase	Energy metabolism
GSU1379	fur	+	GSU1379	ferric uptake regulation protein Fur	Regulatory functions
GSU1380	feoB-1	+	30486	ferrous iron transport protein B	Transport and binding proteins
GSU1381		+	30486	conserved hypothetical protein	No Data
GSU1390		+	30488	helix-turn-helix DNA-binding protein, putative	Unknown function
GSU1391		+	30488	Fic family protein	Unknown function
GSU1392		+	30488	CRISPR-associated protein Cas1	Mobile and extrachromosomal element functions
GSU1393		+	30488	CRISPR-associated protein, CT1978 family	Mobile and extrachromosomal element functions
GSU1424		+	30492	conserved hypothetical protein	No Data
GSU1425	lgt	+	GSU1425	prolipoprotein diacylglyceryl transferase	Protein fate
GSU1426	rsbW	-	30493	putative anti-sigma regulatory factor (serine/threonine protein kinase)	Regulatory functions
GSU1427		-	30493	anti-anti-sigma factor	Regulatory functions
GSU1461	pyrF	+	30502	orotidine 5'-phosphate decarboxylase	Purines, pyrimidines, nucleosides, and nucleotides
GSU1463	aspS	+	30503	aspartyl-tRNA synthetase	Protein synthesis
GSU1464		+	30503	LysM domain protein	Unknown function
GSU1486	tatC	+	30509	twin-arginine protein translocation pathway protein TatC	Unknown function
GSU1487	ribF	+	30509	riboflavin kinase and FAD synthetase	Biosynthesis of cofactors, prosthetic groups, and carriers
GSU1488		+	30509	membrane protein, putative	Cell envelope
GSU1489		+	30509	membrane protein, putative	Cell envelope
GSU1490	aroE	+	30509	shikimate 5-dehydrogenase	Amino acid biosynthesis
GSU1507		+	30514	heptosyltransferase family protein, putative	Central intermediary metabolism
GSU1508		+	30514	conserved hypothetical protein	Hypothetical proteins
GSU1509		+	30514	glycosyl transferase, group 2 family protein	Cell envelope

Table A1, cont.: List of all genes controlled by Fur in *G. sulfurreducens* according to ChIP-chip data.

GSU1510		+	GSU1510	glycosyl transferase, group 2 family protein	Cell envelope
GSU1513		+	30515	conserved possible cytochrome c, 1 heme-binding site	Hypothetical proteins
GSU1566		+	GSU1566	hypothetical protein	No Data
GSU1585		+	30530	protein of unknown function DUF150	Hypothetical proteins
GSU1586	nusA	+	30530	transcription elongation factor NusA	Transcription
GSU1587		+	30530	RNA-binding protein, putative	Unknown function
GSU1588	infB	+	30530	translation initiation factor IF-2	Protein synthesis
GSU1589	rbfA	+	30530	ribosome-binding factor A	Transcription
GSU1590		+	30530	phosphoesterase, putative	Unknown function
GSU1591	truB	+	30530	tRNA pseudouridine synthase B	Protein synthesis
GSU1592	rpsO	+	GSU1592	ribosomal protein S15	Protein synthesis
GSU1598		+	30533	protein of unknown function DUF177	Hypothetical proteins
GSU1599	rpmF	+	30533	ribosomal protein L32	Protein synthesis
GSU1600	plsX	+	30533	fatty acid/phospholipid synthesis protein PlsX	Fatty acid and phospholipid metabolism
GSU1601	fabH-2	+	30533	3-oxoacyl-(acyl-carrier-protein) synthase III	Fatty acid and phospholipid metabolism
GSU1602	fabD-2	+	30533	malonyl CoA-acyl carrier protein transacylase	Fatty acid and phospholipid metabolism
GSU1603	fabG-2	+	30533	3-oxoacyl-(acyl-carrier-protein) reductase	Fatty acid and phospholipid metabolism
GSU1606	rpiB	+	30534	ribose 5-phosphate isomerase B, putative	Energy metabolism
GSU1613		-	30536	endonuclease III family protein	DNA metabolism
GSU1614		-	30536	CoA-binding protein	Unknown function
GSU1639		+	GSU1639	transcriptional regulator, putative	Unknown function
GSU1657		+	30545	DNA internalization-related competence protein ComEC/Rec2	Cellular processes
GSU1658		+	30545	response receiver-modulated diguanylate cyclase (REC, GGDEF)	Regulatory functions
GSU1659	hisS	+	30545	histidyl-tRNA synthetase	Protein synthesis
GSU1714		-	GSU1714	BRO family protein, truncation	Unknown function
GSU1748		-	30564	RNA methyltransferase, TrmA family	Protein synthesis
GSU1749		-	30564	conserved hypothetical protein	No Data
GSU1750	infA	+	GSU1750	translation initiation factor IF-1	Protein synthesis
GSU1771		-	GSU1771	putative DNA/RNA-binding protein	Hypothetical proteins
GSU1782		-	30570	type IV pilus assembly ATPase PilM, putative	Hypothetical proteins
GSU1791	clpX	-	30573	ATP-dependent chaperone and Clp protease specificity component ClpX	Protein fate

Table A1, cont.: List of all genes controlled by Fur in *G. sulfurreducens* according to ChIP-chip data.

GSU1809	ftsH-2	-	30578	cell division ATP-dependent Zn protease FtsH	Cellular processes
GSU1845		-	GSU1845	conserved hypothetical protein	No Data
GSU1847		+	30587	ISGsu7, transposase OrfA	Mobile and extrachromosomal element functions
GSU1853		-	30588	membrane protein, putative	Cell envelope
GSU1854		-	30588	UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase	Cell envelope
GSU1856		+	GSU1856		Mobile and extrachromosomal element functions
GSU1858		-	GSU1858	IPT/TIG domain protein	Unknown function
GSU1886	yfiA	-	GSU1886	ribosomal subunit interface-associated sigma-54 modulation protein	Protein synthesis
GSU1895	pyrG	-	30596	CTP synthase	Purines, pyrimidines, nucleosides, and nucleotides
GSU1932		-	GSU1932	conserved hypothetical protein	No Data
GSU1969		-	30611	hypothetical protein	No Data
GSU1974		-	30611	DegT/DnrJ/EryC1/StrS family protein	Unknown function
GSU1982		-	30612	general secretion pathway predicted ATPase protein ExeA	Unknown function
GSU2009		+	30618	branched-chain amino acid ABC transporter, ATP-binding protein	Transport and binding proteins
GSU2031	pilN	-	30623	type IV pilus biogenesis protein PilN	Cell envelope
GSU2104		-	GSU2104	lipoprotein, putative	Cell envelope
GSU2118		+	GSU2118	integrative genetic element Gsu21, integrase	Mobile and extrachromosomal element functions
GSU2127		+	30644	ISGsu7, transposase OrfA	Mobile and extrachromosomal element functions
GSU2133		-	30645	lipoprotein, putative	Cell envelope
GSU2171		+	30654	ISGsu7, transposase OrfB	Mobile and extrachromosomal element functions
GSU2178		+	GSU2178	conserved hypothetical protein	No Data
GSU2186		-	30658	conserved hypothetical protein	Hypothetical proteins
GSU2187		-	30658	ABC transporter, membrane protein	Transport and binding proteins
GSU2188		-	30658	ABC transporter, ATP-binding protein	Transport and binding proteins
GSU2189		+	GSU2189	sensor histidine kinase (HATPase_c)	Signal transduction
GSU2191		-	GSU2191	aldehyde ferredoxin oxidoreductase, tungsten-containing	Central intermediary metabolism
GSU2193		-	GSU2193	ferritin-like domain protein	Hypothetical proteins
GSU2200		+	GSU2200	conserved hypothetical protein	No Data
GSU2219	cheY-10	-	GSU2219	chemotaxis protein CheY	Signal transduction

Table A1, cont.: List of all genes controlled by Fur in *G. sulfurreducens* according to ChIP-chip data.

GSU2263		-	30673	oxidoreductase, Gfo/Idh/MocA family	Unknown function
GSU2264	lpxA-1	-	30673	acyl-(acyl-carrier-protein)--UDP-N-acetylglucosamine O-acyltransferase	Cell envelope
GSU2265	fabZ	-	30673	(3R)-hydroxymyristoyl-(acyl-carrier-protein) dehydratase	Fatty acid and phospholipid metabolism
GSU2268		-	30675	outer membrane surface protein	Cell envelope
GSU2269	lolD	-	30675	lipoprotein release ABC transporter, ATP-binding protein	Transport and binding proteins
GSU2270	lolE	-	30675	lipoprotein release ABC transporter, membrane protein	Transport and binding proteins
GSU2280		-	30677	ISGsu7, transposase OrfA	Mobile and extrachromosomal element functions
GSU2301		+	30680	conserved hypothetical protein	No Data
GSU2309		+	30683	metallo-beta-lactamase family protein	Unknown function
GSU2310		+	30683	YeeE/YedE family protein	Unknown function
GSU2358		+	30693	1,4-alpha-glucan branching enzyme, possible malto-oligosyltrehalose trehalohydrolase	Energy metabolism
GSU2391		-	30700	ISGsu7, transposase OrfB	Mobile and extrachromosomal element functions
GSU2392		-	30700	ISGsu7, transposase OrfA	Mobile and extrachromosomal element functions
GSU2427		-	GSU2427	conserved hypothetical protein	No Data
GSU2471		-	GSU2471	RNA-directed DNA polymerase and maturase, group II intron origin	Mobile and extrachromosomal element functions
GSU2481	kdpB	+	30714	potassium-transporting ATPase, B subunit	Transport and binding proteins
GSU2593		+	30738	ISGsu6, transposase OrfA (integrase)	Mobile and extrachromosomal element functions
GSU2594		+	30738	ISGsu6, transposase OrfB (ATP/GTP-binding)	Mobile and extrachromosomal element functions
GSU2674		+	GSU2674	hypothetical protein	No Data
GSU2676		-	30758	membrane protein, putative	Cell envelope
GSU2677		-	30758	membrane protein, putative	Cell envelope
GSU2678		+	GSU2678	ATP-independent chaperone, alpha-crystallin/Hsp20 family	Protein fate
GSU2724		-	30767	cytochrome c, 13-15 heme-binding sites	Energy metabolism
GSU2725		-	30767	cytochrome c, 5 heme-binding sites	Energy metabolism
GSU2726		-	30767	conserved hypothetical protein	No Data
GSU2767		-	GSU2767	cytochrome c, 2 heme-binding sites	Energy metabolism
GSU2873		-	30792	iron-sulfur oxidoreductase	Unknown function
GSU2874	argC	-	30792	N-acetyl-glutamyl-5-phosphate reductase	Amino acid biosynthesis

Table A1, cont.: List of all genes controlled by Fur in *G. sulfurreducens* according to ChIP-chip data.

GSU2878	asd	-	30794	aspartate-semialdehyde dehydrogenase, putative	Amino acid biosynthesis
GSU2897		-	30799	conserved hypothetical protein	No Data
GSU2898	omcN	-	30799	cytochrome c, 27-34 heme-binding sites	Energy metabolism
GSU2899		-	30799	cytochrome c, 16-23 heme-binding sites	Energy metabolism
GSU2938		-	30810	conserved hypothetical protein	No Data
GSU2939		-	30810	putative porin	Transport and binding proteins
GSU2940		-	30810	rhodanese-related sulfurtransferase	Cell envelope
GSU2942	mcp032	+	GSU2942	methyl-accepting chemotaxis sensory transducer	Cellular processes
GSU2965		-	GSU2965	conserved hypothetical protein	Hypothetical proteins
GSU2967		+	GSU2967	ferritin-like domain protein	No Data
GSU2969		+	30818	sensor diguanylate cyclase (PAS, GGDEF)	Regulatory functions
GSU2970		+	30818	conserved hypothetical protein	Hypothetical proteins
GSU2977		-	30820	transaldolase	Unknown function
GSU2978		-	30820	conserved hypothetical protein	No Data
GSU2979	folK	-	30820	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	Biosynthesis of cofactors, prosthetic groups, and carriers
GSU2980		-	30820	metal-binding domain transcriptional regulator	Regulatory functions
GSU2981		-	30821	TonB-related periplasmic protein, putative	Transport and binding proteins
GSU2982		-	30821	TonB-dependent outer membrane receptor, putative	Transport and binding proteins
GSU2991		-	30824	sensor histidine kinase (HAMP, HisKA, HATPase_c)	Regulatory functions
GSU3008	cobS	-	30828	cobalamin 5'-phosphate synthase	Biosynthesis of cofactors, prosthetic groups, and carriers
GSU3041	csrA	-	30834	carbon storage regulator	Energy metabolism
GSU3053	fliA	-	30837	RNA polymerase sigm-28a factor for flagellar operon	Cellular processes
GSU3054	flhG	-	30837	flagellar biosynthetic protein FlhG (ATPase)	Cellular processes
GSU3055	flhF	-	30837	flagellar biosynthetic protein FlhF, putative	Cellular processes
GSU3056	flhA	-	30837	flagellar biosynthetic protein FlhA	Cellular processes
GSU3081		-	GSU3081	protein of unknown function DUF164, C-terminal fragment	Disrupted reading frame
GSU3082		+	30842	ISGsu7, transposase OrfA	Mobile and extrachromosomal element functions
GSU3083		+	30842	ISGsu7, transposase OrfB	Mobile and extrachromosomal element functions
GSU3098	hisB	-	30846	imidazoleglycerol-phosphate dehydratase	Amino acid biosynthesis

Table A1, cont.: List of all genes controlled by Fur in *G. sulfurreducens* according to ChIP-chip data.

GSU3257	glgA-2	+	30880	glycogen synthase	Energy metabolism
GSU3267		-	30882	conserved hypothetical protein	No Data
GSU3268	feoB-2	-	30882	ferrous iron transport protein B, putative	Transport and binding proteins
GSU3269		-	GSU3269		Transport and binding proteins
GSU3270		-	GSU3270	FeoA family protein	Unknown function
GSU3271		-	GSU3271	carbohydrate-selective porin OprB	No Data
GSU3272		-	30883	hypothetical protein	No Data
GSU3273		-	30883	conserved hypothetical protein	No Data
GSU3274		-	30883	cytochrome c, 1 heme-binding site	Energy metabolism
GSU3275		+	30884	conserved hypothetical protein	No Data
GSU3276		+	30884	LysM domain protein	Unknown function
GSU3306		+	GSU3306	conserved hypothetical protein	No Data
GSU3342		+	30901	conserved hypothetical protein	Hypothetical proteins
GSU3395	putA	-	GSU3395	proline dehydrogenase/delta-1-pyrroline-5-carboxylate dehydrogenase	Energy metabolism

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSS0001			+				2.095	0.00412		0.706		0.358
GSS0021			-				3.529	0.00371		0.156	7.282	0.024
GSS0028			-					0.163	-2.343	0.0235	-3.026	0.0106
GSS0030			+				2.791	0.00691		0.342		0.611
GSS0036			+				2.754	0.0165		0.431		0.0532
GSU0033		dnaK	+	GSU0033	chaperone protein DnaK	Protein fate	-2.219	0.000899		0.107		0.631
GSU0071			-	GSU0071	conserved hypothetical protein	No Data	2.156	0.00579		0.408		0.321
GSU0077			+	30210	conserved hypothetical protein	No Data	2.016	0.00393		0.177		0.906
GSU0078			+	30210	PilZ domain protein	No Data	2.222	0.00126		0.492		0.755
GSU0085		hdrF	-	GSU0085	heterodisulfide oxidoreductase, NAD(P)H oxidoreductase subunit	Energy metabolism		0.001		0.557	-2.312	0.0468
GSU0086.1			-	GSU0086.1	hypothetical protein			0.000856		0.0672	-3.433	0.0135
GSU0087		hdrE	-	30212	heterodisulfide reductase, iron-sulfur binding subunit, putative	Energy metabolism		0.000773		0.0924	-2.84	0.0166
GSU0088		hdrD	-	30212	heterodisulfide reductase, iron-sulfur cluster-binding subunit, putative	Energy metabolism		0.00231		0.236	-2.156	0.0279
GSU0089		mvhD	-	30212	methyl viologen-reducing hydrogenase, iron-sulfur cluster-containing subunit, MvhD	Energy metabolism		0.00255		0.143	-2.539	0.0152
GSU0090		hdrA	-	30212	heterodisulfide reductase, subunit A	Energy metabolism		0.000862		0.316	-2.087	0.0305
GSU0208			+	GSU0208	conserved hypothetical protein	No Data		0.586	2.711	0.0353	2.804	0.0226
GSU0235		tex	-	30243	S1 RNA-binding domain-containing transcriptional accessory protein	Unknown function		0.00085		0.0782	2.26	0.0151
GSU0252			+	GSU0252	ISGsu1, transposase	Mobile and extrachromosomal element functions		0.042	-2.022	0.0353		0.0614
GSU0264			-	30249	transporter, putative	Cellular processes	-6.681	0.000155		0.0273	-4.399	0.00134
GSU0265			-	30249	membrane protein, putative	Cell envelope	-8.093	0.000275		0.0236	-4.73	0.00145
GSU0339		nuoB	+	30264	NADH dehydrogenase I, B subunit	Energy metabolism	-2.041	0.00216		0.141		0.0545
GSU0340		nuoC	+	30264	NADH dehydrogenase I, C subunit	Energy metabolism	-2.004	0.0028		0.112	-6.884	0.0426
GSU0341	Y	nuoD	+	30265	NADH dehydrogenase I, D subunit	Energy metabolism		0.000868		0.102	-4.788	0.0335
GSU0342		nuoE-1	+	30265	NADH dehydrogenase I, E subunit	Energy metabolism		0.000915		0.134	-4.601	0.046
GSU0343		nuoF-1	+	30265	NADH dehydrogenase I, F subunit	Energy metabolism		0.000853		0.156	-4.863	0.05
GSU0345		nuoH-1	+	30266	NADH dehydrogenase I, H subunit	Energy metabolism	-2.23	0.00104		0.207		0.0711
GSU0346		nuoI-1	+	30266	NADH dehydrogenase I, I subunit	Energy metabolism	-2.067	0.00734		0.289		0.118
GSU0347		nuoJ-1	+	30266	NADH dehydrogenase I, J subunit	Energy metabolism	-2.218	0.00168		0.256		0.0915
GSU0348		nuoK-1	+	30266	NADH dehydrogenase I, K subunit	Energy metabolism	-2.279	0.00234		0.154		0.0524
GSU0349		nuoL-1	+	30267	NADH dehydrogenase I, L subunit	Energy metabolism	-2.314	0.000381		0.2		0.0586
GSU0350		nuoM-1	+	30267	NADH dehydrogenase I, M subunit	Energy metabolism	-2.324	0.000159		0.243		0.0659

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSU0351		nuoN-1	+	GSU0351	NADH dehydrogenase I, N subunit	Energy metabolism	-2.477	0.000304		0.291		0.0841
GSU0356			-	GSU0356		Regulatory functions		0.103	-2.557	0.0183	-2.975	0.00878
GSU0357			-	30269	cytochrome c, 7-8 heme-binding sites	Energy metabolism		0.0669	-6.22	0.0194	-5.617	0.0124
GSU0384			-	GSU0384	ferritin-like domain protein	Hypothetical proteins		0.00292	-2.253	0.0204	-4.481	0.00145
GSU0434			-	GSU0434	thiamine biosynthesis protein ThiI-related adenine nucleotide alpha hydrolase superfamily protein	Unknown function		0.546	2.854	0.025	2.738	0.0185
GSU0466		macA	-	GSU0466	cytochrome c, 2 heme-binding sites	Energy metabolism		0.00921	-6.309	0.0252	-4.51	0.0227
GSU0469			-	30292	hypothetical protein	No Data	-2.182	0.0036		0.0235		0.0423
GSU0495			+	GSU0495		No Data	-2.054	0.0102		0.769		0.55
GSU0509		sfrA	-	GSU0509	Fe(III) reductase, alpha subunit	Energy metabolism	-2.177	0.00213		0.0706	-4.444	0.0122
GSU0510		sfrB	-	GSU0510	Fe(III) reductase, beta subunit	Energy metabolism		0.000115	-2.075	0.00531	-3.596	0.000801
GSU0538		hspA-1	-	GSU0538	ATP-independent chaperone, alpha-crystallin/Hsp20 family	Protein fate	-2.392	0.0012		0.379		0.772
GSU0593			-	30317	cytochrome b, putative	Hypothetical proteins	2.635	0.00566		0.144		0.708
GSU0594			-	30317	cytochrome c, 7 heme-binding sites	Energy metabolism	2.416	0.00255		0.154		0.7
GSU0595.1			-	GSU0595.1	conserved hypothetical protein		2.432	0.00368		0.199		0.78
GSU0596			-	30318	response receiver (REC)	Regulatory functions	2.118	0.00965		0.361		0.984
GSU0618		omcE	+	GSU0618	cytochrome c, 4 heme-binding sites	Energy metabolism	-3.246	0.000299		0.195	-3.888	0.00233
GSU0619			+	GSU0619	conserved hypothetical protein	No Data	-2.044	0.000272		0.312	-2.282	0.00743
GSU0658		clpB	+	GSU0658	chaperone ATPase ClpB	Protein fate	-3.001	0.00118		0.309		0.533
GSU0676			+	30334	lipoprotein, putative	Cell envelope		0.178	2.383	0.0331	2.687	0.0159
GSU0677			+	30334	ABC transporter, membrane protein, putative	Transport and binding proteins		0.198	2.35	0.00884	2.493	0.00154
GSU0678			+	30334	ABC transporter, ATP-binding protein	Transport and binding proteins		0.00642	2.519	0.00694	3.743	0.000832
GSU0724			-	GSU0724	conserved hypothetical protein	Hypothetical proteins	2.061	0.00297		0.739		0.405
GSU0735			-	30350	PTS system IIA protein, fructose subfamily	Signal transduction		0.907	-2.064	0.0266	-2.053	0.0196
GSU0739		ehrA	+	30352	Ech-hydrogenase-related complex NuoL-like integral membrane subunit	Energy metabolism		0.00902		0.0634	-2.848	0.0159
GSU0740		ehrB	+	30352	Ech-hydrogenase-related complex NuoH-like integral membrane subunit	Energy metabolism		0.000634	-2.266	0.0286	-3.362	0.00898
GSU0741		ehrC	+	30352	Ech-hydrogenase-related complex HyfE-like integral membrane subunit	Energy metabolism		0.00377	-2.442	0.0251	-3.713	0.00736
GSU0742		ehrD	+	30352	Ech-hydrogenase-related complex HyfF-like integral membrane subunit subunit	Energy metabolism		0.00956	-2.472	0.0191	-3.36	0.00611
GSU0743		ehrL	+	30352	Ech-hydrogenase-related complex, large subunit	Energy metabolism		0.00785		0.0839	-2.749	0.0398
GSU0744			+	30352	hypothetical protein	No Data		0.0123		0.0564	-4.004	0.0234
GSU0745		ehrS	+	30352	Ech-hydrogenase-related complex, small subunit	Energy metabolism		0.00663	-2.255	0.0413	-3.125	0.0166

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSU0777		fdnG	+	30358	formate dehydrogenase, major subunit, selenocysteine-containing	Energy metabolism	4.341	0.000774		0.65		0.613
GSU0778		fdnH	+	30358	formate dehydrogenase, iron-sulfur subunit	Energy metabolism	2.554	0.00216		0.991		0.524
GSU0779			+	30358	formate dehydrogenase, b-type cytochrome subunit, putative	Energy metabolism	2.011	0.0131		0.965		0.495
GSU0781			+	30358	twin-arginine translocation protein, TatA/E family	Protein fate	2.858	0.000894		0.638		0.554
GSU0782		hybS	+	30359	periplasmically oriented, membrane bound [NiFe]-hydrogenase small subunit	Energy metabolism	57.397	0.000161		0.234		0.184
GSU0783		hybA	+	30359	periplasmically oriented, membrane bound [NiFe]-hydrogenase iron-sulfur cluster binding subunit	Energy metabolism	51.475	0.000372		0.283		0.192
GSU0784	Y	hybB	+	30359	periplasmically oriented, membrane bound [NiFe]-hydrogenase integral membrane subunit	Energy metabolism	34.895	0.000758		0.275		0.205
GSU0785		hybL	+	30360	periplasmically oriented, membrane bound [NiFe]-hydrogenase large subunit	Energy metabolism	52.466	0.000287		0.26		0.156
GSU0786		hybP	+	30360	periplasmically oriented, membrane bound [NiFe]-hydrogenase maturation protease	Protein fate	43.834	0.000354		0.287		0.187
GSU0787			+	30360	twin-arginine translocation protein, TatA/E family	Protein fate	28.682	0.000613		0.251		0.205
GSU0788			+	30360	conserved hypothetical protein	Hypothetical proteins	17.549	0.00146		0.298		0.279
GSU0789			-	GSU0789	response regulator, putative (REC)	Signal transduction	5.182	0.00101		0.435		0.131
GSU0790			-	GSU0790	conserved hypothetical protein	No Data	2.978	0.00119		0.615		0.141
GSU0793			-	30361	conserved hypothetical protein	No Data		0.224	-2.503	0.00659	-2.695	0.00094
GSU0828	Y		+	30370	efflux pump, RND family, outer membrane protein	Cellular processes		0.932	16.73	0.0344	16.778	0.0238
GSU0829	Y		+	30370	efflux pump, RND family, membrane fusion protein	Cellular processes		0.644	20.312	0.0277	20.796	0.0184
GSU0830	Y		+	30370	efflux pump, RND family, inner membrane protein	Transport and binding proteins		0.815	17.325	0.0263	17.182	0.018
GSU0831	Y		+	30370	nitrogen regulatory protein P-II, putative	Regulatory functions		0.811	21.26	0.0206	21.626	0.0127
GSU0832	Y		+	30370	lipoprotein, putative	Cell envelope		0.205	21.608	0.0187	22.318	0.00886
GSU0834			+	GSU0834	conserved hypothetical protein	No Data		0.606	6.508	0.00716	6.315	0.00291
GSU0835			-	30371	conserved hypothetical protein	No Data	2.806	0.00389		0.0195	5.542	0.00354
GSU0836			-	30371	transcription elongation factor C-terminal domain protein	Regulatory functions		0.00391	2.331	0.00688	4.406	0.00189
GSU0837			-	GSU0837	response regulator (REC)	Regulatory functions		0.00355	2.546	0.0184	4.146	0.00393
GSU0840			-	GSU0840		No Data		0.00208	2.511	0.0134	4.286	0.00174
GSU0885			-	GSU0885	phosphotransacetylase N-terminal domain protein	Biosynthesis of cofactors,	2.277	0.00104		0.189	3.033	0.0124

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSU0908			-	30386	moaD family protein	prosthetic groups, and carriers Biosynthesis of cofactors, prosthetic groups, and carriers	2.117	0.000268		0.568		0.452
GSU0914		rhlE-2	+	GSU0914	ATP-dependent RNA helicase RhlE	Transcription		0.716	2.023	0.0312	2.042	0.0215
GSU0919			-	GSU0919	conserved hypothetical protein	No Data	6.428	0.00244	-9.221	0.0373		0.439
GSU0930			+	GSU0930	conserved hypothetical protein, selenocysteine-containing	Unknown function		0.00755	-2.05	0.0137		0.0107
GSU0957			-	30395	ISGsu1, transposase	Mobile and extrachromosomal element functions		0.0264	-2.043	0.0362		0.0587
GSU0975			+	30399	phage tail sheath protein, putative	Mobile and extrachromosomal element functions	-2.001	0.00109		0.742		0.406
GSU0990			+	30402	hypothetical protein	No Data		0.000247		0.297	-2.709	0.0499
GSU0993			-	GSU0993	thioesterase superfamily protein	Hypothetical proteins	-3.303	0.000728		0.929	-3.219	0.0227
GSU0994		fumB	-	GSU0994	fumarate hydratase, class I	Energy metabolism	-2.459	0.00317	-2.593	0.0371	-6.377	0.00423
GSU0996			+	GSU0996	conserved hypothetical protein	No Data		0.0157	-6.902	0.0423		0.0534
GSU1033		mcp014	+	GSU1033	methyl-accepting chemotaxis sensory transducer	Cellular processes		0.000649		0.204	2.465	0.0227
GSU1060	Y		-	GSU1060	conserved hypothetical protein	No Data		0.00246		0.071	5.149	0.0341
GSU1069			-	GSU1069	conserved hypothetical protein	Hypothetical proteins	-2.092	0.00637		0.0331		0.275
GSU1106			+	GSU1106	citrate synthase	Energy metabolism	-2.264	0.00455		0.08	-3.227	0.00153
GSU1108			+	GSU1108	aldehyde dehydrogenase family protein	Energy metabolism	-4.671	0.000357		0.148	-9.077	0.0127
GSU1177		frdA	+	30443	succinate dehydrogenase/fumarate reductase, flavoprotein subunit	Energy metabolism		0.0242		0.0527	-2.329	0.0128
GSU1181			-	GSU1181	conserved hypothetical protein	No Data		0.0209		0.133	-2.024	0.04
GSU1228		omcI	+	30454	cytochrome c, 9-10 heme-binding sites	Energy metabolism		0.0247		0.0783	-2.128	0.046
GSU1229			+	30454	lipoprotein, putative	Cell envelope		0.025		0.0953	-2.164	0.0497
GSU1248			-	GSU1248	transposase, ISL3 family, truncation	Mobile and extrachromosomal element functions	15.446	0.000101		0.0983	9.97	0.0024
GSU1251			+	30460	RRC1 domain protein	Unknown function	16.049	0.000117		0.616	17.284	0.000902
GSU1252			+	30460	multicopper oxidase with phosphopantotheine attachment site	Hypothetical proteins	18.185	0.000139		0.875	17.827	0.000796
GSU1253			+	30460	conserved hypothetical protein	No Data	7.449	0.000269		0.365	6.984	0.000858
GSU1254			+	30460	hypothetical protein	No Data	6.689	0.000269		0.978	6.669	0.00107
GSU1255			+	30461	conserved hypothetical protein	Hypothetical proteins	12.615	0.000105		0.0234	8.172	0.000736
GSU1256			+	30461	conserved domain protein	Hypothetical proteins	10.106	0.0000856		0.0853	6.5	0.0035
GSU1257			+	30461	cytochrome c and periplasmic receptor, 1 heme-binding site	Energy metabolism	5.604	0.000274		0.468	4.874	0.00736

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSU1281		cbiO-2	+	30467	cobalt ABC transporter, ATP-binding protein	Transport and binding proteins	2.071	0.0022		0.335		0.335
GSU1282			+	30467	putative porin	No Data	2.282	0.00157		0.393		0.627
GSU1305		gdhA	-	GSU1305	NADP-dependent glutamate dehydrogenase	Energy metabolism	-2.085	0.000785		0.502		0.672
GSU1338			+	GSU1338	heavy metal transport/detoxification domain protein	Unknown function		0.0477	2.29	0.0222	2.117	0.0167
GSU1339			+	30479	conserved hypothetical protein	No Data		0.00648	2.385	0.018		0.017
GSU1340			+	30479	ABC transporter, membrane protein	Transport and binding proteins		0.333	2.115	0.0323		0.0393
GSU1379	Y	fur	+	GSU1379	ferric uptake regulation protein Fur	Regulatory functions	3.856	0.00785	3.479	0.047	13.419	0.00191
GSU1380	Y	feoB-1	+	30486	ferrous iron transport protein B	Transport and binding proteins	10.871	0.00238		0.0509	48.606	0.00185
GSU1381	Y		+	30486	conserved hypothetical protein	No Data	10.159	0.00192	4.206	0.0327	42.735	0.00135
GSU1382		ideR	+	GSU1382	iron/manganese-dependent transcriptional regulator	Regulatory functions	10.157	0.00241	4.397	0.034	44.669	0.000859
GSU1406			-	GSU1406	conserved hypothetical protein (selenocysteine-containing)	Hypothetical proteins	-3.063	0.00116		0.0775	-6.052	0.00832
GSU1408			-	GSU1408	ParA family protein	Cellular processes	-2.722	0.00103		0.0672	-5.575	0.00851
GSU1409			-	30490	NifU-like domain protein	Unknown function	-2.513	0.012		0.0844	-6.673	0.0123
GSU1410			-	30490	conserved hypothetical protein	Hypothetical proteins	-2.596	0.00575		0.0674	-6.607	0.00895
GSU1411			-	30490	arsenite efflux pump ACR3 and related membrane proteins family member	Transport and binding proteins	-2.148	0.0031	-2.98	0.0232	-6.403	0.00235
GSU1412			-	30490	hypothetical protein	No Data	-2.368	0.00742		0.0636	-7.151	0.0112
GSU1412.1			-	GSU1412.1	iron-molybdenum cluster-binding protein		-2.223	0.0158		0.0542	-9.654	0.0127
GSU1465		icd	+	30504	isocitrate dehydrogenase, NADP-dependent	Energy metabolism	-2.059	0.000265		0.0804	-3.548	0.0131
GSU1466		mdh	+	30504	malate dehydrogenase	Energy metabolism		0.00557		0.0962	-3.333	0.0143
GSU1467			+	30505	2-oxoglutarate:ferredoxin oxidoreductase, ferredoxin subunit	Energy metabolism	-2.037	0.000423		0.0834	-8.916	0.034
GSU1468			+	30505	2-oxoglutarate:ferredoxin oxidoreductase, alpha subunit	Energy metabolism		0.00088		0.0666	-6.241	0.022
GSU1469			+	30505	2-oxoglutarate:ferredoxin oxidoreductase, thiamine pyrophosphate-binding subunit	Energy metabolism	-2.098	0.00108		0.0869	-5.905	0.025
GSU1470			+	GSU1470	2-oxoglutarate:ferredoxin oxidoreductase, gamma subunit	Energy metabolism	-2.79	0.00808		0.195	-6.961	0.0392
GSU1471			+	GSU1471	metal dependent phosphohydrolase, HD superfamily	Unknown function	-2.899	0.00154		0.303	-4.584	0.0341
GSU1472			+	30506	PATAN domain protein	No Data	-2.797	0.00169		0.356	-4.377	0.0451
GSU1473			+	30506	membrane protein, putative	Cell envelope	-2.621	0.00159		0.635		0.0573
GSU1538			+	GSU1538	methylamine utilization cytochrome c peroxidase MauG, 2 heme-binding sites	Energy metabolism		0.161	-4.892	0.0193	-6.03	0.00339

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSU1622			+	GSU1622	membrane protein, putative	Transport and binding proteins	-2.171	0.00385		0.467		0.104
GSU1639	Y		+	GSU1639	transcriptional regulator, putative	Unknown function		0.044	3.115	0.0331		0.0576
GSU1642			-	GSU1642	ferritin-like domain protein	Hypothetical proteins	2.092	0.00242		0.287		0.937
GSU1647			-	GSU1647	conserved hypothetical protein	No Data		0.0484	-4.477	0.0133	-5.216	0.00413
GSU1660		acnB	+	GSU1660	aconitate hydratase 2	Energy metabolism	-3.309	0.00211		0.0955	-8.735	0.0156
GSU1681			-	30549	phosphotransacetylase N-terminal domain protein	Biosynthesis of cofactors, prosthetic groups, and carriers		0.000642		0.354	-2.289	0.0465
GSU1700		maeB	-	GSU1700	malate oxidoreductase (phosphate acetyltransferase-like domain fusion)	Energy metabolism		0.000674		0.0666	-2	0.00334
GSU1705		panB	+	30554	ketopantoate hydroxymethyltransferase	Biosynthesis of cofactors, prosthetic groups, and carriers	-2.237	0.000857		0.0188		0.141
GSU1706		panC	+	30554	pantoate--beta-alanine ligase	Biosynthesis of cofactors, prosthetic groups, and carriers	-2.435	0.000357		0.0672		0.917
GSU1707			+	30554	pyridoxal phosphate-dependent decarboxylase	Energy metabolism	-2.128	0.000357		0.019		0.0748
GSU1740			+	GSU1740	cytochrome c, 1 heme-binding site	Energy metabolism		0.364	-2.457	0.0347	-2.161	0.0219
GSU1844			-	GSU1844	IPT/TIG domain protein, putative	Unknown function	23.982	0.000173		0.109	20.425	0.000224
GSU1855.2			-	GSU1855.2	conserved hypothetical protein		-2.001	0.00832		0.558	-2.215	0.0179
GSU1858	Y		-	GSU1858	IPT/TIG domain protein	Unknown function	-2.051	0.00126		0.127		0.146
GSU1917		uppS	-	GSU1917	undecaprenyl diphosphate synthase	Biosynthesis of cofactors, prosthetic groups, and carriers		0.663	2.031	0.0274		0.0182
GSU1943			-	GSU1943	hypothetical protein	No Data	2.468	0.0031		0.431		0.0761
GSU1945			-	GSU1945	fibronectin type III domain protein	Unknown function	2.153	0.000756		0.182		0.0615
GSU1947			-	30608	hypothetical protein	No Data		0.00919		0.0872	2.789	0.0283
GSU1994			+	GSU1994	hypothetical protein	No Data	2.203	0.00344		0.669		0.125
GSU2012		nifU	-	30619	nitrogen fixation iron-sulfur cluster assembly protein NifU	Unknown function		0.0334		0.0717	2.949	0.0236
GSU2034			-	30624	conserved hypothetical protein	No Data		0.00572		0.208	-2.34	0.045
GSU2036			-	30625	conserved hypothetical protein	No Data		0.0208		0.248	-3.214	0.0306
GSU2037			-	30625	conserved hypothetical protein	No Data	-2.005	0.00231		0.083	-3.37	0.0124
GSU2038			-	30625	Tfp pilus assembly protein tip-associated adhesin-like protein	No Data	-2.093	0.000542		0.453		0.0687
GSU2073			-	GSU2073	EF hand domain/PKD domain protein	Unknown function	-3.94	0.000769		0.271	-8.978	0.046
GSU2074			-	30635	PPIC-type PPIASE domain protein	Unknown function	-2.999	0.00035		0.201		0.0526
GSU2075			-	30635	serine protease, subtilase family	Protein fate	-3.65	0.000837		0.19	-11.908	0.0444
GSU2076		omcZ	-	GSU2076	cytochrome c, 7-8 heme-binding sites	Energy metabolism	-2.204	0.00415		0.0782	-14.14	0.0329

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSU2097			-	30641	carbon monoxide dehydrogenase accessory protein, putative	Energy metabolism	-2.161	0.000964		0.978	-2.169	0.0126
GSU2098		cooS	-	30641	carbon monoxide dehydrogenase, catalytic subunit	Energy metabolism	-2.835	0.000253		0.431	-2.581	0.00728
GSU2129			-	GSU2129	conserved hypothetical protein	No Data		0.176		0.0712	4.472	0.0467
GSU2131			-	30645	conserved hypothetical protein	No Data		0.2	9.318	0.0107	11.576	0.00369
GSU2132			-	30645	hypothetical protein	No Data		0.0461	16.146	0.00569	20.662	0.000924
GSU2133	Y		-	30645	lipoprotein, putative	Cell envelope		0.0454	18.436	0.0106	23.577	0.000714
GSU2134			-	30645	nitrogen regulatory protein P-II, putative	Regulatory functions		0.0422	24.443	0.00701	31.487	0.00218
GSU2135			-	30645	heavy metal efflux pump, RND family, inner membrane protein, CzcA family	Transport and binding proteins		0.0501	30.827	0.0192	34.96	0.00948
GSU2136			-	30645	efflux pump, RND family, membrane fusion protein	Cellular processes		0.155	33.753	0.0195	37.935	0.0103
GSU2137			-	30645	metal ion efflux pump, RND family, outer membrane protein, CzcC family	Cellular processes		0.00644	32.101	0.0233	39.388	0.0135
GSU2138			-	GSU2138	ISGsu1, transposase, interruption	No Data		0.0627	-2.014	0.0205		0.0152
GSU2180			-	GSU2180	ISGsu1, transposase	Mobile and extrachromosomal element functions		0.0313	-2.014	0.0363		0.0776
GSU2200	Y		+	GSU2200	conserved hypothetical protein	No Data		0.0111		0.111	-4.42	0.0375
GSU2201			+	GSU2201	cytochrome c, 8 heme-binding sites	Energy metabolism	-2.07	0.000371		0.143	-4.506	0.037
GSU2203		omcK	+	30661	cytochrome c, 10 heme-binding sites	Energy metabolism	-2.298	0.00155		0.208	-2.764	0.0107
GSU2204		omcL	+	30661	cytochrome c, 1 heme-binding site	Energy metabolism	-2.114	0.00145		0.238	-2.443	0.0102
GSU2320			+	30687	conserved hypothetical protein	No Data	2.317	0.00835		0.67		0.248
GSU2321			+	30687	hypothetical protein	No Data	2.89	0.00125		0.446		0.116
GSU2323			+	30687	conserved hypothetical protein	Cell envelope	2.086	0.00176		0.565		0.405
GSU2324			+	30687	conserved hypothetical protein	No Data	2.482	0.00117		0.523		0.461
GSU2325			+	30687	cation-transport ATPase, E1-E2 family	Transport and binding proteins	2.334	0.000757		0.651		0.217
GSU2326			+	30687	outer membrane lipoprotein	Cell envelope	2.396	0.000756		0.236		0.333
GSU2327			+	30687	conserved hypothetical protein	Hypothetical proteins	2.229	0.000353		0.35		0.251
GSU2328			+	GSU2328		Hypothetical proteins	2.384	0.0031		0.238		0.405
GSU2329		cfa	+	30688	cyclopropane-fatty-acyl-phospholipid synthase	Fatty acid and phospholipid metabolism	2.213	0.0024		0.261		0.366
GSU2352		aplC	-	GSU2352	sodium/solute symporter family protein	Transport and binding proteins	-2.079	0.00275		0.106		0.579
GSU2404			-	GSU2404	low-complexity protein	Unknown function	-2.268	0.00682		0.922		0.191
GSU2405			-	30701	hypothetical protein	No Data	-3.289	0.00126		0.67		0.138
GSU2406			-	30701	DnaJ domain protein	Unknown function	-3.8	0.00231		0.629		0.15
GSU2407			-	30701	conserved hypothetical protein	No Data	-5.197	0.00257		0.997	-5.211	0.0367
GSU2408			-	30702	ATP-independent chaperone, alpha-crystallin/Hsp20 family	Protein fate	-4.213	0.00324		0.668		0.193

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSU2409			-	30702	ATP-independent chaperone, alpha-crystallin/Hsp20 family	Protein fate	-5.472	0.00156		0.453		0.144
GSU2410		hspA-2	-	30702	ATP-independent chaperone, alpha-crystallin/Hsp20 family	Protein fate	-4.699	0.00307		0.588		0.134
GSU2424			+	GSU2424	conserved hypothetical protein	No Data		0.0459		0.0308	2.273	0.0135
GSU2442			-	GSU2442	RelA/SpoT domain protein	Unknown function		0.247	-2.347	0.0234		0.0241
GSU2469			+	GSU2469	hypothetical protein	No Data		0.0211		0.248	-2.312	0.0183
GSU2478			+	30714	hypothetical protein	No Data	3.323	0.0326		0.856		0.29
GSU2479			+	30714	hypothetical protein	No Data	2.515	0.0119		0.922		0.279
GSU2480		kdpA	+	30714	potassium-transporting ATPase, A subunit	Transport and binding proteins	2.348	0.0115		0.996		0.373
GSU2481	Y	kdpB	+	30714	potassium-transporting ATPase, B subunit	Transport and binding proteins	2.319	0.0131		0.973		0.396
GSU2482		kdpC	+	30714	potassium-transporting ATPase, C subunit	Transport and binding proteins	2.059	0.00961		0.955		0.35
GSU2483		kdpD	+	30714	osmosensitive potassium channel sensor histidine kinase KdpD (HisKA, HATPase_c)	Signal transduction	2.08	0.00503		0.998		0.333
GSU2484		kdpE	+	30714	winged-helix transcriptional response regulator KdpE (REC, transregC)	Signal transduction	2.019	0.00349		0.977		0.33
GSU2485			+	GSU2485	Kup system potassium transporter	Transport and binding proteins	2.06	0.00171		0.0226		0.519
GSU2493			-	GSU2493	NHL repeat domain protein	Unknown function		0.00227		0.142	-2.468	0.016
GSU2494			-	30717	cytochrome c, 10 heme-binding sites	Energy metabolism	-2.63	0.000362		0.0668	-5.149	0.00906
GSU2495			-	30717	cytochrome c, 26 heme-binding sites	Energy metabolism	-3.003	0.000155		0.046	-4.685	0.00188
GSU2496			-	30717	conserved hypothetical protein	No Data	-2.335	0.000737		0.994	-2.338	0.00582
GSU2504		omcS	-	GSU2504	cytochrome c, 6 heme-binding sites	Energy metabolism		0.0937	-2.042	0.0211	-2.485	0.0086
GSU2513			+	GSU2513	cytochrome c, 1 heme-binding site	Energy metabolism		0.194		0.0536	-2.449	0.0158
GSU2568			-	30733	MiaB-like tRNA modifying enzyme	Protein synthesis		0.00423	4.609	0.0432	6.995	0.0185
GSU2569		trmU	-	30733	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	Protein synthesis		0.00581	4.541	0.0362	6.39	0.018
GSU2570			-	30733	cysteine desulfurase	Biosynthesis of cofactors, prosthetic groups, and carriers		0.00171	4.535	0.0349	6.672	0.0159
GSU2571			-	30733	Rrf2 family transcriptional regulator	Unknown function		0.0289	5.841	0.031	6.902	0.0177
GSU2572		cysE	-	30733	serine O-acetyltransferase	Amino acid biosynthesis		0.00838	5.108	0.0276	6.793	0.0147
GSU2622			+	GSU2622	HAMP domain/GAF domain/HD domain phosphohydrolase, putative heme-binding site	Unknown function	-2.131	0.00154		0.733		0.0553
GSU2664			-	30755	efflux pump, RND family, inner and outer membrane proteins	Transport and binding proteins		0.711	2.818	0.0332	2.902	0.0226
GSU2665			-	30755	efflux pump, RND family, membrane fusion protein	Transport and binding proteins		0.326	2.534	0.0492	2.819	0.0268

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSU2678	Y		+	GSU2678	ATP-independent chaperone, alpha-crystallin/Hsp20 family	Protein fate	7.685	0.00336	11.288	0.0266	86.752	0.00225
GSU2727			+	GSU2727	conserved hypothetical protein	No Data	3.123	0.000644		0.354		0.602
GSU2730			-	GSU2730	conserved hypothetical protein	No Data		0.00529		0.319	-2.06	0.0399
GSU2731		omcC	-	30768	membrane-associated cytochrome c, 12 heme-binding sites	Energy metabolism		0.00342		0.105	-8.854	0.0469
GSU2732			-	30768	cytochrome c, 8 heme-binding sites	Energy metabolism		0.000585		0.0729	-9.471	0.0348
GSU2737		omcB	-	30769	membrane-associated cytochrome c, 12 heme-binding sites	Energy metabolism	-2.796	0.00295		0.0544	-5.893	0.00655
GSU2738			-	30769	cytochrome c, 8 heme-binding sites	Energy metabolism		0.000643		0.0694	-9.791	0.0305
GSU2739			-	30769	conserved hypothetical protein	No Data		0.00147		0.0859	-7.805	0.046
GSU2742			-	30770	conserved hypothetical protein	No Data		0.0119	-2.884	0.0352	-2.158	0.0457
GSU2743			-	30770	cytochrome c, 1 heme-binding site	Energy metabolism		0.11	-4.556	0.0185	-3.831	0.0131
GSU2780			-	GSU2780	conserved hypothetical protein	No Data		0.551	2.562	0.0244	2.71	0.0127
GSU2808			+	GSU2808	cytochrome c, 6 heme-binding sites	Energy metabolism	-2.247	0.00148		0.481	-2.753	0.0405
GSU2812			+	30784	glutaredoxin family protein	Energy metabolism		0.0052	-2.228	0.0407	-2.948	0.0155
GSU2875		rpsI	-	30793	ribosomal protein S9	Protein synthesis		0.5	2.004	0.0302	2.186	0.0129
GSU2886.1			-	GSU2886.1	cytochrome c, 7 heme-binding sites		-5.253	0.000591		0.274	-9.27	0.0206
GSU2887			-	GSU2887	cytochrome c, 27 heme-binding sites	Energy metabolism	-5.643	0.000386		0.191	-10.652	0.0136
GSU2894			+	GSU2894	conserved hypothetical protein	Hypothetical proteins	-2.45	0.000616		0.2	-3.05	0.00731
GSU2896			-	30798	ankyrin-related protein	Unknown function	-2.15	0.00129		0.313	-3.179	0.0471
GSU2897	Y		-	30799	conserved hypothetical protein	No Data	-2.904	0.0084		0.195		0.0552
GSU2898	Y	omcN	-	30799	cytochrome c, 27-34 heme-binding sites	Energy metabolism	-2.361	0.000478		0.516	-2.728	0.023
GSU2899	Y		-	30799	cytochrome c, 16-23 heme-binding sites	Energy metabolism	-2.186	0.000631		0.629	-2.384	0.0215
GSU2936			-	30809	conserved hypothetical protein	No Data		0.0831	-9.593	0.00529	-12.314	0.000647
GSU2937			-	30809	cytochrome c, 5 heme-binding sites	Energy metabolism		0.00656	-8.704	0.00529	-14.907	0.000797
GSU2938	Y		-	30810	conserved hypothetical protein	No Data		0.999	-7.034	0.00701	-7.035	0.00299
GSU2939	Y		-	30810	putative porin	Transport and binding proteins		0.00259	-8.793	0.00495	-13.683	0.000931
GSU2940	Y		-	30810	rhodanese-related sulfurtransferase	Cell envelope		0.00577	-9.048	0.00467	-14.414	0.000868
GSU2944			+	30811	(R)-2-hydroxyglutaryl-CoA dehydratase D component-related protein	Energy metabolism		0.0157	-2.341	0.0268	-2.723	0.0134
GSU2982	Y		-	30821	TonB-dependent outer membrane receptor, putative	Transport and binding proteins		0.0302	-2.689	0.00848	-2.053	0.0131
GSU3057		fnpA	-	30838	NADPH ferredoxin oxidoreductase (FNOR) alpha subunit	Amino acid biosynthesis		0.00116		0.0968	3.022	0.0226
GSU3059			-	GSU3059	radical SAM domain protein	Unknown function		0.625	6.299	0.0062	6.53	0.00235
GSU3060			+	GSU3060	transcriptional regulator, TetR family	Regulatory functions		0.812		0.0538	5.605	0.04
GSU3061		shc-2	+	GSU3061	squalene cyclase domain protein	Fatty acid and phospholipid metabolism		0.0176	11.795	0.0203	10.403	0.0128
GSU3122			-	GSU3122	metallo-beta-lactamase family protein	Unknown function	-2.002	0.00109		0.12		0.171

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1		2		3	
							Fold Δ	P-value	Fold Δ	P-value	Fold Δ	P-value
GSU3123			-	30851	aldehyde ferredoxin oxidoreductase domain protein	Unknown function	-2.186	0.00213		0.207		0.0535
GSU3125			-	30851	alcohol dehydrogenase, zinc-containing	Energy metabolism	-2.599	0.00174		0.159		0.548
GSU3126			-	GSU3126	oxidoreductase, aldo/keto reductase family	Unknown function	-2.886	0.000363		0.144		0.414
GSU3140			-	30854	conserved hypothetical protein	Protein fate		0.0304		0.0564	2.145	0.0268
GSU3141			-	GSU3141	conserved hypothetical protein	No Data		0.00338	5.57	0.0372		0.0556
GSU3142		aroG-2	-	GSU3142	3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase	Amino acid biosynthesis		0.00399	5.005	0.0497		0.0817
GSU3187		frx-6	+	GSU3187	ferredoxin family protein	Energy metabolism	2.009	0.00808	-2.648	0.0312		0.244
GSU3188			+	GSU3188	rubredoxin	Energy metabolism	2.031	0.0072		0.0859		0.353
GSU3241			-	30878		No Data	2.482	0.00149		0.39	2.188	0.0154
GSU3242			+	GSU3242	conserved hypothetical protein, frameshifted	No Data	-8.774	0.00147		0.528	-11.033	0.0106
GSU3265			+	30881	sulfite reductase, assimilatory-type	Central intermediary metabolism		0.000525		0.205	-2.15	0.0354
GSU3266		uvrD	-	GSU3266	UvrD/REP helicase	DNA metabolism	5.249	0.00343		0.0609	16.729	0.00344
GSU3267	Y		-	30882	conserved hypothetical protein	No Data	26.247	0.00155		0.0631	110.038	0.0019
GSU3268	Y	feoB-2	-	30882	ferrous iron transport protein B, putative	Transport and binding proteins	43.425	0.00183	5.384	0.0478	233.83	0.000678
GSU3269	Y		-	GSU3269		Transport and binding proteins	55.693	0.000991	4.432	0.0407	246.882	0.000561
GSU3270	Y		-	GSU3270	FeoA family protein	Unknown function	31.358	0.000701	7.514	0.00874	235.653	0.000279
GSU3271	Y		-	GSU3271	carbohydrate-selective porin OprB	No Data	84.345	0.000359	3.174	0.0252	267.746	1.38E-05
GSU3272	Y		-	30883	hypothetical protein	No Data	44.416	0.00211	4.986	0.0172	221.504	0.00239
GSU3273	Y		-	30883	conserved hypothetical protein	No Data	34.254	0.00051	5.44	0.018	186.351	0.000235
GSU3274	Y		-	30883	cytochrome c, 1 heme-binding site	Energy metabolism	34.065	0.000263	5.43	0.00742	184.986	1.35E-05
GSU3302			+	30892	methylmalonyl-CoA mutase, isobutyryl-CoA mutase-like catalytic subunit	Energy metabolism	-2.116	0.000255		0.398		0.0528
GSU3303			+	30892	glyoxalase/bleomycin resistance protein/dioxygenase family protein	Energy metabolism	-2.048	0.00344		0.912		0.166
GSU3311			+	GSU3311	sensor globin domain protein	No Data		0.016		0.0772	2.271	0.0236
GSU3329			-	30897	predicted DNA-binding protein with helix-hairpin-helix motif	Unknown function		0.0614		0.0782	2.064	0.045
GSU3340		groEL	+	30900	chaperonin GroEL	Protein fate		0.0546	2.186	0.0355		0.0251
GSU3370			-	GSU3370	transcriptional regulator, GntR family	Regulatory functions	-2.856	0.000895		0.0366	-3.701	0.00157
GSU3398			+	30910	efflux pump, RND family, outer membrane protein	Transport and binding proteins		0.0742	5.837	0.00565	6.273	0.00154
GSU3399			+	30910	efflux pump, RND family, membrane fusion protein	Transport and binding proteins		0.345	6.611	0.00493	7.143	0.00227
GSU3400			+	30910	heavy metal efflux pump, RND family, inner membrane protein, CzcA family	Cellular processes		0.0407	6.834	0.00618	7.744	0.00156

Table A2: List of all differentially expressed genes, with fold change and P-value. 1 = "sufficient"/"high," 2 = "limited"/"sufficient," 3 = "limited"/"high."

Gene	Fur controlled?	Gene name	Strand	DOOR	Gene product	Cellular function	1 Fold Δ	1 P-value	2 Fold Δ	2 P-value	3 Fold Δ	3 P-value
GSU3435			-	30919	ankyrin repeat protein	Unknown function	2.115	0.00191		0.467		0.39
tRNA-Arg-2			+	tRNA-Arg-2			2.543	0.0157		0.197		0.823
tRNA-Arg-4			+	tRNA-Arg-4			2.936	0.00558		0.177		0.649
tRNA-Asp-1			-	tRNA-Asp-1			2.333	0.0327	-6.504	0.0322		0.105
tRNA-Cys-1			-	tRNA-Cys-1			5.366	0.0072		0.107		0.764
tRNA-Glu-2			+	30894			2.08	0.014		0.744		0.313
tRNA-Gly-1			+	tRNA-Gly-1			2.527	0.00414		0.0664		0.355
tRNA-Gly-2			-	30657			9.043	0.00311	-14.3	0.0189		0.322
tRNA-Ile-1			+	30328			2.092	0.00831		0.209		0.0553
tRNA-Ile-2			+	30435			2.07	0.000965		0.0859		0.0351
tRNA-Leu-2			-	tRNA-Leu-2			3.367	0.0124		0.143		0.12
tRNA-Leu-3			-	tRNA-Leu-3			2.445	0.00922		0.169		0.79
tRNA-Lys-1			+	30894			2.193	0.00588		0.207		0.831
tRNA-Met-1			-	tRNA-Met-1			3.781	0.0157		0.111		0.804
tRNA-Met-2			-	tRNA-Met-2			2.055	0.015		0.274		0.862
tRNA-Phe-1			+	tRNA-Phe-1			2.12	0.0469		0.124		0.929
tRNA-Pro-1			+	tRNA-Pro-1				0.571		0.0681	-2.908	0.0342
tRNA-Ser-3			+	tRNA-Ser-3			2.248	0.00418		0.649		0.668
tRNA-Ser-5			-	tRNA-Ser-5			7.158	0.00212	-9.102	0.0169		0.46
tRNA-Val-3			-	30572			2.208	0.0427		0.2		0.939

REFERENCES

1. **Lovley, D.R., J.F. Stolz, G.L. Nord Jr, and E.J.P. Phillips.** 1987. Anaerobic production of magnetite by a dissimilatory iron-reducing microorganism. Nature 330**(6145)**:252-254.

2. **Lovley, D.R.** Chapter 7: Reduction of iron and humics in subsurface environments. <u>Subsurface Microbiology and Biogeochemistry.</u> 2001.

3. **Bond, D.R., D.E. Holmes, L.M. Tender, and D.R. Lovley.** 2002. Electrode-reducing microorganisms that harvest energy from marine sediments. Science 295**(5554)**:483-485.

4. **Lovley, D.R., D.E. Holmes, and K.P. Nevin.** 2004. Dissimilatory Fe(III) and Mn(IV) reduction. Advances in Microbial Physiology 49.

5. **Anderson, R.T., J.N. Rooney-Varga, C.V. Gaw, and D.R. Lovley.** 1998. Anaerobic benzene oxidation in the Fe(III) reduction zone of petroleum-contaminated aquifers. Environ Sci Technol**(32)**:1222-1229.

6. **Holmes, D.E., K.T. Finneran, R.A. O'Neil, and D.R. Lovley.** 2002. Enrichment of members of the family *Geobacteraceae* associated with stimulation of dissimilatory metal reduction in uranium-contaminated aquifer sediments. Appl Environ Microbiol 68**(5)**:2300-2306.

7. **Lovley, D.R.** 2006. Bug juice: Harvesting electricity with microorganisms. Nature Reviews Microbiology 4**(7)**:497-508.

8. **Yi, H., K.P. Nevin, B.C. Kim, A.E. Franks, A. Klimes, L.M. Tender, and D.R. Lovley.** 2009. Selection of a variant of *Geobacter sulfurreducens* with enhanced capacity for current production in microbial fuel cells. Biosens Bioelectron 24**(12)**:3498-3503.

9. **Methé, B.A., K.E. Nelson, J.A. Eisen, I.T. Paulsen, W. Nelson, J.F. Heidelberg, D. Wu, M. Wu, N. Ward, M.J. Beanan, R.J. Dodson, R. Madupu, L.M. Brinkac, S.C. Daugherty, R.T. DeBoy, A.S. Durkin, M. Gwinn, J.F. Kolonay, S.A. Sullivan, D.H. Haft, J. Selengut, T.M. Davidsen, N. Zafar, O. White, B. Tran, C. Romero, H.A. Forberger, J. Weidman, H. Khouri, T.V. Feldblyum, T.R. Utterback, S.E. Van Aken, D.R. Lovley, and C.M. Fraser.** 2003. Genome of *Geobacter sulfurreducens*: metal reduction in subsurface environments. Science 302**(5652)**:1967.

10. **Coppi, M.V., C. Leang, S.J. Sandler, and D.R. Lovley.** 2001. Development of a genetic system for *Geobacter sulfurreducens*. Appl Environ Microbiol 67**(7)**:3180-3187.

11. **Caccavo Jr, F., D.J. Lonergan, D.R. Lovley, M. Davis, J.F. Stolz, and M.J. McInerney.** 1994. *Geobacter sulfurreducens* sp. nov., a hydrogen- and acetate-oxidizing dissimilatory metal-reducing microorganism. Appl Environ Microbiol 60**(10)**:3752-3759.

12. **Ding, Y.H., K.K. Hixson, C.S. Giometti, A. Stanley, A. Esteve-Núñez, T. Khare, S.L. Tollaksen, W. Zhu, J.N. Adkins, M.S. Lipton, R.D. Smith, T. Mester, and D.R. Lovley.** 2006. The proteome of dissimilatory metal-reducing microorganism *Geobacter sulfurreducens* under various growth conditions. Biochim Biophys Acta 1764**(7)**:1198-1206.

13. **Stevens, J.M., Daltrop, O., Allen, J.W.A., and S.J. Ferguson.** 2004. C-type cytochrome formation: chemical and biological enigmas. Acc Chem Res 37**(12)**:999-1007.

14. **Nevin, K.P., and D.R. Lovley.** 2000. Lack of production of electron-shuttling compounds or solubilization of Fe(III) during reduction of insoluble Fe(III) oxide by *Geobacter metallireducens*. Appl Environ Microbiol 66**(5)**:2248-2251.

15. **Liang, S., Squier, T.C., Zachara, J.M., and J.K. Fredrickson.** 2007. Respiration of metal (hydr)oxides by *Shewanella* and *Geobacter:* a key role for multihaem *c*-type cytochromes. Mol Microbiol 65(**1**):12-20.

16. **., Serruto, D., Parkhill, J., Relman, D.A., Donati, C., Moxon, R., Falkow, S., and Rino Rappuoli.** 2008. Microbiology in the post-genomic era. Nat Rev Microbiol 6(6):419-30.

17. **Cho, B.K., Zengler, K., Qiu, Y., Park, Y.S., Knight, E.M., Barrett, C.L., Gao, Y., and Bernhard O Palsson.** 2009. The transcription unit architecture of the *Escherichia coli* genome. Nat Biotech 27**(11)**: 1043-1049.

18. **Qiu, Y., Cho, B.K., Park, Y.S., Lovely, D., Palsson, B.O., and Karsten Zengler.** 2010. Structural and operational complexity of the *Geobacter sulfurreducens* genome. Unpublished manuscript.

19. Sigma factor network in Geobacter. Unpublished manuscript.

20. **Escolar, L., Perez-Martin, J., and Victor de Lorenzo.** 1999. Opening the iron box: Transcriptional metalloregualtion by the Fur protein. J Bacteriol 181**(20):**6223-6229.

21. **Lee, J.W., and John D. Helmann.** 2007. Functional specialization within the Fur family of metalloregulators. Biometals 20:485-499.

22. **Touati, D.** 2000. Iron and oxidative stress in bacteria. Arch Biochem Biophys 373:1-6.

23. **O'Neil, R., Coppi, M., Nevin, K., Woodard, T., Holmes, D., Methe, B., Krushka, J., Yan, B., qian, Y., Rodionov, D., Liu, A., and Derek Lovely.** Investigation of the Fur regulon of *Geobacter sulfurreducens.* Unpublished manuscript.

24. **Bell, P. E., A. L. Mills, and J. S. Herman.** 1987. Biogeochemical Conditions Favoring Magnetite Formation during Anaerobic Iron Reduction. Applied and Environmental Microbiology **53:**2610-2616.

25. **Lovley, D. R.** 2000. Fe(III) and Mn(IV) Reduction, p. 3-30. *In* D. R. Lovley (ed.), Environmental Metal-Microbe Interactions. ASM Press, Washington, D.C. .

26. **Lovley, D. R., and E. J. P. Phillips.** 1987. Rapid Assay for Microbially Reducible Ferric Iron in Aquatic Sediments. Applied and Environmental Microbiology **53:**1536-1540.

27. **Cho, B.K., Knight, E.M., and Bernhard O. Palsson.** 2008. Genomewide identification of protein binding locations using chromatin immunoprecipitation coupled with microarray. Methods Mol Biol 439:131-145.

28. **Perocchi, F., Xu, Z., Clauder-Munster, S., and Lars M. Steinmetz.** 2007. Antisense artifacts in transcriptome microarray experiments are resovled by anctinomycin D. Nucleic Acids Res 35(**19**): e128.

29. **Rudolph, G., Hennecke, H., and Hans-Martin Fischer.** 2006. Beyond the Fur paradigm: iron-controlled gene expression in rhizobia. FEMS Microbiol Rev 30: 631-648.

30. **Ernst, F.D., Bereswill, S., Waidner, B., Stoof, J., Mader, U., Kusters, J.G., Kuipers, E.J., Kist, M., van Vliet, A.H.M., and Georg Homuth.** 2005. Transcriptional profiling of *Heliobacter pylor* Fur- and iron-regulated gene expression. Microbiology 151: 533-546.

31. **Delaney, I., Spohn, G., Rappuoli, R., and V. Scarlato.** 2001. The Fur repressor controls transcription of iron-ativated and –repressed genes in *Heliobacter pylori.* Mol. Microbiol. 42: 1297-1311.

32. **Carton, M.L., Maddocks, S., Gillingham, P., Crave, C.J., and Simon C. Andrews.** 2006. Feo—Transport of ferrous iron into bacteria. BioMetals 19: 143-157.

33. **Fontecilla-Camps, J.C., Volbeda, A., Cavazza, C., Nicolet Y.** 2007. Structure/function relationships of [NiFe]- and [FeFe]-hydrogenases. Chem Rev 107(10): 4273–4303.

34. **Shi, L., Squier, T.C., Zachara, J.M., and James K. Fredrickson.** 2007. Respiration of metal (hydr)oxides by *Shewanella* and *Geobacter:* a key role for multihaem c-type cytochromes. Mol. Microbiol. 65(1):12-20.

35. **O'Neil, R.A., D.E. Holmes, M.V. Coppi, L.A. Adams, M.J. Larrahondo, J.E. Ward, K.P. Nevin, T.L. Woodard, H.A. Vrionis, A.L. N'Guessan, and D.R. Lovley.** 2008. Gene transcript analysis of assimilatory iron limitation in *Geobacteraceae* during groundwater bioremediation. Environ Microbiol 10**(5)**:1218-1230.

36. **Theil, E.C.** 1987. Ferritin: Structure, gene regulation, and cellular function, in animals, plants, and microorganisms. Ann. Rev. Biochem 56:289-315.

37. **Krewulak, K.D.** 2007. Structural biology of bacterial iron uptake. Biochimica et Biophysica Acta 1778: 1781-1804.

38. **Methé, B.A., K. Webster, K.P. Nevin, J.E. Butler, and D.R. Lovley.** 2005. DNA microarray analysis of nitrogen fixation and Fe(III) reduction in *Geobacter sulfurreducens.* Appl Environ Microbiol 71**(5)**:2530-2538.

39. **Stojiljkovic, I., Baumler, A.J., and K. Hantke.** 1994. Fur regulon in Gram-negative bacteria: Identification and characterization of new iron-regulated *Escherichia coli* genes by a Fur titration assay. J. Mol. Biol. 236:531-545.

40. **Lavrrar, J.L. and M.A. McIntosh.** 2003. Architecture of a Fur binding site: a comparative analysis. J. Bacteriol. 185(7):2194-2202.

41. **Narberhause, F.** 2002. Alpha-crystallin-type heat shock proteins: Socializaing minichaperones in the context of a multichaperone network. Microbiol Mol Biol Rev. 66(1):64-93.

42. Miles S., Carpenter B.M., Gancz, H., and D.S. Merrel. 2010. Helicobacter pylori apo-Fur regulation appears unconserved across species. J. Microbiol. 48(3):378-386.

43. Robison, K., A. M. McGuire, and G. M. Church. 1998. A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete Escherichia coli K-12 genome. J Mol Biol **284:**241-54.

44. Rodionov, D. A., I. Dubchak, A. Arkin, E. Alm, and M. S. Gelfand. 2004. Reconstruction of regulatory and metabolic pathways in metal-reducing delta-proteobacteria. Genome Biol **5:**R90.

45. **Krushkal, J., B. Yan, L. N. DiDonato, M. Puljic, K. P. Nevin, T. L. Woodard, R. M. Adkins, B. A. Methe, and D. R. Lovley.** 2007. Genome-wide expression profiling in Geobacter sulfurreducens: identification of Fur and RpoS transcription regulatory sites in a relGsu mutant. Funct Integr Genomics **7:**229-55.

46. Thompson DK, Beliaev AS, Giometti CS, Tollaksen SL, Khare T, Lies DP, Nealson KH, Lim H, Yates J 3rd, Brandt CC, Tiedje JM, and J Zhou. 2002. Transcriptional and proteomic analysis of a ferric uptake regulator (fur) mutant of *Shewanella oneidensis*: possible involvement of fur in energy metabolism, transcriptional regulation, and oxidative stress. Appl Environ Microbiol. 68(2):881-892.

47. Wan XF, Verberkmoes NC, McCue LA, Stanek D, Connelly H, Hauser LJ, Wu L, Liu X, Yan T, Leaphart A, Hettich RL, Zhou J, and DK Thompson. 2004. Transcriptomic and proteomic characterization of the Fur modulo in the metal-reducing bacterium *Shewanella oneidensis.* J Bacteriol. 186(24):8385-8400.

48. Yang Y, Harris DP, Luo F, Wu L, Parsons AB, Palumbo AV, Zhou J. 2008. Characterization of the *Shewanella oneidensis* Fur gene: roles in iron and acid tolerance response. BMC Genomics. 9 Suppl 1:S11.

49. Bender KS, Yen HC, Hemme CL, Yang Z, He Z, He Q, Zhou J, Huang KH, Alm EJ, Hazen TC, Arkin AP, Wall JD. 2007. Analysis of a ferric uptake regulator (Fur) mutant of *Desulfovibrio vulgaris* Hildenborough. Appl Environ Microbiol. 73(17): 5389-5400.

50. Jackson, L.A., Ducey, T.F., Day, M.W., Zaitshik, J.B., Orvis, J., and D.W. Dyer. 2010. Transcriptional and functional analysis of the *Neisseria gonorrhoeae* Fur regulon. J. Bacteriol.192(1): 77-85.

51. Danielli, A., and V. Scarlato. 2010. Regulatory circuits in *Helicobacter pylori:* network motifs and regulators involved in metal-dependent responses. FEMS Microbiol Rev. Epub ahead of print.

52. Danielli A., Roncarati, D., Delany, I., Charini, V., Rappuoli, R., and V. Scarlato. 2006. In vivo dissection of the *Helicobacter pylori* Fur regulatory circuit by genome-wide location analysis. J. Bacteriol. 188(13):4564-4662.

53. Wade, J.T., Reppas, N.B., Church, G.M., and K. Struhl. 2005. Genomic analysis of LexA binding reveals the permissive nature of the *Escherchia coli* genome and identifies unconventional target sites. Gene Dev 19:2619-2630.