Doctoral Dissertations

Graduate School

12-2007

# Integrating Mass Spectrometry Based Proteomics and Bioinformatics Technologies for the Molecular Level Characterization of *Shewanella oneidensis* to Chromate Exposure

Melissa Renee Thompson
*University of Tennessee - Knoxville*

To the Graduate Council:

I am submitting herewith a dissertation written by Melissa Renee Thompson entitled "Integrating Mass Spectrometry Based Proteomics and Bioinformatics Technologies for the Molecular Level Characterization of *Shewanella oneidensis* to Chromate Exposure." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Life Sciences.

Robert L. Hettich, Major Professor

We have read this dissertation and recommend its acceptance:

Dorothea Thompson, Elizabeth Howell, Loren Hauser, Frank Larimer

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Melissa Renee Thompson entitled "Integrating Mass Spectrometry Based Proteomics and Bioinformatics Technologies for the Molecular Level Characterization of *Shewanella oneidensis* to Chromate Exposure". I have examined the final electronic copy of this dissertation for form and content and recommended that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Life Sciences.

<div align="right">

Robert L. Hettich  _
Major Professor

</div>

We have read this dissertation
and recommended its acceptance:

Dorothea Thompson   _

Elizabeth Howell    _

Loren Hauser    _

Frank Larimer    _

<div align="right">

Accepted for the Council:

_ Carolyn R. Hodges _
Vice Provost and
Dean of the Graduate School

</div>

# INTEGRATING MASS SPECTROMETRY BASED PROTEOMICS AND BIOINFORMATICS TECHNOLOGIES FOR THE MOLECULAR LEVEL CHARACTERIZATION OF *SHEWANELLA ONEIDENSIS* TO CHROMATE EXPOSURE

A Dissertation Presented for the Doctor of Philosophy Degree
The University of Tennessee, Knoxville

Melissa Renee Thompson
December, 2007

# DEDICATION

This dissertation is dedicated to everyone who helped me along the way from childhood to adulthood.  I specifically wish to thank my husband, Brad, who has supported me during my entire graduate school career and sacrificed a great deal for me.  I also wish to thank my family Bob and Brenda Rugen, Michelle Rugen, Brett and Kristin Rugen, H.C. Rugen, Jr., Ruth Dove, Gertrude Connor, and Lendell and Janet Thompson.  To Seymore and Putty, who are the best support and distraction after a long day in the laboratory or writing at the computer.  I would like to dedicate this work in memory of my grandfather Charles Dove and grandmother Maxine Dove.  Finally, I would like to dedicate this work in memory of my grandmother, Eleanor Rugen, who was a great source of encouragement of my dreams and goals during her life.

# ACKNOWLEDGMENT

# ABSTRACT

The research outlined in this dissertation involves the development and demonstration of a mass spectrometry-based proteomics approach to characterize the global level molecular response of *Shewanella oneidensis* MR-1 to chromate exposure. The proteomics approach is centered on a high performance technique of multi-dimensional on-line liquid chromatographic separations with subsequent tandem mass spectrometric detection. Since very complex proteome samples are digested into peptides and then directly measured by MS, this technique is termed shotgun proteomics. This approach affords the identification and quantification of complex mixtures by directly analyzing their proteolytic peptides and then using computational techniques to reassemble the protein information. The research goals for this dissertation project were two-fold: (1) enhancement of the experimental and computational methodologies to permit deeper and more confident proteome characterizations, and (2) demonstration of this optimized approach for the comprehensive investigation of the molecular level response of the bacterium *S. oneidensis* to chromate insult. To address research needs, we developed a single-tube lysis method for cell lysis-proteome digestion to enable investigations of small amounts of cellular biomass, and identified suitable bioinformatic approaches to mine post-translational modifications from proteome datasets. These advancements were then utilized to examine the molecular level response of *S. oneidensis* to chromate insult, which was accomplished by varying chromate concentrations, dosages, and time points. These measurements provided the first global proteome-level observation of the dynamic changes of *S. oneidensis* in response to chromate insult.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF SYMBOLS AND ABBREVIATIONS

| | |
|---|---|
| µg | Micrograms |
| µL | Microliter |
| µm | Micrometer |
| 2-D | Two dimensional |
| 2D-PAGE | Two-dimensional polyacrylamide gel electrophoresis |
| Å | Angstrom |
| AC | Alternating current |
| ACN | Acetonitrile |
| AFM | Atomic Force Microscopy |
| AMD | Acid mine drainage |
| AMT | Accurate mass and time |
| AQUA | Absolute quantification isotopically labelled peptides |
| b ion | Fragmentation of a peptide from the N-terminus between the carboxy and amide bonds |
| BCA | Bicinchoninic acid solution |
| $CaCl_2$ | Calcium chloride |
| CID | Collision induced dissociation fragmentation method |
| cm | Centimeter |
| Cr | Chromium |
| Cr(III) | Trivalent chromium |
| Cr(V) | Pentavalent chromium |
| Cr(VI) | Hexavalent chromium |
| $Cr_2O_7^{2-}$ | Dichromate |
| $CrO_4^{2-}$ | Chromate |
| Da | Dalton |
| DC | Direct current |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| DOE | Department of Energy |
| DOM | Dissolved organic matter |
| DPC | Diphenylcarbazide |
| DTA | Flat file of MS/MS spectrum converted from RAW file |
| DTT | Dithiothreitol |
| EDTA | Ethylenediaminetetraacetic acid |
| ESI | Electrospray ionization |
| FA | Formic acid |
| FASTA | Text file containing a nucleotide or amino acid sequence |
| FDR | False discovery rate |
| Fe(II) | Divalent iron |
| Fe(III) | Trivalent iron |
| FT-ICR | Fourier transform ion cyclotron resonance |
| FTICR MS | Fourier transform ion cyloctron resonance mass spectrometry |

| | |
|---|---|
| g | Gravity |
| h or hr | Hour |
| HCl | Hydrogen chloride |
| $H_2O$ | Chemical formula for water |
| HPLC | High performance liquid chromatography |
| ICAT | Isotope coded affinity tags for peptide/protein quantification |
| ID | Internal diameter or identification |
| IEF | Isoelectric focusing |
| iTRAQ | Isobaric amine-reactive tagging reagent |
| $K_2CrO_4$ | Potassium chromate |
| KEGG | Kyoto encyclopedia of genes and genomes |
| kDa | Kilodaltons |
| kV | Kilovolts |
| L | Liter |
| LB | Luria-Bertani medium |
| LC | Liquid chromatography |
| LC/LC-MS/MS | 2-dimensional liquid chromatography tandem mass spectrometry |
| LC-MS | Liquid chromatography-mass spectrometry |
| LC-MS/MS | Liquid chromatography-tandem mass spectrometry |
| LCQ | Liquid chromatography quadrupole ion trap, 3-D ion trap |
| LTQ | Linear trapping quadrupole |
| LTQ-FTICR | Linear trapping quadrupole Fourier transform ion cyclotron resonance hybrid mass spectrometer |
| LTQ-MS | Linear trapping quadrupole mass spectrometry |
| M | Molarity |
| m/z | Mass to charge ratio |
| MALDI | Matrix assisted laser desorption ionization |
| MALDI-MS | Matrix assisted laser desorption ionization mass spectrometry |
| MASPIC | DBDigger scorer |
| mg | Milligrams |
| MIC | Minimum inhibitory concentration |
| min | Minute |
| mL | milliLiters |
| mM | Millimolar |
| Mn(IV) | Tetravalent manganese |
| MS | Mass spectrometry |
| MS/MS | Tandem mass spectrometry |
| ms2 | Flat file of MS/MS data converted from RAW file |
| MudPIT | Multidimensional protein identification technology |
| mzXML | Flat file of MS/MS data converted from RAW file |
| NA or N/A | Not available |
| NaCl | Sodium Chloride |
| nL | Nanoliter |
| OD | Optical density |
| ORF | Open reading frame |
| ORNL | Oak Ridge National Laboratory |

| | |
|---|---|
| pH | Measure of the acidity of a solution |
| pI | Isoelectric point |
| PMF | Peptide mass fingerprinting |
| ppm | Parts per million |
| PTM(s) | Post translational modification(s) |
| p-value | Probability value that MS/MS identification is true |
| QIT | Quadrupole ion trap |
| QTOF | Quadrupole time of flight |
| RF | Radio frequency |
| RNA | Ribonucleic acid |
| rpm | Revolutions per minute |
| ROS | Reactive oxygen species |
| RP | Reverse phase |
| s | Second |
| SCX | Strong cation exchange |
| SDS | Sodium dodecyl sulphate |
| SDS-PAGE | Sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| sqt | Results file of MS/MS data searched with database or de novo |
| TCA | Tricarboxylic acid cycle |
| TFE | Tetrafluoroethylene |
| TIC | Total ion current |
| TIGR | The Institute for Genome Research |
| TOF | Time of flight |
| Tris | Trishydroxymethylaminomethane |
| txt | Text file |
| U - V cos $\Omega$t | Equation for the DC and RF voltage of a quadrupole ion trap |
| UV | Ultraviolent radiation |
| w/w | Weight for weight |
| WT | Wild-type |
| Xcorr | SEQUEST cross-correlation score |
| y ion | Fragmentation of a peptide from the C-terminus between the carboxy and amide bonds |
| $\Delta$CN or DelCN or DeltaCN | Sequest delta correlation score |

**Chapter 1**

**Introduction to *Shewanella oneidensis* Microbiology with Respect to Chromate**

**Exposure using Mass Spectrometry Technology**


**<u>Introduction</u>**

Utilizing mass spectrometry technology to decipher global metabolic pathways

and regulatory networks has increased understanding of these cellular processes in

biological systems.  The application of mass spectrometry based proteomics falls under

the area of systems biology.  Systems biology as defined by Ideker *et al* [1] is the

integration of data from a number of technologies in order to build a comprehensive

model to predict cellular pathway responses in a variety of organisms.  This approach

generally utilizes metabolomic, genomic, transcriptomic, and proteomic datasets in some

fashion in order to catalog potential metabolic pathway members.  The proteomics

portion of systems biology encompasses three areas of emerging technology (1)

improvement of mass spectrometry instrumentation, (2) improvements in separation of

complex mixtures, and (3) continuous advancement in the computational software

utilized for mining the resulting datasets.  The research in this dissertation contributed to

a Department of Energy (DOE) project designed to utilize a systems biology approach

with the proteomics portion of the project presented in this dissertation.  The goal of this

dissertation is to understand the chromate response in *Shewanella oneidensis* (Figure 1.1)

for the purposes of bioremediation.

The use of proteomics based analyses began with two-dimensional

polyacrylamide gel electrophoresis (2-D PAGE) in 1975 [2].  This method separates
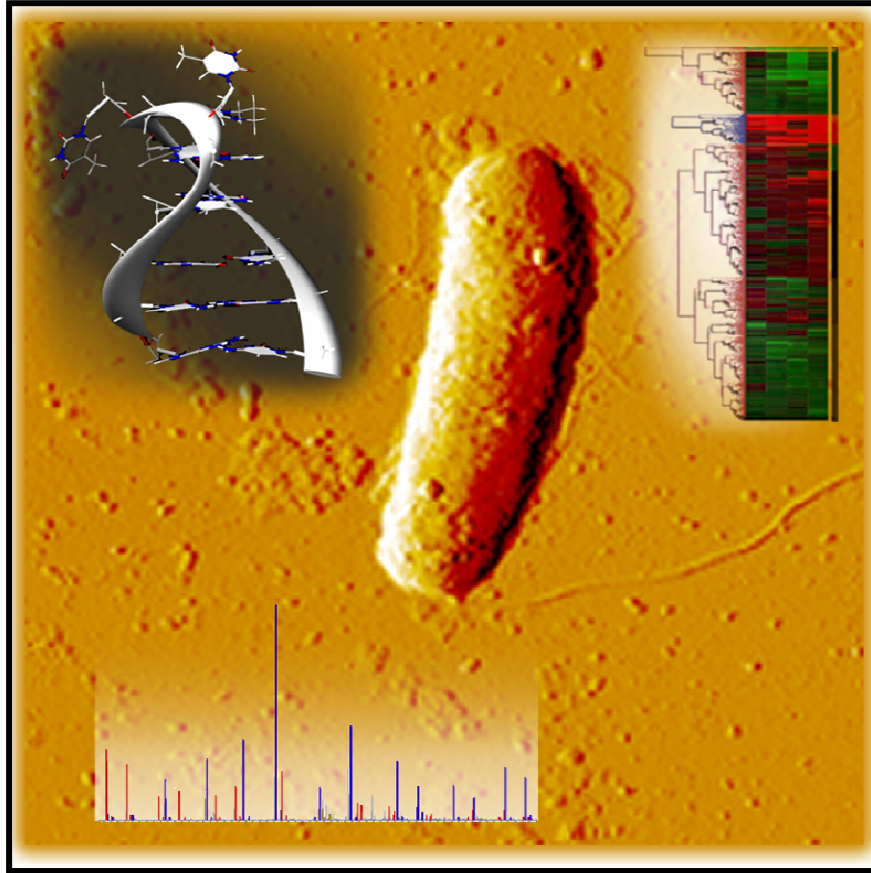
**Figure 1.1.** An integrated approach to understand chromate exposure in *S. oneidensis*. The dissertation comprised the proteomic portion of the project depicted by the tandem mass spectrum in the figure. The bacterial cell is *S. oneidensis* imaged by atomic force microscopy (AFM) (courtesy of K. Chourey). The double helix DNA strand depicts the availability of the genome sequence (courtesy of G. Wickham) and the RNA expression cluster depicts the transcriptomic portion of the project (courtesy of D. Thompson).

mixtures of proteins first by the isoelectric point of each individual protein and then by the molecular mass of the protein [2-4]. This initial work demonstrated the utility of separating complex mixtures of proteins in order to answer questions pertaining to the expression levels and subsequent metabolic pathway information of cellular proteins responding to a perturbation. The initial work with 2-D PAGE was limited by the inability to identify the differential spots on the gel easily. Identification of the protein species was accomplished initially by N-terminal Edman degradation [5, 6]. Edman degradation was developed by Edman and Begg [6] and operates by progressively cleaving the N-terminal amino acid residue from the protein. This residue is then separated chromatographically in order to determine its identity. The first application of Edman degradation required 5.0 mg of purified protein [6].

## Mass Spectrometry Based Proteomics

Mass Spectrometry was not originally amenable to working with more fragile biomolecules such as proteins due to limitations in the types of ionization sources used previously (i.e. chemical and electron ionization [7]). Electron ionization was the first ionization source created for mass spectrometric detection of organic molecules and was developed in the late 1920's [8]. This ionization method employs the use of electrons ejected off a heated filament which bombards a gas phase analyte that has been injected into the ionization space. This method of ionization works well for molecules that have a high vapor pressure [8]. The mechanism of chemical ionization occurs through the interaction of the sample analyte desorbed from a probe with ions that are produced in the source (i.e. a reagent gas) leading to a reduction in the fragmentation of the molecular

3

species [8].  This method of ionization was created to circumvent the harshness of

electron ionization, which fragmentation of the molecular species is common.

During the early 1990's, Hillenkamp *et al* [9] described a new, more gentle means

of creating ionized proteins, defined as matrix assisted laser desorption ionization

(MALDI).  Now, gel spots identified in the 2-D PAGE gels could be excised, the

protein(s) digested, and analyzed using MALDI-MS with a method designated as peptide

mass fingerprinting (PMF).  The number of intact unique peptide masses to be identified

for a given protein can be as few as three to four using PMF [10].  This method works

well for less complex mixtures such as gel spots, however, this methodology is not

amenable to complex mixtures of proteins without the initial 2-D PAGE separation of the

proteins.  Therefore, a gel-free methodology was developed due to the cumbersome

nature of 2-D PAGE.

In MALDI, the peptides/proteins are spotted onto a conductive metal plate and

dried within the embedded matrix that is ablated by a laser [9].  The laser causes

ionization of the matrix, which is generally an acid, creating positively charged ions by

desorbing both the matrix and the embedded analyte off the surface.  Following

desorption, there are ion-molecule reactions between the matrix (the ion) and the analyte

(the molecule), resulting in an ionized analyte.  The charged analyte can now be injected

into the mass analyzer for subsequent detection.  Primarily, MALDI ionization is

conducted under vacuum [8], however, there is an atmospheric pressure source available

that was developed by Laiko *et al* [11].

Electrospray is also a gentle ionization source similar to MALDI-MS, but the

peptides/proteins remain in solution.  Electrospray ionization (ESI) was first described

and developed by Fenn *et al* [12] as a method of atmospheric ionization for large

biomolecules such as DNA and proteins. ESI is a solution-based ionization method and

creates ions by applying high voltage (in the range of kilovolts) to a silica capillary

causing ionized droplets to be sprayed from the tip of the silica [8]. Due to Columbic

repulsion of the ionized analyte within the droplet, each droplet formed breaks apart into

smaller droplets causing desolvation of the ionized analyte and creating a Taylor cone.

With respect to electrospray ionization in mass spectrometry, the Taylor cone is a plume

of charged droplets emitted from the electrospray tip. The droplets form due to the

voltage difference between the electrospray tip and the opening of the mass spectrometer

[8]. The droplets undergo further desolvation within a heated metal capillary that

transfers the ions into the vacuum of the mass spectrometer [8] (Figure 1.2). This method

of ionization is easily coupled to a number of mass spectrometers as described below. In

addition, the primary advantage of this ionization technique for analysis of biomolecules

is the ability to couple this ionization source directly to the chromatographic separation of

complex mixtures of proteins/peptides as discussed below [13-16].

There are a number of mass spectrometer designs amenable to acquiring large

amounts of data on peptides or proteins with a relatively rapid data acquisition speed.

However, the mass spectrometer chosen for a given proteomics experiment is dependent

on the properties of the sample to be analyzed (intact proteins or digested peptides) and

the capabilities of the mass analyzer such as the ability to perform data-dependent

MS/MS, the mass resolution, the dynamic range, and the mass range. Data-dependent

MS/MS allows for an unbiased selection of ions to be isolated and subsequently

fragmented. This capability is necessary when attempting to catalog the members of a
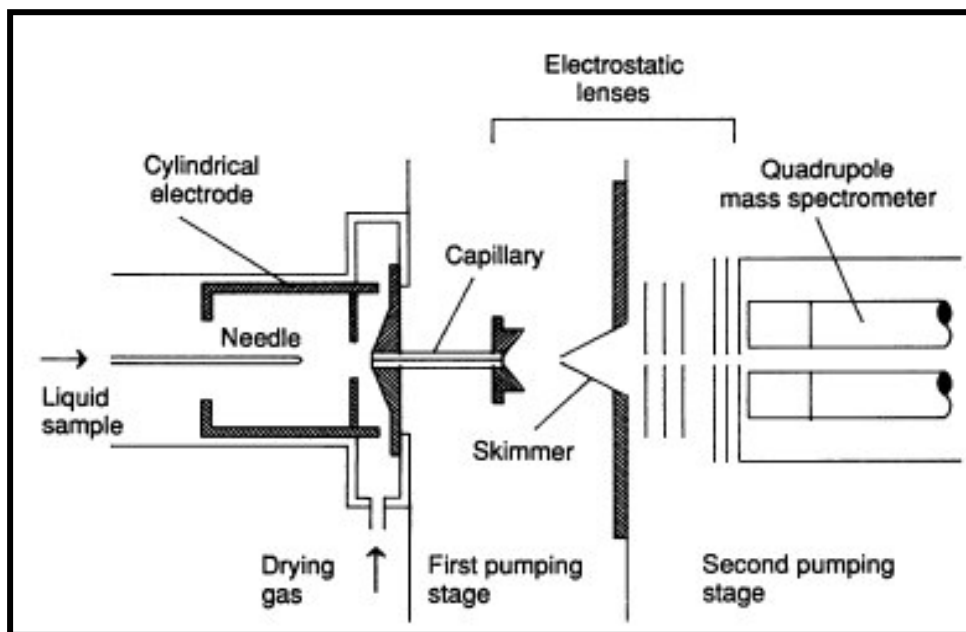
**Figure 1.2.** Original design of the electrospray ionization source.

Reprinted Figure 1 from Science, Vol. 246, Fenn *et al*, Electrospray ionization for mass

spectrometry of large biomolecules, pages 64-71, Copyright 1989 with permission from

AAAS.

complex mixture (i.e. a proteome).  The mass resolution for a mass analyzer is defined as

the peak width at half of the maximum height for that peak [17, 18].  The dynamic range

is the ability of an ion lower in concentration compared to another ion to be detected in

concurrence with the abundant ion within the same experiment [19, 20].  The mass range

of a mass analyzer is a continuum of $m/z$ values from lower $m/z$ to greater $m/z$ detected by

the mass spectrometer and determines the type of biomolecule, intact protein or digested

peptide, which can be utilized for an experiment.

Mass analyzers that have the ability to detect and analyze intact proteins include

the Fourier transform ion cyclotron resonance (FT-ICR) [21, 22], Orbitrap [23], and time

of flight (TOF) [24].  The benefits of using an FT-ICR include the high mass accuracy (in

the ppm range), high mass resolution, and ability to be coupled with chromatographic

separation [25, 26].  In addition, the FT-ICR has been coupled to the linear trapping

quadrupole (LTQ) to form a hybrid instrument with enhanced capabilities [22].  The

Orbitrap operates in a similar fashion as the FT-ICR, utilizing the ions' oscillation

frequency as the method of detection.  This mass analyzer differs from the FT-ICR in the

respect that the Orbitrap does not use a super-conducting magnet.  The Orbitrap is also

interfaced as a hybrid instrument with the LTQ, allowing high resolution full mass

spectra to be acquired in the Orbitrap and ion fragmentation with subsequent MS/MS to

be acquired in the LTQ.  The time of flight mass analyzer is also found as a hybrid

instrument with a quadrupole or an ion trap [27, 28].

Some mass analyzers do not have the capability to detect intact proteins, but are

very useful for peptide detection and analysis, including the three-dimensional ion trap

(LCQ) [29] and the LTQ [30].  The LTQ is a second generation design based on the

principles of the LCQ.  Improvements made by Stafford *et al* [29] allowed the LCQ to be

coupled readily with peptide detection.  The LCQ has been used in many studies on

shotgun proteomics due to it's capabilities of data-dependent MS/MS, dynamic range,

and data acquisition speed allowing the LCQ to be coupled to liquid chromatographic

separation.  However, improvements to the design of the ion trap led to the

commercialization of the LTQ [30].  The LTQ improves on the capabilities of the LCQ

with an even faster data acquisition speed leading to the detection of more than five times

as many protein identifications and four times as many MS/MS spectra acquired during

the same time period [31, 32].  The increased dynamic range observed with the LTQ in

comparison to the LCQ is due to the increased data acquisition speed and the increased

trapping volume leading to a larger ion population without the space-charging effects of

the three dimensional ion trap [30].

**Liquid Chromatographic Separation of Biomolecules**

The development of electrospray led to improvements in the second area of

proteomics:  separation of complex mixtures of peptides/proteins.  Electrospray allows

for online separation of peptides/proteins to be coupled to mass spectrometry.  This gel-

free method of separating complex mixtures is less time consuming, more sensitive, and

more reproducible than gel-based methods.  A large number of studies published in

proteomics utilize reverse phase (RP) separation online with mass spectrometry.  This is

due to the liquid phase used in RP separation consisting of an aqueous to organic

gradient.  These two types of solvents readily disperse into droplets that desolvate in the

heated capillary rapidly.  The addition of strong cation exchange (SCX) online with RP

was introduced by Washburn *et al* [33].  This allowed unbiased separation of peptides

first based on their affinity for the negatively charged resin of the SCX column followed

by separation based on hydrophobicity.  2-D PAGE separation is limited based on the

protein molecular weight (proteins > 180 kDa) and isoelectric point (pI within the range

of 5-10) [33].  The unbiased separation of peptides/proteins is required for global

detection studies like shotgun proteomics experiments.  Another form of separation,

isoelectric focusing (IEF), has emerged as an alternative to SCX as the second dimension

of separation [34, 35].  This form of separation is performed with either intact proteins or

peptides.  This is a gel-based method predominantly [34] utilizing IPG strips available

commercially, but can also be liquid-based with a commercially available apparatus (i.e.

the MicroRotofor cell from Bio-Rad, Hercules, CA).

**Advancements in Computational Biology**

Computational biology or algorithm development is an emerging area of

development due to the large advancements described above, however there is still more

progress that must be made in order increase the confidence of subsequent MS/MS data

analysis.  The development of data-dependent MS/MS coupled to liquid chromatographic

separation increased the number of acquired spectra [36, 37] so now the average LC-MS

experiment acquires hundreds of thousands of MS/MS [20, 38].  Due to the large number

of MS/MS acquired, manual interpretation of spectra is unfeasible.  Therefore, the

development of computational algorithms created an automated approach for MS/MS

data analysis.

The algorithm Sequest [39] was the first written for mining peptide identifications

from data-dependent MS/MS.  This algorithm requires the presence of a database that

contains the predicted protein sequences present in the complex peptide mixture being

9

analyzed.  Subsequent to the release of Sequest, a number of other algorithms have been

released freely or commercially packaged [40-44].  The key for an algorithm to be

deemed successful is based on a number of criteria including (1) accuracy of MS/MS

identification, (2) the length of time for searching a MS/MS, (3) ease of setting the

criteria parameters for an algorithm, and (4) availability of the algorithm for use to the

general scientific community.  Each subsequent algorithm attempts to improve on the

speed and accuracy of matching the MS/MS acquired against a theoretical spectrum

generated from the provided protein database.  Increasing the accuracy of matching the

MS/MS acquired has been a challenge with respect to identifying post-translational

modified proteins [45].  Three search algorithms, DBDigger [44], InsPecT [40], and

Sequest [39] are evaluated in this dissertation and their abilities to search shotgun

proteomics data for post-translational modifications (PTMs) are discussed in detail in

Chapters 2 and 5.

PTMs are chemical modifications found on the amino acid residues of proteins

and can be due to consequences of either biological or inorganic chemical processes [46-

51].  These modifications, when due to an enzymatic process, are sometimes involved in

cellular signaling and transcriptional activation control [52-55].  Oxidation is an example

of an inorganic chemical modification process and can be due to an increase in the

presence of free radicals within the cytosol that are subsequently quenched by proteins

[46, 47].  Determining the identity and relative stoichiometry of the modified version of

the residue in contrast to the unmodified counterpart is an emerging challenge in mass

spectrometry based proteomics [45, 56].  This area is an emerging challenge due to

advancements that have led to the routine cataloguing of the unmodified protein

10

complement in both bacterial cultures and tissue samples. However, the low

stoichiometry of the modified counterpart has led to a large effort in enriching for the

PTM of interest [55, 57, 58]. There are efforts being made to determine from a global

proteome dataset the identity of post-translationally modified proteins [59, 60].

Bacteria are amenable subjects for large-scale studies, such as the shotgun

proteomics studies described in this dissertation for a number of reasons. First, many

bacterial lab strains are easily cultivatable under laboratory conditions, where cellular

material is not a limiting factor. Second, the environmental importance of a growing list

of bacterial species have been implicated in biogeochemical cycling [61-65] and fouling

of energy pipelines [66]. Finally, a large number of bacteria have been characterized

both physiologically and biochemically creating a vast literature database to search. In

addition, the availability of an immense number of bacterial genome sequences from both

isolates and environmental communities [67-78] is required for properly searching the

MS/MS data.

### *Shewanella oneidensis* **MR-1**

*S. oneidensis* MR-1 is a gram-negative γ-proteobacteria first isolated from Lake

Oneida, NY by Myers and Nealson [61] as *Alteromonas putrefaciens*. *A. putrefaciens*

was renamed *Shewanella putrefaciens* by Myers and Nealson [79]. Then, in 1999 *S.*

*putrefaciens* was renamed *S. oneidensis* MR-1 by Venkataswaran *et al* [80]. *S.*

*oneidensis* was enriched from the lake sediments as a manganese and iron reducing

bacterial species. The genome was published by Heidelberg *et al* [69], which facilitated

work in better understanding the metabolic capabilities of this bacterium. *S. oneidensis* is

facultatively anaerobic indicating that the bacterium prefers molecular oxygen as the

terminal electron acceptor for respiration; however, a number of electron acceptors can be utilized under anaerobic growth conditions [61]. In addition to manganese and iron, *S. oneidensis* was found to utilize many other compounds as terminal electron acceptors for respiration under anaerobic conditions [81-86]. Some other electron acceptors found to be utilized by *S. oneidensis* include fumarate and nitrate [81], elemental sulfur [85], and nitrite [87]. In fact, as a result of other unpublished work, the estimation is that of over a dozen electron acceptors may be utilized by *S. oneidensis* [88, 89]. The energetics of various electron acceptors have been determined with oxygen being most favorable followed by trivalent iron, nitrate, tetravalent manganese, and nitrite [90].

Since *S. oneidensis* has such a vast repertoire of electron acceptors for utilization, there must be an extensive regulatory and sensory system in place for recognizing this array of respiratory molecules. Upon completion of the genome sequence annotation [69], the identification of a large number of genes involved in environmental responses were identified. This included the annotation of 88 two-component response regulatory genes [69], three separate pathways for chemotaxis signaling, and 29 methyl-accepting chemotaxis protein receptors (chemoreceptors) [91]. The large number of response regulatory proteins is expected to be due to the diverse environments in which *S. oneidensis* may be found including marine, freshwater, and soil sediments [69]. Nealson *et al* [88] performed the first comprehensive study on chemotaxis response in *S. oneidensis*. The authors found *S. oneidensis* responded dramatically to the presence of nitrate and nitrite by accumulating around wells containing the electron acceptors [88]. In addition, the authors found that the bacterium did not demonstrate any type of response to the transition metals Mn(IV) oxide and Fe(III) citrate. However, a

subsequent study by Bencharit and Ward [92] found that *S. oneidensis* indeed demonstrates a chemotactic response using swarm plate assays to a variety of transition metals utilized as terminal electron acceptors including Mn(IV) and Fe(III).

*S. oneidensis* also contains a number of transcriptional regulatory proteins that have been elucidated in a number of studies [93-97]. Understanding the role of transcriptional regulators in response to gene expression is necessary to understand global regulatory and metabolic pathway function. EtrA, electron transport regulator A, was found to be 73.6% identical to the *Escherichia coli* protein Fnr, fumarate/nitrate regulator, and was found to functionally complement the *fnr* mutant in *E. coli* [93]. Another study using a gene replacement strategy [94] knocked out *etrA* in *S. oneidensis* and found that growth with fumarate and nitrate as terminal electron acceptors was reduced in the mutant cultures. The corresponding terminal reductase activities for the electron acceptors was also reduced in the mutant strain, indicating that EtrA may control transcription of the corresponding reductase genes [94]. A comparative transcriptomic study found that the mutant strain of the *etrA* gene knockout affected the transcriptional levels of 69 genes in *S. oneidensis* [95]. In addition, putative regulatory sequences demonstrating conservation with *fnr* regulatory sequences were identified upstream of 26 operons with affected transcript levels. The corresponding mRNAs from fumarate reductase were not only found to be repressed in the mutant strain, but contained a putative *fnr* sequence as well [94]. Another global regulator, Fur (ferric uptake regulator), has been studied in detail by transcriptomics and proteomics [96, 97]. These studies identified that Fur was involved in the negative regulation of iron transport genes and a putative Fur box was identified upstream of a number of affected genes [96].

Therefore, the presence of perhaps other transcriptional regulators, which respond to transition metals, is likely to be encoded in the genome as well.

As a result of the apparent respiratory versatility of *S. oneidensis,* initial work focused on understanding the enzymatic mechanism involved in using these transition metals as electron acceptors [66, 79, 81-84, 90, 98-103]. This work included attempts to isolate and characterize the terminal electron acceptor reductases [86, 102, 104-110]. First, prior to isolation of *S. oneidensis* MR-1, the utilization of metals for growth was not known to occur. Myers and Nealson [79] determined that a proton motive force was being generated, which is direct evidence for respiratory growth in response to reduction of Mn(IV) and Fe(III) when added as the sole electron acceptors to the growth media. Now, that growth coupled to metal reduction was known, the next step was to determine the enzymes involved in the direct reduction of the metals. A number of studies followed that demonstrated the role of a number of cytochromes and reductases from *S. oneidensis* that were essential to the electron transport chain activated in response to Mn(IV) and Fe(III) [81, 100, 102, 104-119]. Primarily, the identification of outer membrane cytochromes that were required for metal reduction indicated that *S. oneidensis* may require direct contact with the surface for reduction to occur [104-107]. After the genome sequence had been published [69], Meyer *et al* [117] and Yang *et al* [119] found that the *S. oneidensis* genome encodes a large number of cytochromes. In addition, terminal reductase activity for a number of electron acceptors was also found to be located in membrane protein preparations [86, 102, 120]. Direct contact with the surface may explain the lack of metal ion transporters in the genome compared to other bacterial species [69].

## S. oneidensis Cr(VI) Toxicity and Transformation

Another transition metal utilized as a reductant by *S. oneidensis* is the hexavalent species of chromium [Cr(VI)] in the form of chromate ($CrO_4^{2-}$) or dichromate ($Cr_2O_7^{2-}$) [86]. Chromate is a serious pollutant caused by human activities and discharged as liquid waste at many industrial and governmental facilities [121-123]. Traditionally, the process for remediating these contaminated waste sites has involved costly chemical methods [124, 125]. Therefore, the idea of using a natural environmental bacterium that demonstrates an enzymatic ability to reduce Cr(VI) to insoluble Cr(III) hydroxides in addition to a tolerance for various concentrations of Cr(VI) is more economical and environmentally friendly [126]. The use of bacteria for remediation has been proposed and implemented on a number of pollutants in the United States [126-131].

*S. oneidensis* has demonstrated Cr reductase activity in the cytoplasmic membrane [86], however, there are not any known reports on the growth of *S. oneidensis* utilizing Cr(VI) as the sole electron acceptor [132]. Myers *et al* [86] demonstrated the Cr reductase activity in anaerobically grown cultures utilizing fumarate as the electron acceptor with Cr(VI) added to the purified membrane fractions subsequently. The hypothesized pathway for reduction of Cr(VI) proceeds first via a one electron transfer yielding Cr(V) [86]. This is based on evidence from Myers *et al* [86], which found the reductase activity was inhibited by substances known to inhibit members of multi-component electron transport chains. In addition, Viamajala *et al* [133] demonstrated that the kinetic activity of the Cr reductase in *S. oneidensis* must involve at least three separate enzymes. The three enzymes are unknown, but one is described as being relatively slow and there are two that are enzymatically fast but susceptible to inhibition

by nitrite when nitrate is used as the electron acceptor [133]. This indicates that Cr reductase activity in *S. oneidensis* may be non-specific. These kinetics measurements were taken on *S. oneidensis* grown anaerobically with nitrate or fumarate as the electron acceptor. Direct evidence for Cr(VI) reduction in *S. oneidensis* was accomplished by Daulton *et al* [134] identifying extracellular precipitates of Cr(III) surrounding the outside of the bacterial cells. In addition, Cr(III) precipitates in the cytoplasm of *S. oneidensis* have been imaged by transmission electron microscopy [135]. However, there is also evidence that Cr(VI) reduction may occur indirectly via reduction of Fe(III) to Fe(II) first [131, 136, 137].

Cr reductase activity is not novel to *S. oneidensis*, rather a number of other bacteria have shown this activity previously. This includes *S. alga* BrY, a relative of *S. oneidensis*, where *S. alga* was found to be proficient in Cr(VI) reduction following a starvation period [137]. The starvation period causes the cells to decrease in volume allowing them to penetrate further into subsurface environments where Cr(VI) concentrations may be greater. In addition to *S. alga,* other bacterial species encode soluble Cr(VI) reductase proteins that have been purified, including *Pseudomonas putida* MK1 [138], *P. putida* PRS2000 [139], and *Pseudomonas ambigua* G-1 [140].

Even though many bacterial species have demonstrated either direct enzymatic or indirect chemical reduction of Cr(VI), the bacterial species must demonstrate a minimum level of resistance to Cr(VI) toxicity in order to be used as an agent of bioremediation. The toxic effects of Cr(VI) include inhibition of sulfate uptake [141] and the generation of reactive oxygen species (ROS). The ROS cause oxidative damage to DNA [126] leading to mutations that can manifest as cancer in humans [142, 143]. Therefore, a great

16

deal of work has been dedicated to understanding the mechanisms of resistance in addition to the level of resistance of various bacterial species [144-146].

The minimum inhibitory concentration (MIC) for aerobically grown *S. oneidensis* was found to be 2 mM [147]. For anaerobically grown *S. oneidensis*, inhibition was found to be at a much lower concentration [135]. An extensive physiological study on the toxicity effects of Cr(VI) and Cr(III) was performed by Viamajala *et al* [146]. This study found that cultures grown under aerobic or anaerobic conditions would cease to divide at a measurable rate until all Cr(VI) had been reduced; indicating an inhibition of some metabolic process caused by the presence of Cr(VI). Interestingly, the authors hypothesize that the primary toxicity of Cr may be due to the trivalent species remaining bound to the reductases and causing precipitates to form in the cytoplasm [146].

*S. oneidensis* is not the only bacterium that demonstrates resistance to Cr toxicity. Members of the genus *Pseudomonas* have demonstrated resistance to Cr(VI) toxicity via chromosomal and plasmid-borne genes [148-153]. In addition, members of the genus *Alcalignes* were also found to demonstrate Cr(VI) resistance via plasmid-borne genes [154, 155]. Specifically, the ChrA gene was isolated from *Pseudomonas* and *Alcalignes* as responsible for chromate resistance [151, 155] and may encode an active efflux pump that expels Cr(VI) from the bacterium. There is a protein, SO0986, encoded in the genome of *S. oneidensis* that demonstrates significant sequence similarity to the ChrA gene in *Peudomonas aeruginosa*, however, there is no expression evidence to indicate that active efflux of Cr(VI) is used as the mechanism in *S. oneidensis*. In fact, as described above, the molecular mechanisms involved in chromate toxicity and reduction in *S. oneidensis* remain largely unknown.

Due to the fact that very little is known at the molecular level about chromate exposure in *S. oneidensis*, the first goal of the dissertation involved the completion of a number of proteomic studies to improve understanding regarding chromate exposure in *S. oneidensis* (Figure 1.3). This was accomplished by first a global study of acute chromate shock found in Chapter 3 in order to understand the initial response of *S. oneidensis* to sub-lethal levels of chromate. Also in Chapter 3, a global chronic exposure study is outlined demonstrating how the bacterium might respond to sub-lethal concentrations of chromate over an extended period of time. This work was followed by a dosage response study in Chapter 4, the purpose of which was to understand how the bacterium responds to various levels of sub-lethal concentrations of chromate. Finally in Chapter 6, the time during transformation of chromate from the hexavalent species to the trivalent species is explored; giving possible evidence of proteins that may be transporting and reducing the sub-lethal concentration of chromate in wildtype and Δ2426 mutant cultures. The Δ2426 mutant was created based on evidence from previous studies [147, 156] that the protein, SO2426, was highly up-regulated in response to acute chromate exposure. In addition, Cr(VI) reduction assays indicated that the Δ2426 mutant was deficient in Cr(VI) transformation in addition to a number of other transition metals [157].

The second goal of the work was more technology driven and focused on different method development areas to build upon the proteomics pipeline developed at Oak Ridge National Laboratory. Improvements to two of the technologically challenged areas of proteomics (sample preparation for proteome analysis and computational analysis of MS/MS data) enhanced both the experimental design and the resulting information obtained from performing shotgun proteomics experiments. Figure 1.4

**Global Molecular Response of *Shewanella oneidensis* MR-1 to Chromate Exposure**

**Acute exposure**
1mM 45 and 90 min
Initial response to Cr(VI)
•Transporters(TonB receptors, TonB, siderophore, etc.
•Response regulator(2426)
•Reductases(3585)

**Dosage response**
0, 0.3, 0.5, 1 mM 30 min
Environment = gradient
• 0.3mM= 4718(sigma54-response regulator) up
• 0.5 and 1mM= sulfate transport system up
• 1mM= 7 stress proteins (PepB, SO1117, IbpA, ClpA, RecN, ClpB, SOA0048)

**Cr(VI) Reduction**
0.3mM over 4hr
Observe Cr(VI) disappear
1. Proteins involved
   • Transporters
   • Reductases
2. Stress proteins
   • Chemotaxis
3. Role of 2426

**Chronic exposure**
0.3 mM 24 hr
all Cr(VI) disappeared
• lytic phage activated (14 proteins)
• flagellar proteins down
• cell division machinery failing (cells become elongated and fibrous)

**Post-translational modifications**
• important protein activity regulators
• in shotgun proteomic data
   • no additional instrument time needed
• chromate exposure = stress
   • should readily observe chemotactic response
   • affect "peptide count" if most peptides are modified

**Biomass Lysis Tests**
• use biological replicates instead
• better mimic natural growth conditions and reduction chemistry
   • Cr(VI) reduction rate increases with bacterial population size
• no established method to deal with these demands

**Figure 1.3**. Flowchart diagram of the proteome studies of *S. oneidensis* Cr(VI) exposure.
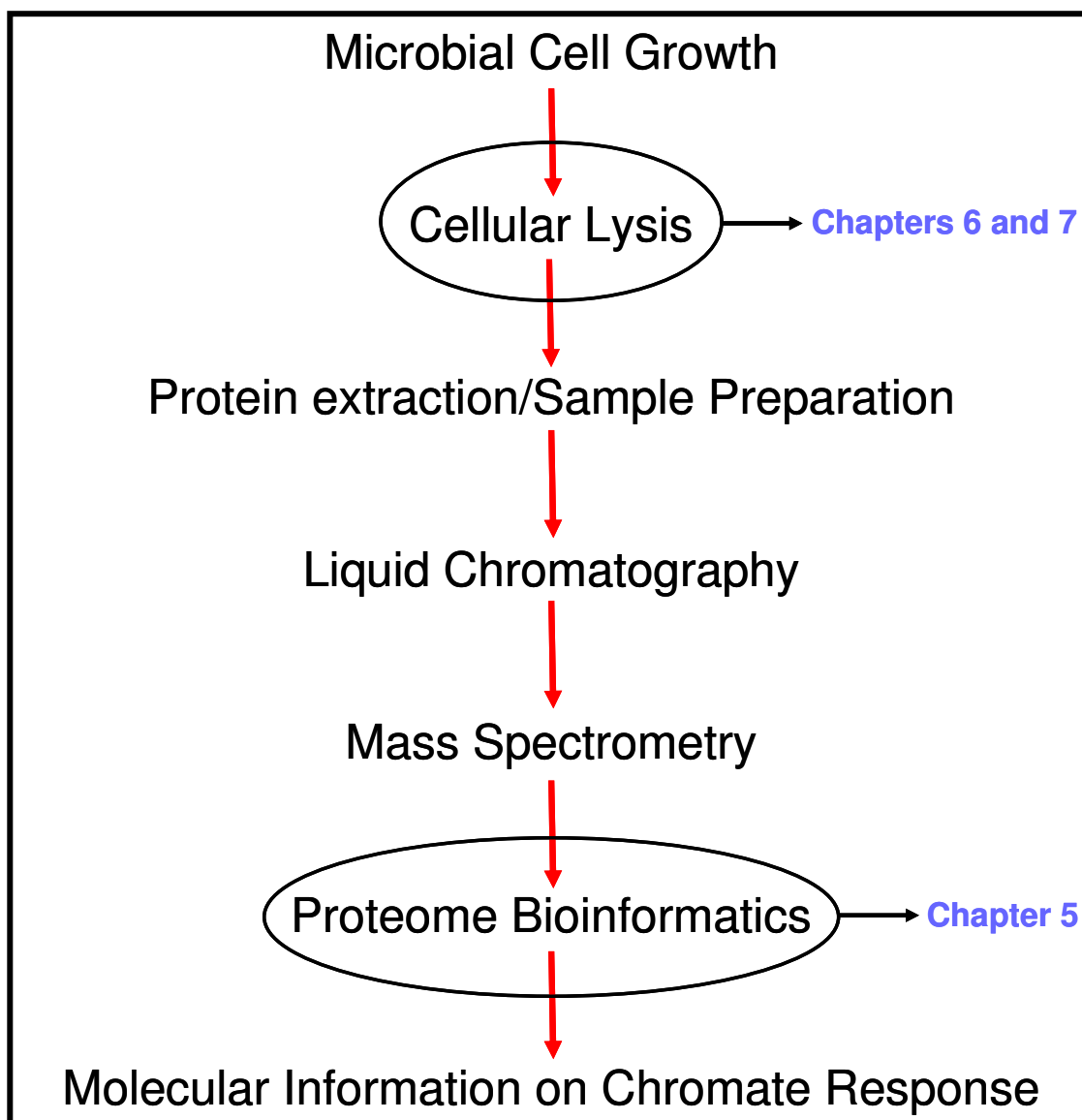
**Figure 1.4.** Representative shotgun proteomics experimental design.

Outlined in the flow chart are the steps in a representative experiment from whole-cell lysate to the resulting data analysis of the acquired MS/MS data. Improvements to the setup are highlighted and details are outlined in Chapters 5, 6, and 7.

depicts the steps involved in a conventional shotgun proteomics experiment and is representative of the pipeline used at Oak Ridge National Laboratory from cell lysis to data analysis. Highlighted are the method development areas described in this dissertation that have been accomplished to improve on the design of the experimental setup. Chapter 5 demonstrates a novel systematic study of mining shotgun proteomics data for PTMs. This was accomplished by comparing the performance of three algorithms (DBDigger [44], InsPecT [40], and Sequest [39]) by searching MS/MS data from whole-cell lysates of *S. oneidensis* for PTMs. Chapter 6 is not only about understanding chromate transformation in *S. oneidensis*, but comprises one of the first comprehensive proteome measurements from biological replicate cultures of a bacterium. In addition, Chapter 7 describes a novel single-tube lysis method that has been successfully applied in the results of Chapter 6. The single-tube lysis method permits global proteome detection from a few milliliters of bacterial cell culture versus the traditional liters of cell culture. This is a greater than 1000x reduction in the amount of cellular material needed for proteome detection and characterization. All of the research described in this dissertation comprises both novel and thorough evaluations performed to understand acute and chronic chromate exposure as well as during chemical transformation of chromate at the molecular level utilizing methodologies in the systems biology area of shotgun proteomics.

## Chapter 2

## Experimental Design of Shotgun Proteomics Experiments and Bioinformatic

## Platforms using Multidimensional Protein Identification Technology

### Introduction

The work presented in this dissertation utilized a mass spectrometry based platform for global protein characterization from microbial isolate cultures and a natural microbial community. This chapter describes the experimental details and components of this platform utilized for large scale proteome measurements. Each chapter summarizes the specific mass spectrometry method used (i.e. LCQ or LTQ), with this chapter describing the overall methodological approach explored and developed. The instruments utilized during the course of the dissertation were a LCQ (liquid chromatography quadrupole ion trap, three-dimensional ion trap) and a LTQ (linear trapping quadrupole) coupled to an online two-dimensional chromatography method. In addition to acquisition of mass spectrometry data, the lack of a feasible means of searching the data for post-translational modifications was addressed through optimizing the scoring filters from three bioinformatics platforms and adapting an in-house script.

### Microbial Cultures and Communities

Microbial cultures (Chapters 3-7) were obtained from isolates stored as glycerol stocks at -80 °C and were provided by Dr. Dorothea Thompson from Purdue University (*Shewanella oneidensis* MR-1) and Dr. Dale Pelletier from the Biosciences Division at Oak Ridge National Laboratory (*Rhodopseudomonas palustris* CGA0010). The acid mine drainage biofilm microbial community used in Chapter 7 was a gift from Dr. Jillian

Banfield at University of Berkley. *S. oneidensis* was cultivated in batch culture under aerobic conditions and *R. palustris* was cultivated under photoheterotrophic conditions. *S. oneidensis* was grown in 500 mL cultures for [147, 156, 158] or 100 mL cultures for [159] under aerobic conditions to mid-exponential phase ($A_{600}$, 0.5) followed by continued monitored growth or the addition of $K_2CrO_4$ for a given period of time. For cellular harvest, cells were pelleted by centrifugation (5,000 x *g* for 5 min), resuspended in ice-cold LB medium, washed two times in 50 mM Tris, 10 mM EDTA (pH 7.6), and centrifuged at 5,000 x *g* for 10 min. The cell pellets were stored at -80 °C until cellular lysis and digestion. Details of microbial growth can also be found in [147, 156, 158-160].

**Preparation of Proteomes for HPLC-MS/MS Analysis**

All chemical reagents were obtained from Sigma unless stated otherwise. Modified sequencing grade trypsin (Promega, Madison, WI) was used in all digestions. The modified trypsin used for digestion was methylated on lysines and arginines, thereby reducing the autolytic behavior of the enzyme, which may interfere with detection of the peptides of interest [161]. HPLC grade water and acetonitrile were acquired from Burdick & Jackson (Muskegon, MI), and 99% formic acid was purchased from EM Science (Darmstadt, Germany).

*Lysis by Sonication and Tryptic Digestion*

For proteome analyses in Chapters 3, 4, 5, and 7, the *S. oneidensis* cells were placed on ice and lysed by sonication using a microprobe at high power with 30-s pulses five times with a 30-s cooling period between each sonication. *R. palustris* and the acid mine drainage biofilm in chapter 7 were also lysed by sonication, however due to the

invaginated membrane of *R. palustris* and the cellulose matrix structure of the biofilm, there were ten 30-s sonication pulses followed by the 30-s cooling period. Cellular debris was removed by centrifugation at 5,000 x *g* for 10 min. The supernatant was centrifuged at 100,000 x *g* for 60 min in an ultracentrifuge to separate a soluble fraction from a pellet for Chapters 3, 4, and 5. The pellet (membrane fraction) was washed with 50 mM Tris, 10 mM EDTA (pH 7.6) and centrifuged at 100,000 x *g* for 60 min; this fraction was then resuspended in 50 mM Tris, 10 mM EDTA (pH 7.6) by brief sonication. Both proteome fractions were quantified using bicinchoninic acid (BCA) analysis [162], aliquoted, and stored at -80 °C until ready for digestion. Approximately 2 mg of each proteome fraction (soluble and membrane) was denatured and disulfide bonds reduced in 6 M guanidine and 10 mM DTT (60 °C for 1 h). The denatured/reduced proteome mixture was diluted 6-fold with 50 mM Tris, 10 mM $CaCl_2$ (pH 7.8), and sequencing grade trypsin was added at 1:100 [protease/protein (w/w)]. The digestions were run with gentle shaking at 37 °C for 18 h, followed by a second addition of trypsin at 1:100 and an additional 5 h incubation. The samples were treated with 20 mM DTT for 1 h at 37 °C as a final reduction step to remove remaining disulfide bonds and then immediately desalted using Sep-Pak Plus C18 solid phase extraction (Waters, Milford, MA). A second reduction step with DTT was performed instead of using the cysteine alkylation reagent iodoacetemide. Iodoacetemide blocks the cysteine residues through the addition of a carboxymethyl group (57 Da), which will appear as a static modification in the resulting MS/MS searches. The added complexity of this modification in the resulting data filter levels must be taken into consideration as described below for PTMs. All samples were concentrated and solvent-exchanged into 0.1% formic acid in water by centrifugal

evaporation to ~10 µg/µl starting material, filtered, aliquoted, and stored at -80 °C until ready for LC-MS/MS analysis.

*Single-Tube Lysis and Tryptic Digestion*

For the proteome analyses presented in Chapters 6 and 7, a single-tube lysis method was employed. This method differs from the above sonication method and was performed as follows. Microbial cell pellets ranging in size from 1 mg to ~200 mg were lysed using 6 M Guanidine/10 mM DTT dissolved in 50 mM Tris/ 10 mM $CaCl_2$ pH 7.6 (Tris Buffer). The $Ca^{2+}$ ion is necessary to enhance activity by promoting autocatalytic activity converting trypsinogen into trypsin, the active enzyme [163]. In addition, the resulting protein content is denatured with the guanidine and the mixture was incubated overnight at 37 °C. Following lysis, the samples were diluted 6-fold with the Tris Buffer and trypsin was added in an optimized amount for 6 hr at 37 °C with gentle rocking followed by a second trypsin aliquot overnight. The optimized amount of trypsin was determined in order to obtain the fewest autolytic tryptic peptide identifications after the LC-MS/MS experiment. When too much trypsin is added, the resulting proteome dataset will have greater than 50% sequence coverage of trypsin identified in the resulting dataset, in contrast to the ~20% sequence coverage observed if an optimized amount is used. A final 20 mM DTT reduction step was performed following the proteolytic digestion for 2 hr at 37 °C with gentle rocking. After lysis, proteolytic digestion, and the final reduction step a high speed centrifugal step was performed to pellet cellular debris. Samples were then desalted with a Sep-Pak $C_{18}$ Lite or Plus cartridge and solvent exchanged into 100% $H_2O$, 0.1% formic acid; followed by filtration using an Ultrafree-MC centrifugal filter device (Millipore, Billerica, MA) and stored at -80 °C until LC/LC-

MS/MS analysis. The final filtration step is necessary to remove any aggregate cellular material that will clog the subsequent LC column.

Lysis using Tetrafluoroethylene (TFE) was performed in the same manner as the Guanidine HCl method, except TFE was added initially at a concentration of 50:50 TFE:Tris Buffer/10 mM DTT. For the microbial community (acid mine drainage biofilm) discussed in Chapter 7, a different single-tube lysis method, freeze/grinding, was attempted. The freeze/grinding method was accomplished by first flash-freezing the biofilm in liquid nitrogen, followed by mechanical grinding of the sample into a fine powder. Then, the cell pellets were resuspended in 6 M Guanidine/10 mM DTT in Tris Buffer with the remaining steps being as above for the Guanidine lysis method.

*Small Sample Lysis by Bead-Beating*

The bead-beating method of lysis [164] is not a single-tube lysis method, however bead-beating is similar in the respect that a smaller than traditional amount of biomass can be lysed with this method. The bead-beating method used 0.5 mm glass beads (Biospec Products, Bartlesville, OK) first sterilized in HPLC-grade methanol (Burdick and Jackson) overnight. Next, an approximately similar amount of glass beads as the cell pellet for lysis is transferred to a separate 2 mL eppendorf tube with a spatula cleaned with methanol. About 250 µL of 6 M Guanidine/10 mM DTT is added to the glass beads and vortexed. The biofilm (as in Chapter 7) is transferred to the glass beads and vortexed 5 times for 30 s with a 30 s cooling period. The biofilm/glass bead slurry is then incubated over night at 37 °C. Following this step, the Guanidine is diluted to 1 M with 50 mM Tris/10 mM $CaCl_2$ and centrifuged to pellet the glass beads. The supernatant is transferred to a separate eppendorf tube and 10-20 µg of trypsin was added for 5 hr at 37

°C with gentle rocking, and another aliquot of trypsin was added with an overnight incubation. Finally, a final reduction step for 2 hr with 20 mM DTT, followed by centrifugation to pellet debris, and the samples were then desalted with a Sep-Pak C18 Plus cartridge. As described in Chapter 7, this method has not been optimized and the results for this method are worse than the other lysis methods due to the buoyancy of the biofilm floating on the top of the glass beads.

## Liquid Chromatographic Separation of Peptides

Separation of complex mixtures of peptides using an online two-dimensional liquid chromatographic separation has many advantages over other methods of separation. In particular, the use of orthogonal separations such as strong cation exchange (SCX) and reverse phase (RP) provides coupled but independent separation dimensions. An orthogonal separation method is defined as use of two or more chromatography types, which separate peptides based on two or more different chemical properties (i.e. hydrophobicity, size, ionic properties, etc.). There are other methods of two-dimensional separation available (see Chapter 1 for discussion); however the coupling of SCX and RP online with mass spectrometry analysis of peptides has become the most successful implementation of two-dimensional orthogonal online separation of peptides. SCX separates peptides based on their affinity for the negatively charged benzene sulfonic acid bonded resin and RP separates peptides based on their affinity for a highly hydrophobic 18-carbon chain bound resin. The traditional method of SCX separation is not considered to be compatible with mass spectrometry due to the use of a strong salt (i.e. NaCl) for peptide elution. However, the work demonstrated here utilizes the volatile salt ammonium acetate, which exhibits greater compatibility with MS

allowing for online separation.  In addition, the total capacity of the column must be considered prior to loading the peptides.  Overloading leads to diminished separation of the peptides, as well as a decreased confidence in the prediction of the elution time.

The proteome fractions (soluble and membrane) prepared from control and chromate-treated samples were analyzed in duplicate for Chapters 3, 4, and 5 and from whole cell lysates in triplicate for samples in Chapters 6 and 7 via a 24-hr two-dimensional (2-D) LC-MS/MS experiment using an Ultimate HPLC system (LC Packings, a division of Dionex, San Francisco, CA).  The HPLC pump provided a flow rate of ~100 µl/min that was split pre-column using a fused silica setup as shown in Figure 2.1 to achieve a final flow rate of ~300 nl/min at the nanospray tip. The flow at the tip was measured using a 5 µL capacity calibrated micropipet (Drummond Scientific Company, Broomall, PA) with 1 µL increments over at least a 3 min time frame.  This was chosen due to the flow rate variation over time and so an average flow rate is taken over the specified time period.  A split phase column (150 µm inner diameter fused silica) was packed via a pressure injection platform (New Objective, Woburn, MA) as follows:  first with ~3.5 cm of strong cation exchange (Luna SCX, 5 µm particle size, 100 Å distance between particles when packed together; Phenomenex, Torrance, CA) followed by ~3.5 cm of $C_{18}$ reverse phase (Aqua C18, 5 µm, 200 Å; Phenomenex). Subsequently, ~500 µg of proteolytic peptide sample was loaded onto the split phase column via the pressure injection platform.  The sample size of 500 µg was chosen in order to take advantage of the dynamic range capabilities of the ion trap mass spectrometers used.  This amount of protein slightly overloads the theoretical capacity of the column, but allows us to identify proteins that may be present at a concentration of
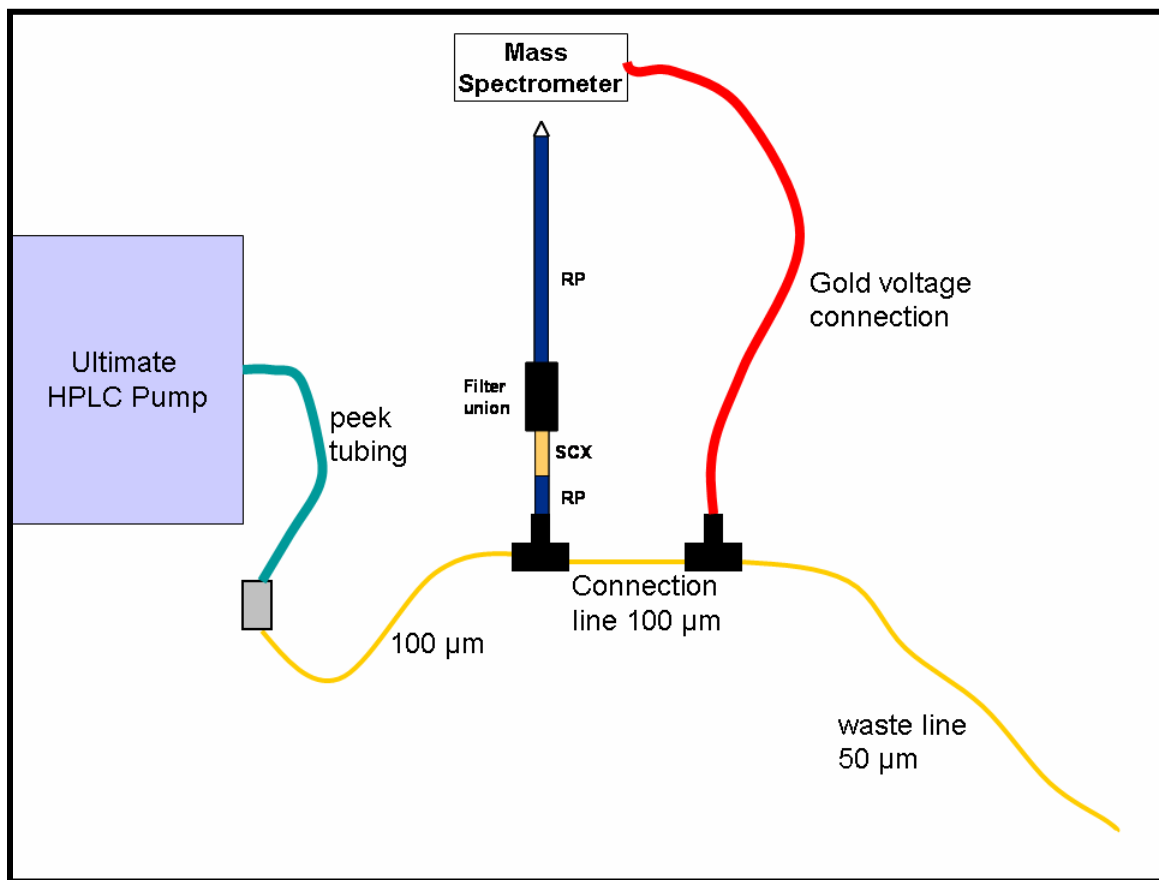
**Figure 2.1.** Schematic diagram of the two-dimensional column setup.

The Ultimate pump is connected to the two dimensional nanocolumn via peek tubing

connected to 100 µm fused silica by a metal fitting, which acts as a ground for the

voltage. The 50 µm waste line is depicted as well extending from the second micro tee

connecting the mass spectrometer voltage to the nanocolumn.

nanograms per milliliter [165].  The loaded split phase column was then inserted behind a

PicoFrit tip (100 µm inner diameter, 15 µm inner diameter at the tip; New Objective)

packed via the pressure injection platform with ~15 cm of $C_{18}$ reverse phase (Jupiter C18,

5 µm, 300 Å or Aqua C18, 5 µm, 200 Å; Phenomenex).  The $C_{18}$ reverse phase resin on

the split phase column acts as a trapping cartridge for the peptides to initially bind to,

with the first chromatographic step being a desalting step as described by McDonald *et al*

[166].  This acts to remove impurities from the sample preparation process and move the

peptides during the gradient [100% Buffer A (95% $H_2O$, 5%ACN, 0.1% formic acid) to

100% Buffer B (30% $H_2O$, 70% ACN, 0.1% formic acid)] from the $C_{18}$ resin to the SCX

resin for subsequent separation based on charge affinity.

Following the first chromatographic step (the desalting step), proteome fractions

(Chapters 3, 4 and 5) or cell lysates (Chapters 6 and 7) were separated using the

orthogonal methods of SCX followed by RP for 11 subsequent salt steps (a step

gradient).  Table 2.1 depicts the time and amount of ammonium acetate added to dislodge

peptides from the SCX resin for each salt step and subsequent organic phase separation.

The chromatographic separation of each salt step is as follows:  2 min of Buffer A for re-

equilibration of the column followed by 2 min (steps 2-11) of the specified amount of salt

or 10 min for step 12.  After the salt step, there was another Buffer A equilibration period

of 5-10 min followed by the Buffer A to Buffer B RP separation.  The re-equilibration

period with the second Buffer A step is critical; if this step is not performed the salt

introduced will precipitate out of solution in the presence of organic solvent.  This RP

separation method consisted of the gradient from 100% Buffer A to 50% Buffer B for

steps 2-11 and 100% Buffer A to 100% Buffer B for step 12.  The goal of step 12 is to

**Table 2.1.**  The Consecutive LC Steps Employed in the Two-Dimensional Separation

| Chromatography Step | Time (min) | Ammonium Acetate Added (mM) |
| --- | --- | --- |
| 1 | 60 | 0 |
| 2 | 120 | 50 |
| 3 | 120 | 75 |
| 4 | 120 | 100 |
| 5 | 120 | 125 |
| 6 | 120 | 150 |
| 7 | 120 | 175 |
| 8 | 120 | 200 |
| 9 | 120 | 225 |
| 10 | 120 | 250 |
| 11 | 120 | 300 |
| 12 | 100 | 500 |

remove all peptides that may still be binding to the resin following step 11, which explains the longer salt step followed by the complete organic phase endpoint during the RP separation. The complete organic phase acts to diminish carryover between LC-MS/MS experiments due to peptides remaining bound to the 15 cm RP analytical column, which is used for multiple chromatographic separations.

Typical total ion current (TIC) and base peak chromatograms are depicted in Figure 2.2 for chromatography steps 1, 6, and 12. These steps were chosen as representatives of how the chromatograms should look for a LC separation. Note how the relative intensity disappears during the injection of salt across the chromatography column. The loss of intensity is a result of the greater conductivity of the salt raising the current of the nanospray tip causing cessation of electrospray. Also, there is a notable difference in peptide elution between steps 6 and 12. In step 12, elution of a majority of peptides occurs by 100 min into the chromatography step indicating complete elution of peptides from the analytical column, which leads to less carry over between sample analyses. Originally, step 12 was 120 min in length; however as noted above, the peptides primarily elute by 100 min and so the gradient was shortened accordingly (See Table 2.1).

**Mass Spectrometry Experiments for the Detection of Peptides**

During the entire chromatographic process, the three-dimensional ion trap (LCQ) or linear trapping quadrupole (LTQ) was operated in a data-dependent MS/MS mode detailed below. The chromatographic methods and HPLC columns were similar for all analyses. The LC-MS/MS system was fully automated and under direct control of the Xcalibur software system (Thermo Scientific, San Jose, CA). Operation of the mass
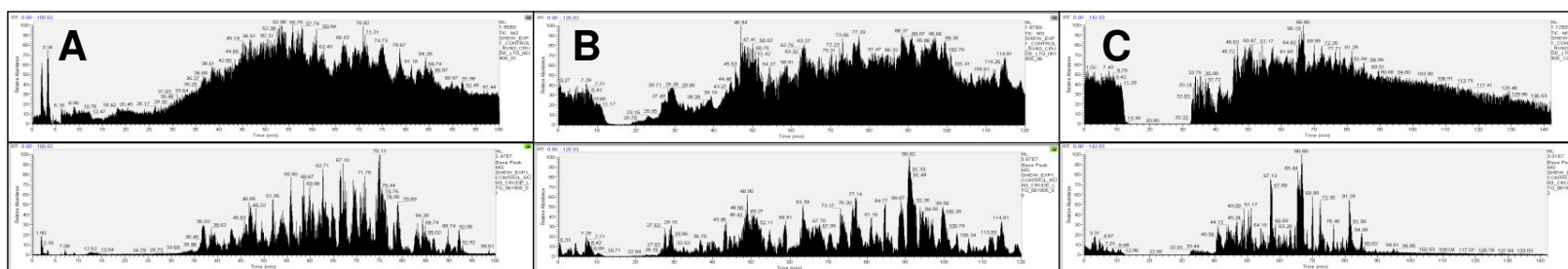
**Figure 2.2.** TIC and base peak chromatograms from selected chromatography steps.

The TIC chromatograms are on the top with the base peak chromatogram below for (A) chromatography step 1, (B) step 6, and

(C) step 12 from Table 2.1. Even though there appears in the TIC chromatogram to be an overwhelming amount of peptides

eluting at any given time, the base peak chromatogram illustrates that there is actually respectable separation of the peptides.

spectrometer in a data-dependent mode allows for increased dynamic range through the use of dynamic exclusion. Dynamic exclusion is used to reduce the sampling of abundant ions being triggered for MS/MS fragmentation; when an abundant ion is detected, it is subsequently placed on a list that prohibits that ion from being isolated for a set period of time (i.e. 3 min). The data-dependent mode of operation is independent of operator control, where the control software (Xcalibur) is programmed such that during a set period of time, a full MS scan is acquired followed by a specified number of MS/MS scans. This loop is repeated until the set period of time is completed (i.e. the liquid chromatographic separation). The MS/MS scans are dependent on ion intensity in the respect that following the MS scan, the *n*th most intense ion is isolated and fragmented where *n* is equal to 1, 2, 3, etc. Once the ion is chosen for fragmentation, the *m/z* value is placed on the dynamic exclusion list and is not isolated or fragmented until removed from the list. Discussed below are the details of how each mass spectrometer was operated and the process by which ions are detected and isolated by each instrument.

*Operation of a three-dimensional ion trap (LCQ)*

The LCQ is a three-dimensional ion trap composed of a ring electrode in the middle and two end-cap electrodes acting to trap ions in the center of the ring electrode. This instrument was first developed in 1960 by Paul and Steinwedel and known as the Paul trap [167]. Advancements to the design leading to successful commercialization of the instrument came in 1984 by Stafford *et al* [29]. Ions are injected into the trap via focusing multipoles transferring the ions as a coherent packet from the electrospray ionization source operated at atmospheric pressure through a heated capillary for desolvation of the ions and into the vacuum of the mass spectrometer. The ions become

34

energetically stable in the center of the trap according to the stability diagram represented in Figure 2.3 by applying DC (U) and AC (V) voltages as well as a frequency ($\cos\Omega t$) to the ring electrode, thereby creating a dynamic trapping field over a specific mass to charge (*m/z*) range [29]. The dynamic trapping field is depicted mathematically with the following equations. The DC voltage (usually kept at 0) is reflected by $a = 4zU/mr_0^2\Omega^2$, where a is a unitless value comprising: U, DC voltage; *z*, charge of the analyte; m, mass of the analyte, r, radius of the ion trap; and $\Omega$, rf frequency. The AC voltage is shown mathematically by $q = 2zV/mr_0^2\Omega^2$, where q is a unitless value comprising: V, voltage, and all other components are the same as above. In addition, a dampening He gas is added to the trap to remove energy from the ions through collisions. In order to eject and thereby detect the trapped ions, the RF voltage is linearly increased to destabilize the trajectory of the ions inversely proportional to *m/z* according to $(m/z)_{eject} = 4V/0.908\ r_0^2\Omega^2$. Therefore, ions lighter in mass will be ejected from the trap prior to the heavier ions. Once an ion is ejected from the trap, it impinges on a conversion dynode, which dislodges many electrons that in turn impinge on an electron multiplier giving an intensity signal for a particular *m/z* value. The resulting mass spectrum displayed depicts the *m/z* values ejected from the ion trap and their resulting relative intensities.

This process of trapping ions is known as mass selective instability [29] in contrast to the original design of the Paul trap [168] using mass selective detection. The mass selective instability used in the Stafford ion trap [29] allows for faster scan speeds, which permits this instrument to function on liquid chromatographic separation time scales. In addition, Stafford *et al* [29] found that the addition of a low molecular weight gas (i.e. helium) improves the resolution, sensitivity, and dynamic range of the LCQ. By
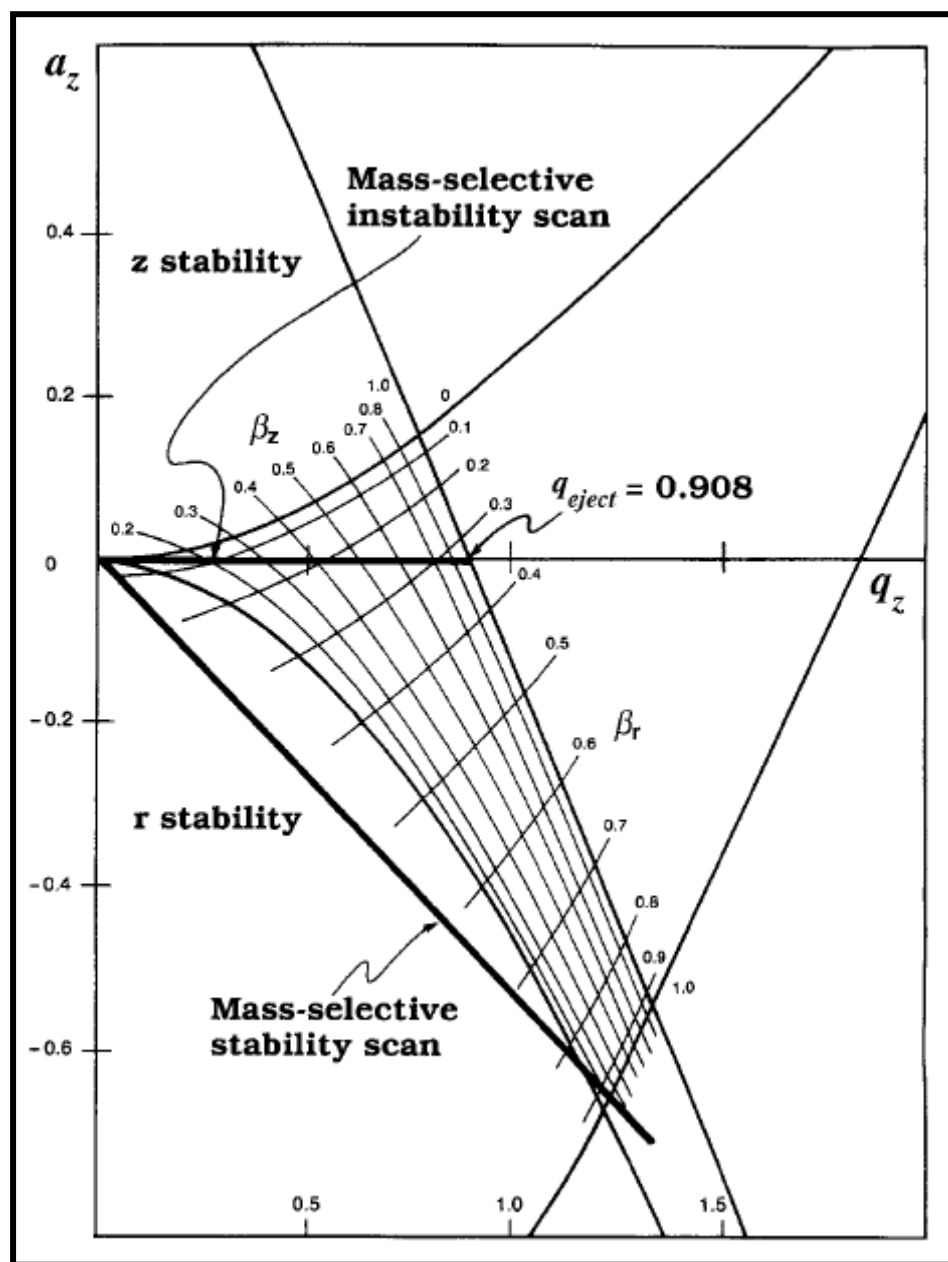
**Figure 2.3.** Instability diagram of the three-dimensional ion trap.

Reprinted Figure 3 from Analytical Biochemistry, Vol. 244, K. R. Jonscher and J. R.

Yates III, The quadrupole ion trap mass spectrometer--a small solution to a big challenge,

pages 1-15, Copyright 1997, with permission from Elsevier.

the addition of He gas, the ions injected into the trap collapse towards the center

minimizing field imperfections caused by structural errors of the ion trap. Tandem mass

spectrometry operates in a similar manner as above, where the AC voltage is set as a

window of *m/z* values that destabilizes the trajectory of all other *m/z* values of either

lower or higher values except the *m/z* of interest. Fragmentation occurs through

increasing the supplemental RF frequency causing the selected ions to be more energetic,

colliding with the He gas molecules within the trap. This creates higher vibrational

energy states for the selected ion, leading to dissociation of the relatively weaker covalent

bonds within that ion, yielding fragments of the original ion detected as *m/z* peaks in the

MS/MS [169].

The following parameters were applied to the LCQ analyses: nanospray voltage

of 2.6 kV, heated capillary temperature of 200 °C, and a full mass scan range of 400-

1700. MS/MS were acquired in a data-dependent mode as follows: 4 MS/MS were

obtained following every full scan; 5 microscans were averaged for every full MS and

MS/MS; a 5 *m/z* isolation width was employed; 35% collision energy was used for

fragmentation, and the dynamic exclusion was set to 1 with the duration being 3 min in

length.

*Operation of a linear trapping quadrupole (LTQ)*

The fundamental operation of the LTQ is very similar to the LCQ, especially

considering both instruments are ion traps. This instrument was developed by Jae

Schwartz at Thermo Finnigan (now Thermo Scientific) in 2002 [30]. The principal

difference between the LCQ (a three-dimensional ion trap) and LTQ (a two-dimensional

ion trap) is improvements that were made to the mechanical design of the trap itself. The

LTQ has a larger trapping volume for injected ions where the ring electrode has been extended in length and broken apart into 4 hyperbolic rods with slits in two for ion ejection and subsequent impingement on the conversion dynode. This is in contrast to the LCQ design where the trapping ring electrode is a single piece. The endcaps of the LCQ have been converted into smaller quadrupoles, in which RF and DC voltages are applied repelling the ions into the center quadrupole. The primary advantage to this instrument is the increased ion storage space provided by the lengthened trap. The design of the LCQ limited the storage capacity of the ion trap, thereby limiting the dynamic range of the instrument, as well as mass accuracy problems arising from space charging [30]. Space charging is defined as the effect of over-filling the ion trap with ions, causing a repulsive internal force on the ion packet. This repulsive force leads to ions on the outer edges of the packet to eject at a different resonance frequency than the *m/z* value predicts [170]. In other words, the *m/z* that is detected has a wider ion peak and the *m/z* is higher than the true *m/z* value for that particular ion leading to lower mass resolution. An example of the increased dynamic range of the LTQ is found in Figure 2.4, which depicts a low-abundant ion that could be misinterpreted as noise giving a *m/z* peak-rich MS/MS.

The LTQ was operated with a nanospray voltage of 2.6 kV, heated capillary temperature of 200 °C, and a full scan *m/z* range of 400–1700. The data-dependent MS/MS mode was operated as follows. Five MS/MS were acquired following every full scan and two microscans were averaged for every full MS and MS/MS. An isolation width of 3 *m/z* was used and 35% relative collision energy was used for fragmentation. The dynamic exclusion was set to 1 with an exclusion duration of 3 min.
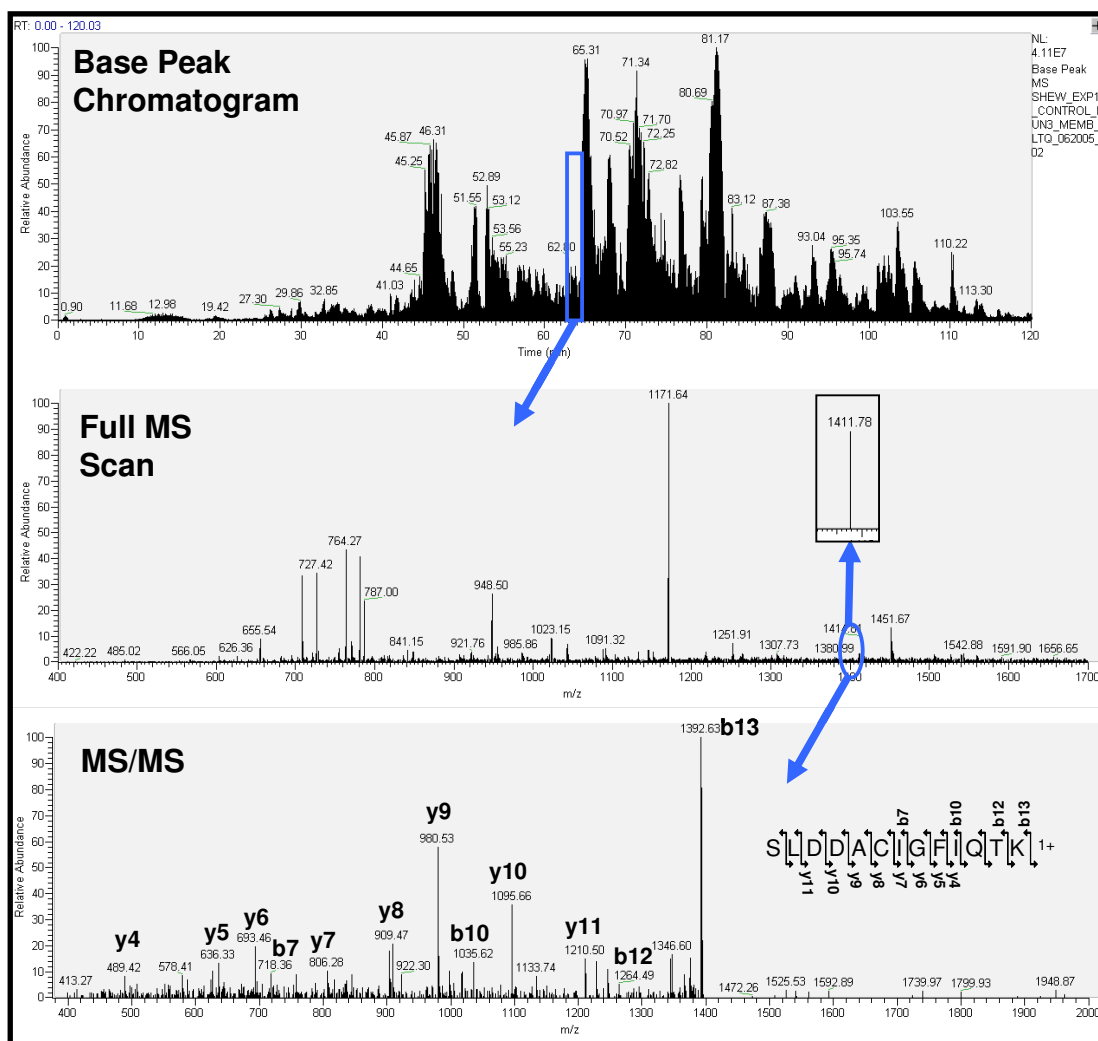
**Figure 2.4.** A MS/MS from a low-abundant peptide using the LTQ.

The peptide, SLDDACIGFIQTK, is from SO3587 a putative transmembrane domain

hypothetical protein in *S. oneidensis*. This peptide was found as singly charged in the

membrane fraction of a LC-MS/MS experiment. The top panel is the base peak chrom-

atogram of chromatography step 2 from the experiment and the middle panel is the full

MS spectrum with a zoomed-in region of the spectrum showing the corresponding intact

peptide with an *m/z* value of 1411.78. The ion was subsequently isolated and fragmented

yielding the MS/MS of the bottom panel with the respective fragment ions labeled.

39

**Proteome Bioinformatics**

Due to the hundreds of thousands of MS/MS acquired during the LC-MS/MS experiments performed, there is a paramount necessity for the development and optimization of bioinformatics platforms to distinguish the identities of the peptide ions fragmented, and filter the resulting datasets acquired by removing false identifications, also known as false positives. False identifications are the result of the misidentification of a MS/MS, which has a scored value that is greater than the threshold filters. A false negative identification is the result of a true peptide identification that has not scored high enough to pass the threshold filters. To this end, a number of search algorithms were assessed and filtering levels for identifications were optimized based on the characteristics of both the datasets (i.e. instrumentation platform) and the algorithm's performance for searching MS/MS data.

*The Sequest algorithm*

Sequest was the first algorithm written to identify peptides fragmented by MS/MS [39]. As the performance and speed of both mass spectrometers and liquid chromatographic separations increased, the bottleneck to be addressed was computational automation of MS/MS identification. To this end, Sequest was written and published in 1994 by Jimmy Eng at the University of Washington-Seattle. This now is part of the Bioworks software licensed by Thermo Scientific. First, Sequest "digests" the provided protein FASTA database *in silico* into peptides based on the enzyme specificity provided by the parameters file (i.e. fully tryptic, no enzyme specificity, etc.). This creates a list of possible peptides that are represented in the experimental MS/MS. Next, Sequest generates theoretical MS/MS using the provided digested peptides from the protein

40

database in order to determine identifications of the experimental spectra. The scoring

process first determines the number of matching peaks between the two spectra and

increases the resulting score if the experimental peak identifications comprise a

consecutive sequence for the peptide. Next, a cross correlation score (Xcorr) is given

based on a comparison of the top 500 scoring peptide identifications [39]. The highest

Xcorr is then compared to the Xcorr of the second best identification and a DeltaCN

value is given. The DeltaCN ($\Delta$CN) value indicates a confidence in the identification of

the MS/MS as belonging to the best scoring peptide from the database.

Sequest is widespread in proteomics; therefore results found in Chapters 3, 4, 5, 6,

and 7 were all published with this algorithm. The database used in the searches

performed in this dissertation can be found in the experimental section of the relevant

chapter. The MS/MS spectra from individual RAW files were first converted to

.mzXML format by using ReAdW software written at the Institute for Systems Biology

in Seattle, WA (www.systemsbiology.org) and can be downloaded from the SourceForge

repository (sashimi.sourceforge.net). Individual spectra were then converted to DTA

files by mzXML2Other, also from the Institute for Systems Biology. DTA files are the

required format for input into Sequest (see Ref. [2]). The parameters for searching the

MS/MS data with Sequest are: enzyme type, trypsin; Parent Mass Tolerance, 3.0;

Fragment Ion Tolerance, 0.5; up to 4 missed cleavages allowed; and fully tryptic peptides

only. The following filter levels were applied to the Sequest search results using the

algorithms DTASelect and Contrast [171]: tryptic peptides only, $\Delta$CN value of at least

0.08, and Xcorr values of at least 1.8 (+1), 2.5 (+2), 3.5 (+3). These values were chosen

and described in Chapters 3 and 4 due to the acceptable false discovery rate (FDR) of

~2% at the 2-peptide level of protein identification that result from this filter level. No chemical modifications were added to the Sequest searches for Chapters 3, 4, 6, and 7. As discussed in detail in Chapter 5, the addition of chemical modifications to a Sequest search is cumbersome.

*The DBDigger algorithm*

The search algorithm DBDigger [44] was written by David Tabb at Oak Ridge National Laboratory to help alleviate some of the inflexibilities present in the Sequest algorithm. In addition, the speed at which a search of MS/MS against a protein database can be performed using DBDigger on a desktop computer is greater than Sequest. DBDigger uses a similar method of scoring experimental MS/MS against a provided protein database to the algorithm Sequest. However, the method of searching by creating a theoretical spectrum once for a given set of experimental MS/MS yields the enhanced speed provided by this algorithm. DBDigger also allows for the additional flexibility of including an unlimited number of chemical modifications to the search parameters. This permits the ability to start searching for post-translational modifications embedded in the MS/MS datasets. However, there are a number of limitations to this algorithm that are discussed in detail in Chapter 5.

DBDigger was used as a search algorithm due to the advantages listed above for datasets presented in Chapters 5 and 7. The resulting raw MS/MS files were first converted to .ms2 files using the algorithm RAW2MS2 [172] with simple charge state assignment accomplished by MS2ZAssign. The resulting ms2 files were then searched against a provided protein database using DBDigger with the MASPIC scorer [173] with the following parameters: digestion sites after K and R; two digestion ends; Parent Mass

Tolerance, 3.0; Fragment Ion Tolerance, 0.5; and any PTMs that were specified as described in Chapter 5. DBDigger results files are in the form of .sqt files; which are imported into DTASelect for the following filtering rules for unmodified datasets: 2 peptides required for protein identification, 25 for +1 peptides, 30 for +2 peptides, and 45 for +3 peptides. Chapter 5 has a discussion on the optimization of the filtering rules when considering DBDigger searches including PTMs.

*The InsPecT algorithm*

InsPecT was specifically designed to search MS/MS datasets containing PTM information [40]. The Bafna group at University of California-San Diego developed this algorithm. This algorithm takes a different approach for searching MS/MS data against a provided search database. The first stage generates a set of three amino acids in length tags (25 were specified in the searches performed in this dissertation) using fragment ions in the experimental MS/MS and enables a relatively short list of peptides to be searched. A list of candidate peptides are scored based on seven different criteria: the number of predicted (1) b and (2) y ions that match to the MS/MS, how well the intensity of the identified (3) b and (4) y ions match the predicted intensity, (5) trypsin specificity, (6) the length of the candidate matching peptide where the presence of PTM(s) indicates a shorter peptide, and (7) the fragment ion profile of the spectrum (presence of an isotope, higher fragment ion intensity in the middle of the spectrum, and properties of the neighboring residue) [40]. The resulting top score (MQScore) is then compared to the distribution of the lower scores (DeltaScore) and a p-value is calculated. The p-value used by InsPecT is based on the p-value devised for the algorithm Peptide Prophet [174].

A p-value representing an ~2% FDR is chosen based on the p-value distribution for a

particular MS/MS dataset as discussed in detail in Chapter 5.

An in-house perl script used to automate Sequest searches, gitrseq.pl, was adapted

for automation of InsPecT searches and named gitrinspect.pl.  The perl script

gitrinspect.pl operates in a similar manner to gitrseq.pl with a few notable exceptions as

follows.  Figure 2.5 is a flow chart outlining the procedure of gitrinspect.pl from a

Xcalibur raw file to a tab-delimitated results file.  The script automates the process from

raw file to InsPecT results in four stages, creating a queue for each raw file to be searched

sequentially.  First, the conversion of the .raw file to the .mzXML format is performed

using the ReAdW software (Institute for Systems Biology), which is the proper file

format for InsPecT input.  Next, the mzXML file is moved to a sub-folder named

temp_dir.  Once the mzXML file is in temp_dir, searching and scoring the MS/MS using

InsPecT is completed.  Finally, the mzXML and the .txt InsPecT results files are removed

from temp_dir and placed in the directory above (inspect_searches).  Once these files are

removed, temp_dir is deleted and the loop starts over again and repeated for each raw file

present in the directory.  This process is necessary due to the limited memory availability

of the InsPecT algorithm on a desktop computer.  InsPecT attempts to load all MS/MS

spectra in a given folder to be searched, thus leading to over 100,000 MS/MS being

loaded into memory.  Feasibly, only one raw file is searched at a time, leading to the

above necessity of creating a sub-folder to perform the actual InsPecT search.
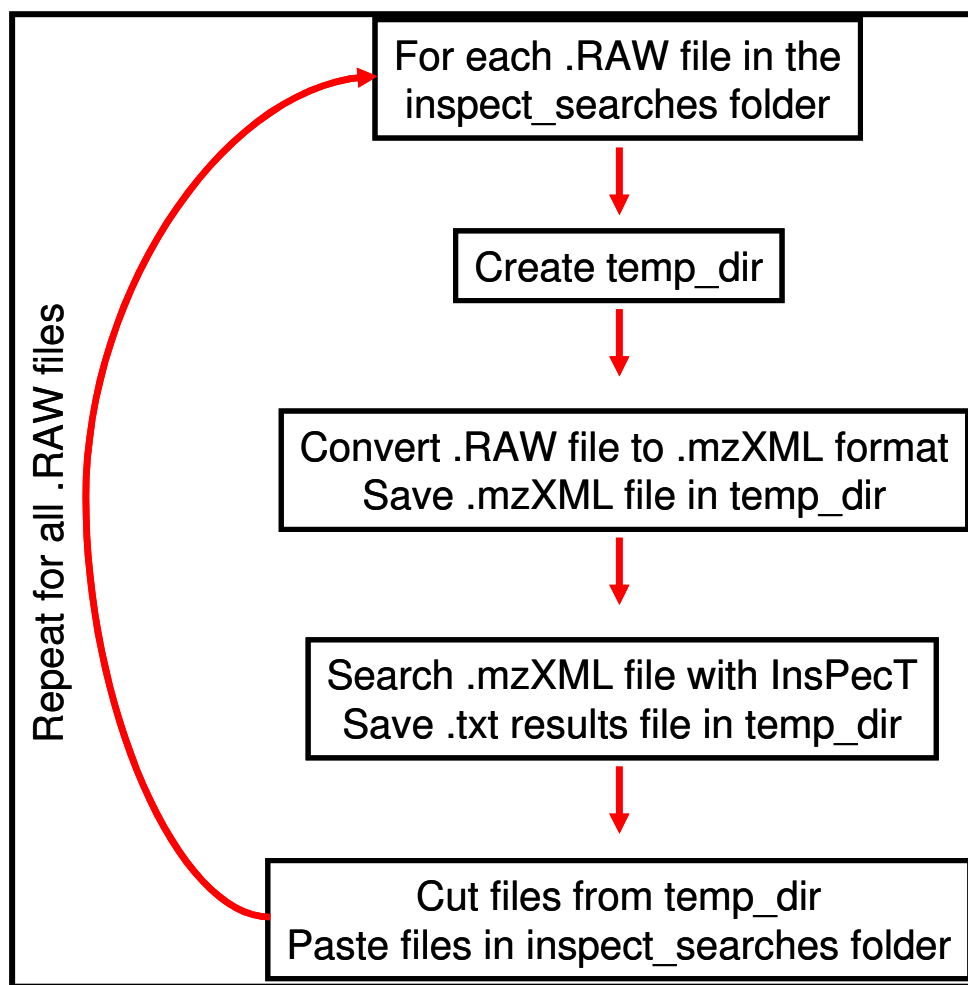
**Figure 2.5.** Flow diagram outlining the scheme of the gitrinspect.pl perl script.

The .RAW files are located in the inspect_searches folder. Creating the temp_dir folder

allows InsPecT to search one .mzXML file at a time without overwhelming

computational capacity on a desktop computer.

*Semiquantitation of protein abundance*

There are many different methods of quantifying MS/MS data available; where stable isotope labeling [175, 176], isotope-coded affinity tags (ICAT) labeling cysteine residues [177, 178], and isobaric tagging for relative and absolute quantification (iTRAQ) [178-180] are some of the common examples of relative quantification by changing the mass of a peptide from one growth state to compare to a separate growth conditions. One method of absolute quantification is the use of AQUA peptides [181], which are synthetic peptides created with heavy isotopes that have the exact amino acid sequence as an endogenous peptide of interest. However, the microbial samples utilized in this dissertation were not amenable to the aforementioned methods as discussed in detail in Chapter 3. Thus, the semiquantitation method was used, which takes into account reproducible differences in spectra count, peptide count and sequence coverage. Spectral count is the total number of MS/MS identified as encompassing amino acid sequence from a particular protein and can be a measurement for relative quantification [182, 183]. The definition of peptide count is the total number peptides identified for a given protein and sequence coverage is the total percent of the protein sequence identified from the MS/MS data. The use of dynamic exclusion does not bias spectra count measurement, because all ions that are isolated are placed on the dynamic exclusion list.

Due to differences in dynamic range discussed in Chapter 4 of this dissertation, different quantification parameters are applied for the LCQ and LTQ datasets. LCQ datasets require the following parameters: $\geq$30% change in sequence coverage, $\geq$4 unique peptides identified, and/or at least a 2x difference in spectra count between the treated and control samples (adapted from Refs. [184] and [97]). LTQ datasets require

the following semi-quantitation rules applied: $\geq$40% change in sequence coverage, $\geq$5 unique peptides identified, and/or at least a 2x difference in spectra count [147, 158].

*Peptide and protein false discovery rates*

The peptide false discovery rate (FDR) can be calculated from using a distracter database (an unrelated proteome database), a reversed database (all protein sequences have been reversed) or a shuffled database (all protein sequences are scrambled). The most popular database types for calculating the peptide FDR are using a reversed database or a shuffled database. The argument for using either one of these two database types is both contain nonsense sequences to calculate the FDR. If the distracter database is used, caution must be taken that the homology of the sequences between the true database and the distracter database do not result in true identifications from sequences comprising the distracter database (i.e. ribosomal proteins, DNA polymerase, etc.). A study by Peng *et al* [185] describe the formula used to determine the peptide FDR for a MS/MS dataset using a reversed database. The formula is %FDR = $2 * (n_{rev}/n_{tot} * 100)$, where $n_{rev}$ is the number of peptides identified that match the reversed database and $n_{tot}$ is the sum of the peptides identified matching the reversed and forward databases [185]. The basis for multiplying the end value by 2 is the assumption that if a peptide sequence score from the reversed database is large enough to pass the filter thresholds then there is an incorrect peptide sequence score from the forward database that also was large enough to pass the filter threshold.

The resulting FDR at the protein level can be much higher (~10-fold) than the corresponding peptide FDR, resulting from the MS/MS identification confidence level. The protein FDR can be calculated from the peptide FDR as a function of the probability

of the possibilities. Illustrated in Figure 2.6 is the predicted FDR at the protein level

based on a 2% FDR at the peptide level using the requirement of 2 and 3 peptides for a

positive identification. Figure 2.6 displays the equation that determines the FDR for the

resulting proteins. For example, the requirement of at least two peptides for protein

identification leads to four different possibilities and eight total peptide possibilities.

Applying a 2% peptide FDR corresponding to 98% true identifications, results in a

protein FDR of 11.5%. This value agrees with the experimental FDR values found in

Chapters 3-5. If a one-peptide requirement were applied to the Figure 2.6 equation, there

would be two possibilities for both values x and z, yielding a FDR at the protein level to

be 48%. This is in agreement with the protein FDRs found in Chapter 4. The ability to

predict the protein FDR based on the value of the peptide FDR is helpful in determining

proper threshold filters to be applied to a MS/MS dataset.

      However, if for instance the cutoff was 5 or 6 peptides to be required for protein

identification, the resulting FDR value at the protein level would be 0.9% and 0.64%,

respectively. This decreases the resulting protein FDR as a consequence of requiring a

larger number of true peptide identifications, with the protein FDR no longer of much

consequence. These calculations only apply to results from one single MS experiment.

If a protein is required to be identified from peptides from multiple MS experiments, the

resulting FDR is even lower than as calculated above. As an example, the two-peptide

requirement protein FDR is calculated to be 11.5%, however if the two peptides are

required to be identified in two MS experiments the resulting FDR is 1.3%. If the one-

peptide requirement were to be applied to this same scenario leading to one peptide in

each MS experiment, the resulting protein FDR is now 23%. Therefore, using

$$\frac{(y)^x}{z} \times 100 = \text{protein FDR}$$

2 peptides required
TT, TF, FT, FF
y = 0.98
x = 4
z = 8

3 peptides required
TTT, TFT, TTF, TFF,
FTT, FFT, FTF, FFF
y = 0.98
x = 8
z = 24

$$\frac{(0.98)^4}{8} \times 100 = 11.5\%$$

$$\frac{(0.98)^8}{24} \times 100 = 3.5\%$$

**Figure 2.6.** Predicted protein FDR based on the peptide FDR classification.

The values in the equation represent the peptide true discovery rate (y), the total number

of true and false identification possibilities for the peptide requirement (x), and the total

number of true and false identification possibilities for a given peptide (z).

a more stringent filter leads to a total protein FDR lower if the peptide requirements are from two MS experiments.

## Conclusions

By improving MS sample preparation methods and data mining filter thresholds, we achieved a more in-depth proteome that would not have been possible otherwise. Through the incorporation of two-dimensional liquid chromatographic separation and the improved dynamic range capabilities of the LTQ, the global proteome studies presented in this dissertation would not have yielded as much depth into the microbial life processes. In addition, the increased dynamic range of the LTQ provided the ability to perform quantification comparisons between cultures of *S. oneidensis* MR-1 following acute and chronic Cr(VI) exposure yielding interesting information on differentially expressed proteins that otherwise would be overlooked using other technologies.

**Chapter 3**

**Understanding Global Chromate Response of *Shewanella oneidensis* MR-1 under**

**Acute and Chronic Exposure Using Shotgun Proteomics**

*All of the data presented below has been adapted from the following published journal articles' Results and Discussion text sections*

Stephen D. Brown, Melissa R. Thompson, Nathan C. VerBerkmoes, Karuna Chourey, Manesh Shah, Jizhong Zhou, Robert L. Hettich, and Dorothea K. Thompson. Molecular Dynamics of the *Shewanella oneidensis* Response to Chromate Stress. *Molecular and Cellular Proteomics,* 2006; 5, 1054-1071. *Sample preparation, LTQ measurements, and data analysis for the proteomics portion of the manuscript were performed by Melissa R. Thompson with assistance from Nathan C. VerBerkmoes. Supplemental material can be found at http://compbio.ornl.gov/shewanella_metal_stress/stress.*

Karuna Chourey, Melissa R. Thompson, Jennifer Morrell-Falvey, Nathan C. VerBerkmoes, Stephen D. Brown, Manesh Shah, Robert L. Hettich, Jizhong Zhou, Mitchel Doktycz, and Dorothea K. Thompson. Global Molecular and Morphological Effects of 24-h Chromium Exposure on *Shewanella oneidensis* MR-1. *Applied and Environmental Microbiology*, 2006; 72, 6331-6344. *Sample preparation, LTQ measurements, and data analysis for the proteomics portion of the manuscript were performed by Melissa R. Thompson with assistance from Nathan C. VerBerkmoes on data analysis. Supplemental material can be found at http://compbio.ornl.gov/shewanella_ metal_stress/chronic.*

**Introduction**

The global molecular response of *Shewanella oneidensis* MR-1 to chromate

exposure was unexplored prior to this dissertation work. The need to understand both the

initial response and prolonged exposure to hexavalent chromium [Cr(VI)] in the form of

chromate ($CrO_4^{2-}$) for the purposes of using *S. oneidensis* as an agent for bioremediation

has not been examined previously. The goal of the work presented in this chapter was to

determine the global protein response in the form of a mass spectrometry based proteome

dataset to reveal novel proteins that are indicative of Cr(VI) response. The proteome

studies described here impart a global insight into temporal alterations in protein content occurring in response to a toxic acute level (1mM) and following long term chronic exposure of a lower dosage of chromate (0.3 mM). The overall protein expression pattern of an organism will not change a great deal in response to a particular stimulus; however this chapter identifies novel proteins found to be differentially expressed during the bacterial response to a sub-lethal dosage of chromate. The responses identified will increase the understanding of possible protein expression mechanisms that allow Cr(VI) reduction, detoxification, and adaptation under aerobic growth conditions. Alterations in protein expression were complemented results from the microarray analysis. However, microarray analysis discussion is limited to only where it is relevant to the proteome results.

## **Materials and Methods**

*Preparation of Proteomes for HPLC-MS/MS Analysis*

For the acute chromate shock large scale proteomic characterization, 500-ml cultures of *S. oneidensis* MR-1 in 4-liter flasks (a total of 1 liter of culture for treatment and control) were grown to midexponential phase ($A_{600}$, 0.5) under aerobic conditions and then either exposed to a final $K_2CrO_4$ concentration of 1 mM or allowed to continue growing in the absence of added chromate. At time points of 45 and 90 min after treatment with $K_2CrO_4$, cells were harvested from each of the following conditions for HPLCMS/MS analysis: 1) Control 1 (untreated midlog phase cells after 45 min of further growth), 2) Treatment 1 (45 min post-Cr addition), 3) Control 2 (untreated midlog phase cells after 90 min of further growth), and 4) Treatment 2 (90 min post-Cr addition). For this, cells were pelleted by centrifugation (5,000 x *g* for 5 min), resuspended in ice-cold

LB medium, washed two times in 50 mM Tris, 10 mM EDTA (pH 7.6), and centrifuged at 5,000 x $g$ for 10 min. For the chronic exposure large scale proteome characterization, *S. oneidensis* cells were first adapted for 24 h to chromate by growth in the presence of 0.3 mM $K_2CrO_4$ in a starter culture. An aliquot of the starter culture was used to inoculate fresh media with or without 0.3 mM $K_2CrO_4$. Cells were then incubated for 24 h and harvested by centrifugation for mass spectrometry analysis. The *S. oneidensis* cells were then placed on ice and lysed by sonication as detailed in Chapter 2. Approximately 2 mg of each proteome fraction (soluble and membrane) prepared from all of the growth conditions was denatured and reduced in 6 M guanidine and 10 mM DTT. Denaturation and reduction were followed by an overnight digestion (see Chapter 2).

*LC/LC-MS/MS Analysis*

The proteome fractions (soluble and membrane) prepared from control and chromate-treated samples were analyzed in duplicate via two-dimensional (2-D) LC-MS/MS experiments using an Ultimate HPLC system (LC Packings, a division of Dionex, San Francisco, CA) coupled to a linear trapping quadrupole (LTQ) mass spectrometer (ThermoFinnigan, San Jose, CA). See Chapter 2 for details of chromatographic separation of peptides. The multiphasic column system was positioned in front of the LTQ on a nanospray source (ThermoFinnigan). All samples were analyzed via a 24-h 12-step 2-D analysis consisting of increasing concentration (0–500 mM) salt pulses of ammonium acetate followed by 2-h reverse phase gradients (see Chapter 2 for details). During the entire chromatographic process, the LTQ was operated in a data-dependent MS/MS mode detailed in Chapter 2.

*Proteome Bioinformatics*

A protein database was created by combining the most recent version of the *S. oneidensis* MR-1 database (Version 8; www.tigr.org/) containing a total of 4,798 predicted proteins with 36 common contaminants (trypsin, keratin, etc.). The database can be downloaded from the website compbio.ornl.gov/shewanella_metal_stress/ databases/. For all database searches, the MS/MS spectra were searched using Sequest (Ref. [39]; ThermoFinnigan) with the following parameters: enzyme type, trypsin; parent mass tolerance, 3.0; fragment ion tolerance, 0.5; up to four missed cleavages allowed; fully tryptic peptides only. The output data files were then filtered and sorted with the DTASelect algorithm [171] using the following parameters: fully tryptic peptides only with ΔCN of at least 0.08 and cross-correlation scores (Xcorrs) of at least 1.8 (+1), 2.5 (+2), and 3.5 (+3). These threshold scores have been tested rigorously in our laboratory and provide a high confidence of identification (see Refs. [184] and [65] for more discussion) with a maximum false-positive rate of 1–2%. Post-translational modifications and other fixed modifications (*i.e.* due to addition of iodoacetamide) were not included in the search parameters. DTASelect files are available on the analysis page (compbio.ornl.gov/shewanella_metal_stress/) under the corresponding acute shock or chronic dataset and are filtered at one peptide and two peptides per protein. The files are presented in a text format or a viewable html version where every identified spectrum can be viewed by clicking on the spectral number (first column, labeled by filename). The DTASelect results from all control and chromate-treated samples were then compared with the Contrast program [171] for each time point. These results are located under the global contrast heading on the analysis page. A list was made of all proteins showing a

54

reproducible significant change of at least 40% sequence coverage, five or more unique peptides, and/or a reproducible spectral count difference of 2x between the control and chromate-treated samples at each time point (adapted from Refs. [184] and [97]). The analysis page also contains inter-run contrast files (compares duplicate runs on same sample) as well as fractionation comparisons (compares replicate runs of the same proteome broken down by fraction).

## Results and Discussion

*Acute Shock Proteome Dataset*

Cellular fractions from each of the four *S. oneidensis* growth conditions (*i.e.* Control 1, Control 2, Treatment 1, and Treatment 2) were analyzed in duplicate using 2-D LC-ES-MS/MS with a linear trapping quadrupole mass spectrometer (see Chapter 2 for details). A total of 2,370 of the 4,931 total predicted genes in the *S. oneidensis* MR-1 genome were identified with at least two peptides (Table 3.1), representing 48% of the theoretical proteome. Due to the large number of false positives possible at the one-peptide filter level [65], we present a rigorous analysis of the two-peptide dataset only. High stringency filtering was used in this study, giving a maximum false-positive rate of 1–2%. The reproducibility between duplicate protein analyses on the LTQ was as follows: 78.6% (chromate-shocked) and 78.4% (control) for the 45-min poststress time point and 77.7% (chromate-shocked) and 73.5% (control) for the 90-min time point. This level of reproducibility is necessary for semiquantification. Variation is likely due to low abundance proteins identified with two peptides in one of the analyses and only one peptide in another, thereby being filtered out. Although previous studies using 2-D PAGE and LC-MS/MS have been used for global *S. oneidensis* proteome studies

55

**Table 3.1.** Proteome Analysis of Chromate-Shocked *S. oneidensis* MR-1

| Condition | No. proteins identified 1 pep[a] | No. proteins identified 2 pep[b] | Av. Sequence coverage[c] |
|---|---|---|---|
| 45 min Control | 2610 | 1911 | 37.24% |
| 45 min Shock | 2644 | 1959 | 37.20% |
| 90 min Control | 2595 | 1892 | 36.38% |
| 90 min Shock | 2664 | 1992 | 36.45% |
| **Total** | **2954** | **2370** | |

[a]Total proteins identified with at least 1 peptide per protein from duplicate runs. [b]Total proteins identified with at least 2 peptides per protein from duplicate runs. [c]Average sequence coverage per protein at the 2 peptide level.

[25, 97, 186] this study represents, to our knowledge, one of the largest measurement of the *S. oneidensis* proteome published to date. The entire list of identified proteins with total sequence coverage, functional categories, pI, and molecular weight information is given in Supplemental Table S2. No major biases were found between the pI and molecular weights of the predicted proteome from the genome and the observed proteome (Supplemental Table S2). The identified proteins with their peptide count (number of identified peptides), spectral count (number of MS/MS spectra identified per protein), and percent sequence coverage (total percentage of the protein sequence covered by tryptic peptides) for the different growth conditions and individual analyses can be found in Supplemental Table S3.

Proteins identified at the two-peptide level were grouped according to the functional categories assigned by the J. Craig Venter Institute annotation [formerly The Institute for Genomic Research (Ref. [69]; www.tigr.org, Comprehensive Microbial Resource)] (Table 3.2). Proteins with assigned functions in amino acid biosynthesis, cellular processes, protein fate, protein synthesis, nucleotide metabolism, and transcription were found with greater than 70% identified. More than 90% of the proteins comprising the functional classes of protein synthesis, nucleotide metabolism, and transcription were identified, representing an almost complete characterization of these categories at the proteome level. Proteins generally thought to be of lower abundance, such as those with assigned functions in signal transduction and transcriptional regulation, were identified at levels of 64 and 52%, respectively, of the total number of proteins per category. A total of 624 of the 2,039 predicted hypothetical proteins were identified. Of the 624 proteins in this functional category, 209 were

**Table 3.2.** Functional Categories of Identified Proteins with at Least Two Peptides

| Number | Functional Category | Observed Proteome | Predicted Proteome | Percent Identified |
|---|---|---|---|---|
| 1 | Amino acid biosynthesis | 70 | 91 | 76.92% |
| 2 | Biosynthesis of cofactors, prosthetic groups, and carriers | 94 | 121 | 76.69% |
| 3 | Cell envelope | 121 | 180 | 67.22% |
| 4 | Cellular processes | 188 | 260 | 72.31% |
| 5 | Central intermediary metabolism | 28 | 51 | 54.90% |
| 6 | DNA metabolism | 97 | 144 | 67.36% |
| 7 | Energy metabolism | 199 | 308 | 64.61% |
| 8 | Fatty acid and phospholipid metabolism | 48 | 65 | 73.85% |
| 9 | Hypothetical proteins | 624 | 2039 | 30.60% |
| 10 | Mobile and extrachromosomal element functions | 34 | 317 | 10.73% |
| 11 | Protein fate | 137 | 185 | 74.05% |
| 12 | Protein synthesis | 131 | 141 | 92.91% |
| 13 | Purines, pyrimidines, nucleosides, and nucleotides | 56 | 62 | 90.32% |
| 14 | Regulatory functions | 104 | 199 | 52.26% |
| 15 | Signal transduction | 39 | 61 | 63.93% |
| 16 | Transcription | 49 | 54 | 90.74% |
| 17 | Transport and binding proteins | 148 | 274 | 54.01% |
| 18 | Unknown function | 203 | 379 | 53.56% |
| | **Total** | **2370** | **4931** | **48.06%** |

annotated as hypothetical proteins, and 415 were annotated as conserved hypothetical proteins. This represents one of the largest identifications of hypothetical protein expression for a microbial proteome to date.

Although absolute quantification at the global proteome level was not feasible, semiquantification of differentially expressed proteins between the chromate-treated and control samples could be accomplished by using a combination of percent sequence coverage, number of unique peptides, and spectral count from the mass spectra [20, 97, 184] (see Chapter 2 for details). Other methods for relative quantification, such as ICAT [177], were considered before the method used here was selected. ICAT labels the cysteine residues of a protein, and in *S. oneidensis* greater than 50% of the predicted proteins have two or fewer cysteine residues present. Moreover almost 20% of the predicted proteome does not contain any cysteine residues. This method would not detect almost 70% of the predicted proteome confidently; thus ICAT was not deemed appropriate for this study.

Comparisons were made between the transcriptome and proteome data to determine the relationship between gene and protein expression. However, in this dissertation, only the proteome data will be described in detail. Both up- and down-regulated proteins were measured by comparing the chromate-shocked cell samples to their respective control samples. Using the semiquantitative criteria discussed above, we identified 78 proteins (Supplemental Table S4) as being differentially expressed in response to chromate (Tables 3.3 and 3.4, Supplemental Table S4). Supplemental Table S4 divides the proteins identified at 45 min post-shock from the proteins identified at 90 min post-shock. We propose that reproducibility of at least 70% between replicate

**Table 3.3.** Up-Regulated Proteins Identified in the 90 min Chromate-Shocked Sample

| Gene | Transcriptome | Control 1 | Control 2 | Chrom 1 | Chrom 2 | Category | Description |
|---|---|---|---|---|---|---|---|
| SO0343 | No change[a] | 20.8 | 18.3 | 43.3 | 45.7 | 7 | aconitate hydratase 1 (*acnA*) |
| SO0423 | Induced[b] | 11.2 | 0.0 | 43.6 | 39.2 | 14 | pyruvate dehydrogenase complex repressor (*pdhR*) |
| SO0798 | Induced | 0.0 | 0.0 | 24.7 | 21.7 | 9 | conserved hypothetical protein |
| SO0934 | No change | 12.2 | 55.5 | 87.8 | 76.3 | 9 | conserved hypothetical protein |
| SO1045 | No change | 0.0 | 0.0 | 22.3 | 24.5 | 9 | hypothetical protein |
| SO1114[c] | Induced | 0.0 | 0.0 | 28.3 | 55.2 | 18 | DNA-damage-inducible protein P (*dinP*) |
| SO1178 | No change | 11.7 | 11.3 | 45.7 | 41.6 | 17 | magnesium and cobalt efflux (*corC*) |
| SO1190[c] | Induced | 39.0 | 20.2 | 55.5 | 51.8 | 9 | conserved hypothetical protein |
| SO1482[c] | Induced | 30.2 | 28.8 | 82.7 | 80.5 | 17 | TonB-dependent receptor, putative |
| SO1576 | No change | 20.7 | 0.0 | 39.2 | 44.6 | 4 | glutathione S-transferase family protein |
| SO1580[c] | Induced | 5.0 | 14.0 | 29.0 | 45.3 | 17 | TonB-dependent heme receptor |
| SO1755 | Induced | 22.3 | 10.6 | 34.6 | 49.6 | 7 | phosphoglucomutase/phosphomannomutase family protein |
| SO2290 | No change | 24.9 | 17.4 | 49.8 | 48.0 | 18 | rhodanese domain protein |
| SO2426[c] | Induced | 0.0 | 0.0 | 37.1 | 48.1 | 15 | DNA-binding response regulator |
| SO2577 | No change | 6.3 | 22.3 | 41.3 | 48.7 | 4 | septum site-determining protein (*minD*) |
| SO2912[c] | Induced | 38.3 | 25.9 | 64.7 | 55.1 | 7 | formate acetyltransferase (*pflB*) |
| SO2915 | Induced | 15.5 | 12.3 | 49.4 | 43.6 | 5 | acetate kinase (*ackA*) |
| SO3030[c] | Induced | 12.9 | 0.0 | 62.4 | 66.1 | 17 | siderophore biosynthesis protein (*alcA*) |
| SO3032[c] | Induced | 5.7 | 0.0 | 31.4 | 32.3 | 17 | siderophore biosynthesis protein, putative |
| SO3033[c] | Induced | 12.0 | 6.7 | 57.6 | 57.3 | 17 | ferric alcaligin siderophore receptor |
| SO3061 | Induced | 6.1 | 3.1 | 31.3 | 37.9 | 6 | DNA topoisomerase III (*topB*) |
| SO3407 | Induced | 9.3 | 8.8 | 24.7 | 19.1 | 9 | conserved hypothetical protein |
| SO3462[c] | Induced | 0.0 | 6.0 | 35.1 | 34.1 | 6 | DNA repair protein RecN (*recN*) |
| SO3585[c] | Induced | 0.0 | 0.0 | 20.1 | 23.0 | 4 | azoreductase, putative |
| SO3586[c] | Induced | 0.0 | 0.0 | 60.1 | 20.3 | 18 | glyoxalase family protein |
| SO3599 | Induced | 12.8 | 18.5 | 68.7 | 64.8 | 17 | sulfate ABC transporter, periplasmic protein (*cysP*) |
| SO3667[c] | Induced | 0.0 | 0.0 | 91.9 | 96.8 | 9 | conserved hypothetical protein[d] |
| SO3669[c] | Induced | 9.3 | 11.6 | 71.6 | 77.9 | 17 | heme transport protein (*hugA*) |
| SO3670[c] | Induced | 0.0 | 0.0 | 13.5 | 18.6 | 17 | TonB1 protein (*tonB1*) |
| SO3671 | Induced | 0.0 | 0.0 | 20.4 | 26.5 | 17 | TonB system transport protein (*exbB1*) |
| SO3673[c] | Induced | 0.0 | 0.0 | 63.1 | 61.2 | 17 | hemin ABC transporter, periplasmic (*hmuT*) |
| SO3675[c] | Induced | 0.0 | 0.0 | 72.9 | 59.0 | 17 | hemin ABC transporter, ATP-binding protein (*hmuV*) |
| SO3723 | Induced | 0.0 | 0.0 | 50.2 | 29.8 | 5 | adenylylsulfate kinase (*cysC*) |

**Table 3.3.** Continued

| Gene | Transcriptome | Control 1 | Control 2 | Chrom 1 | Chrom 2 | Category | Description |
|------|---------------|-----------|-----------|---------|---------|----------|-------------|
| SO3726 | Induced | 22.1 | 7.8 | 49.7 | 49.3 | 5 | sulfate adenylyltransferase, subunit 1 (*cysN*) |
| SO3727 | Induced | 38.7 | 20.2 | 68.5 | 68.5 | 5 | sulfate adenylyltransferase, subunit 2 (*cysD*) |
| SO3737 | Induced | 30.8 | 22.1 | 62.1 | 73.3 | 5 | sulfite reductase (NADPH) hemoprotein (*cysI*) |
| SO3738 | Induced | 5.9 | 0.0 | 30.8 | 27.3 | 5 | sulfite reductase (NADPH) flavoprotein (*cysJ*) |
| SO3907 | No change | 0.0 | 0.0 | 48.1 | 60.0 | 9 | conserved hypothetical protein |
| SO3913[c] | Induced | 0.0 | 0.0 | 27.2 | 32.6 | 9 | conserved hypothetical protein |
| SO3914[c] | Induced | 18.5 | 15.8 | 68.1 | 67.8 | 17 | TonB-dependent receptor, putative |
| SO4077 | Induced | 14.5 | 14.0 | 36.7 | 37.1 | 17 | TonB-dependent receptor, putative |
| SO4516 | Induced | 0.0 | 8.8 | 37.7 | 32.6 | 17 | ferric vibriobactin receptor (*viuA*) |
| SO4523 | Induced | 39.4 | 50.5 | 70.8 | 74.4 | 17 | iron-regulated outer membrane virulence (*irgA*) |
| SO4651[c] | Induced | 0.0 | 0.0 | 41.4 | 77.6 | 9 | conserved hypothetical protein |
| SO4652[c] | Induced | 0.0 | 0.0 | 46.6 | 62.0 | 17 | sulfate ABC transporter, periplasmic (*sbp*) |
| SO4655[c] | Induced | 0.0 | 0.0 | 54.5 | 55.1 | 17 | sulfate ABC transporter, ATP-binding (*cysA-2*) |
| SO4743 | Induced | 50.9 | 49.9 | 71.7 | 67.7 | 17 | TonB-dependent receptor, putative |
| SOA0042 | No change | 0.0 | 0.0 | 46.9 | 46.3 | 9 | hypothetical protein |

[a]No change in expression at the 90 min time point in the microarray analysis. No change defines a gene that was found to not have an induction or repression of expression at any time point on the microarray. [b]Induced defines a gene that exhibited at least two-fold induction at the 90 min time point. [c]Proteins also found to be up-regulated in the 45 min chromate-exposure samples. [d]The functional annotation of SO3667 was revised recently to a heme iron utilization protein, HugZ [187].

**Table 3.4.** Down-Regulated Proteins Identified in the 90 min Chromate-Shocked Samples

| Gene | Transcriptome | Control 1 | Control 2 | Chrom 1 | Chrom 2 | Category | Description |
|------|---------------|-----------|-----------|---------|---------|----------|-------------|
| SO0398 | Repressed[a] | 42.4 | 18.1 | 0.0 | 4.6 | 7 | fumarate reductase flavoprotein subunit (*frdA*) |
| SO0404 | Repressed | 78.8 | 77.9 | 55.4 | 57.2 | 9 | hypothetical protein |
| SO0548 | Repressed | 70.0 | 76.7 | 47.8 | 51.1 | 6 | DNA-binding protein, HU family |
| SO0847 | Induced | 45.9 | 28.7 | 0.0 | 0.0 | 7 | iron-sulfur cluster-binding protein (*napG*) |
| SO0848[b] | Induced[c] | 73.2 | 52.9 | 43.8 | 46.1 | 7 | periplasmic nitrate reductase (*napA*) |
| SO0902 | Repressed | 49.8 | 37.5 | 26.5 | 18.4 | 7 | NADH:ubiquinone oxidoreductase, (*nqrA-1*) |
| SO0970 | Repressed | 59.9 | 57.4 | 34.7 | 34.9 | 7 | fumarate reductase flavoprotein subunit precursor |
| SO1111 | Repressed | 75.8 | 70.1 | 35.7 | 42.7 | 17 | bacterioferritin subunit 2 (*bfr2*) |
| SO1405 | No change[d] | 40.1 | 31.4 | 11.1 | 0.0 | 18 | transglutaminase family protein |
| SO1429 | No change | 42.8 | 26.8 | 0.0 | 0.0 | 7 | anaerobic dimethyl sulfoxide reductase, A (*dmaA-1*) |
| SO1430 | No change | 31.2 | 36.2 | 0.0 | 0.0 | 7 | anaerobic dimethyl sulfoxide reductase, B (*dmsB-1*) |
| SO1490 | Repressed | 67.3 | 70.2 | 27.2 | 39.8 | 7 | alcohol dehydrogenase II (*adhB*) |
| SO1518 | Repressed | 80.4 | 81.5 | 48.1 | 52.4 | 9 | conserved hypothetical protein |
| SO1776 | Repressed | 48.5 | 41.9 | 25.7 | 31.0 | 3 | outer membrane protein precursor (*mtrB*) |
| SO1777 | Repressed | 8.7 | 6.9 | 0.0 | 0.0 | 7 | decaheme cytochrome c MtrA (*mtrA*) |
| SO1778 | Repressed | 43.4 | 55.1 | 25.0 | 25.2 | 7 | decaheme cytochrome c (*omcB*) |
| SO1779 | Repressed | 51.8 | 53.5 | 26.7 | 29.9 | 7 | decaheme cytochrome c (*omcA*) |
| SO2469 | Repressed | 22.1 | 24.7 | 10.0 | 12.1 | 18 | conserved hypothetical protein |
| SO2490 | Repressed | 51.8 | 66.9 | 33.5 | 28.9 | 14 | transcriptional regulator, RpiR family |
| SO2929[b] | Repressed | 73.8 | 73.0 | 36.3 | 40.0 | 9 | hypothetical protein |
| SO3538[b] | Repressed | 46.9 | 34.7 | 0.0 | 0.0 | 14 | transcriptional regulator HlyU (*hlyU*) |
| SO3565 | Repressed | 72.0 | 59.9 | 36.8 | 39.6 | 13 | 2,3-cyclic-nucleotide 2-phosphodiesterase (*cpdB*) |
| SO3920 | Repressed | 21.7 | 27.6 | 0.0 | 0.0 | 7 | periplasmic Fe hydrogenase, large subunit (*hydA*) |
| SO3967 | No change | 71.6 | 56.4 | 23.0 | 24.5 | 17 | molybdenum ABC transporter, periplasmic protein |
| SO4513[b] | No change | 44.3 | 36.2 | 3.3 | 3.5 | 7 | formate dehydrogenase, alpha subunit |
| SO4561 | No change | 37.7 | 47.4 | 0.0 | 0.0 | 9 | conserved hypothetical protein |

[a]Repressed is a gene exhibiting at least two-fold repression at the 90 min time point. [b]Proteins down-regulated in both the 45 and 90 min chromate-exposure experiments. [c]Induced is a gene exhibiting at least two-fold induction at the 90 min time point.

[d]No change in expression at the 90 min time point in the microarray analysis.

analyses at the protein level is necessary for a successful determination of differentially expressed proteins where a protein must be detected in both of the replicate analyses to be included as a candidate for semiquantification. For the 45-min chromate treatment analysis, a total of 24 proteins were found to be up-regulated with six additional proteins being down-regulated relative to the control sample (Supplemental Table S4). The genes for 23 of these 24 up-regulated proteins showed corresponding induction levels at the 30- and 60-min time points based on microarray hybridization. A putative formate acetyltransferase (encoded by *pflB*), which was identified as being up-regulated at the protein level, was found to be induced at the transcript level at the 30-min treatment time but repressed at the 60-min time point. The subset of proteins identified as up-regulated were dominated by species (12 total) assigned to the functional category of transport and binding proteins and included TonB-dependent receptors, siderophore biosynthesis proteins, heme transport proteins, and TonB1. Proteins classified as hypothetical were also dominant with a total of six conserved hypothetical proteins identified as up-regulated. For the six proteins down-regulated in the 45-min shocked sample relative to the control condition, five were found to be repressed at the mRNA level by microarray analysis, and the other protein, a hypothetical protein (SOA0141), revealed no change at the transcript level. Proteins that were measured as being down-regulated in the 45-min chromate-shocked samples consisted of three hypothetical proteins (SO1124, SO2929, and SOA0141), a periplasmic nitrate reductase (NapA), a transcriptional regulator (HlyU), and the α subunit of formate dehydrogenase (SO4513).

Proteomic analysis of the 90-min chromate treatment samples revealed 48 up-regulated proteins and 26 down-regulated proteins relative to the control sample

63

(Supplemental Table S4 and Tables 3.3 and 3.4). After 90 min of chromate exposure, 39 of the 48 proteins up-regulated under Cr(VI) conditions also were induced at the transcript level as identified by microarray analysis. Nine of the up-regulated proteins (AcnA, SO0934, SO1045, CorC, SO1576, SO2290, MinD, SO3907, and SOA0042) revealed no significant change ($p < 0.05$) in expression at the transcription level, and two of the proteins (AcnA and SO2290) expressed at higher abundance levels under chromate stress conditions did not exhibit a change in their mRNA expression levels at the 90-min time point but did at earlier time points (5, 30, and 60 min). Of the 48 up-regulated proteins, 19 proteins were from the transport and binding category, six proteins were from central intermediary metabolism, and 10 were annotated as hypothetical proteins. All of the transport and binding proteins identified as up-regulated in the 45-min chromate-shocked samples were also identified as up-regulated in the 90-min samples, whereas only four of the six conserved hypothetical proteins showed up-regulation at both postexposure time points.

Eighteen of 26 proteins down-regulated 90 min after chromate addition were also repressed at the transcript level. Of the remaining down-regulated proteins, five showed no change at the mRNA level, one protein of which was found to exhibit no mRNA change at the 90-min time point. Two down-regulated proteins (NapG and NapA) were found to be induced at the transcript level. Four of the six proteins down-regulated at the 45-min time point were also repressed at 90 min. The transcriptome data revealed an increase in the number of down-regulated energy metabolism genes over time in response to chromate. This trend was also reflected in the proteomic data in which 13 proteins

with functions in energy metabolism were identified as repressed at the 90-min time point relative to the 45-min interval.

Differences between the transcriptomic and proteomic responses of *S. oneidensis* to chromate shock are due to the stringency of filtering used in the proteome study and inherent measurement differences between microarray *versus* proteome technology. Supplemental Table S5 demonstrates the correlation of the identified and/or differentially expressed proteins to the top 100 mRNAs induced at each time point. The list is composed of a total of 194 mRNAs, which is a concatenated list representing the top 100 mRNAs at each time point (initial list was 400 genes in total before removal of redundant genes). A total of 67% of the genes in Supplemental Table S5 did not meet the differential criteria at the protein level where 45% were not identified by mass spectrometry and another 22% were not considered differentially expressed. The 67% of genes found not to meet differential expression criteria corresponds to 130 genes of 194, which is comparative in percentage to the total global analysis.

As revealed by transcriptome and proteome analyses, a major feature of the molecular response of *S. oneidensis* MR-1 to acute chromate challenge was the differential regulation of the TonB1-ExbB1-ExbD1 complex, an integral inner membrane system for iron transport, as well as other genes involved in exogenous iron acquisition. Genes that were highly induced based on time series microarray experiments and were identified as up-regulated proteins at the 45- and/or 90-min chromate treatment conditions (Table 3.3) included two putative siderophore biosynthesis proteins (AlcA and SO3032), a ferric alcaligin siderophore receptor (SO3033), HugA heme transport protein (SO3669), TonB1 (SO3670), a putative TonB-dependent receptor (SO3914), a TonB-

dependent heme receptor (SO1580), ViuA (SO4516), and HmuT (SO3673) and HmuV (SO3675) of the hemin ABC transporter. Wang and Newton [188] demonstrated that mutants harboring a deletion of the *tonB-trp* region of the *E. coli* chromosome were sensitive to chromic ion ($Cr^{3+}$) due to defective iron transport systems, and residual iron uptake by these strains was shown to be inhibited by chromic ion. Regulation of iron homeostasis is primarily carried out by the Fur protein (for a review, see Ref. [189]). It has been suggested previously that iron uptake regulation may not be the only function of Fur but that it may also serve to sequester iron to prevent the generation of highly reactive hydroxyl radicals via Fenton reactions [190]. The putative MR-1 ferritin genes, but not bacterioferritin genes, were induced in response to chromate, and these respective iron storage proteins have been suggested to have roles in short term iron flux and long term iron storage in *E. coli* [191].

In addition to iron transport genes, our global analyses demonstrated that CysP, CysC, CysD, CysN, CysI, CysJ, Sbp, and CysA-2 are up-regulated at both the mRNA and protein levels in response to chromate treatment (Table 3.3). The enhanced expression of genes encoding proteins involved in sulfate transport and assimilatory sulfate metabolism suggests the possibility of chromate-induced sulfur limitation in *S. oneidensis*, perhaps through competitive inhibition of sulfate uptake by chromate, as has been shown previously [149, 151]. Partial reduction of Cr(VI) to Cr(V) produces ROS [126, 140, 143], leading to chromate-mediated oxidative stress. Researchers working with a number of different bacteria have observed induction of genes involved in sulfur and iron homeostasis following different oxidative stresses [192-195]. A variety of explanations have been proposed including disruption of intracellular redox cycling

66

leading to insufficient sulfite reduction, a reduction in cysteine biosynthesis correlated with cell envelope damage and subsequent leakage of sulfide [196], and increased demand for low molecular weight protective thiol-containing compounds such as glutathione [197]. Alternatively induction of genes involved in sulfur metabolism and iron sequestration might represent an adaptive response to sulfur and iron limitation in MR-1 following chromate exposure.

One of the potentially interesting findings to emerge from this integrated global investigation was the co-regulated expression of a cluster of three genes (*so3585*, *so3586*, and *so3587*) at both the mRNA and protein levels. All three genes are transcribed in the same direction on the MR-1 chromosome and show a similar transcriptional profile in response to chromate with the peak in up-regulated expression occurring at the 30-min time point. The proteins encoded by two of these genes (*so3585* and *so3586*) were detected in the chromate-treated samples only (Table 3.3), suggesting that expression of SO3585 and SO3586 were differentially regulated in response to chromate stress conditions. By contrast, hypothetical protein SO3587 was found in both the control and chromate-shocked samples at the two-peptide level (Supplemental Table S3) even though the gene encoding this protein was shown to be up-regulated over the entire time course in response to chromate stress. In addition, SO3587 was found only in the membrane fractions, and a hydrophobicity plot analysis using the computer program SOSUI [198] identified a putative transmembrane domain (IGIALIFADVSLYLAYFFVGLGV) in SO3587. SO3585 and SO3586 were detected in both the soluble and membrane fractions. Based on their proximity in genome location and co-regulated expression profile within the context of chromate stress, we predict that SO3585, SO3586, and

67

SO3587 function together as a protein complex associated with the cell membrane and play an important role in the response of the cell to chromate toxicity.

*S. oneidensis* SO3585 and SO3586 are annotated as a putative azoreductase and glyoxalase family protein, respectively [69]. Glyoxalase systems are known to serve as key detoxification routes for preventing the intracellular accumulation of methylglyoxal, a natural metabolite with toxic electrophilic properties (for a review, see Ref. [199]). Azoreductases are responsible for the reductive cleavage of azo dyes, synthetic organic colorants used extensively in the textile, food, and cosmetics industries. Synthetic azo dyes are not readily reduced under aerobic conditions and are considered pollutants. Protein database searches using BLAST Version 2.2.12 [200] with the derived SO3585 primary sequence revealed ~28% sequence identity with *P. putida* ChrR and *E. coli* YieF, two soluble flavoproteins that have been demonstrated to exhibit chromate reductase activity [128, 138, 201]. Regions of conservation in the derived amino acid sequence of SO3585 included the characteristic signature, LFV$\underline{\text{TPE}}$Y$\underline{\text{N}}X_6\underline{\text{LKN}}$AID$X_2\underline{\text{S}}$ (conserved residues in SO3585 are underlined), of the NADH_dh2 family of NAD(P)H oxidoreductases (Ref. [201]; results not shown). Recently further investigation demonstrated that the *P. putida* ChrR functions as a quinone reductase and minimizes oxidative stress induced by intracellular $H_2O_2$, which is generated during the course of chromate reduction [202]. An MR-1 strain carrying an in-frame deletion mutation in the *so3585* locus has been created, and future studies will characterize it to gain insight into the functional role of SO3585 and to assess the importance of azoreductase in the *S. oneidensis* response to chromate.

Nine proteins (SO0798, SO0934, SO1045, SO1190, SO3667, SO3907, SO3913, SO4651, and SOA0042) annotated as hypothetical or conserved hypothetical met our criteria of significance for differential expression and were identified as being up-regulated in response to chromate exposure at the 45- and/or 90-min time points (Table 3.3). For four of these unknown or conserved unknown proteins (SO0934, SO1045, SO3907, and SOA0042), we observed no significant change in expression at the mRNA level, suggesting post-transcriptional regulation of these proteins in response to chromate. Our integrated transcriptome and proteome study implicates these differentially regulated proteins of unknown function in the initial response of MR-1 to toxic chromate, thus revealing gene candidates for future functional analysis.

A recent study by Kolker *et al.* [187] analyzed expression for a subset of 538 hypothetical proteins that were confidently identified in *S. oneidensis* MR-1 as a result of large scale microarray and proteomic analyses of cell samples generated under different growth conditions. A total of 788 hypothetical proteins have been identified based on the study by Kolker *et al.* [187] and the present study: 368 of these functionally undefined proteins were found in both studies, 170 were found only by Kolker *et al.* [187], and 256 were found only in this study (Supplemental Table S6). The 368 hypothetical proteins identified independently by both studies should be considered as expressed proteins, and their annotations should be changed to unknown or conserved unknown [184]. Most of these proteins were found under all growth conditions and identified in most of the replicates. Differences between the datasets revealed by this proteomic study and the one reported by Kolker *et al.* [187] are likely due to differences in the growth conditions used.

*Chronic Exposure Proteome Dataset*

Proteome measurements using LC/LC-MS/MS were taken to characterize the response of *S. oneidensis* MR-1 at the 24-h time point following initial exposure to 0.3 mM chromate. At this time point of sampling the proteome for differentially expressed proteins, the bacterium is no longer in the presence of Cr(VI) and is now presumably resuming growth with cytoplasmic Cr(III). Gene and protein expression profiles of cells exposed to chromate were compared to those of untreated control cells grown in parallel (see Chapter 2 for experimental details). The microarray data will only be presented in this section if the dataset pertains to results obtained by the proteome measurements. At the protein level, a total of 2,313 gene products, representing 47% of the predicted MR-1 proteome, were identified at the two-peptide level in duplicate analyses of control samples and samples exposed to chromate for 24 h (Table 3.5). A total of 3,051 proteins were identified using the less stringent one-peptide filter level; however, a one-peptide filter level for the identification of proteins results in a dramatically higher false-positive discovery rate [65], so a thorough analysis of only the two-peptide data is presented here. The levels of reproducibility between replicate analyses of the control proteome and the experimental (chromate-treated) proteome on the LTQ instrument were 75.6% and 77.2%, respectively. For the observed proteome identified in this study, 109 protein species were found to be differentially expressed under prolonged Cr(VI) exposure, with 56 proteins displaying increased abundance and 53 showing decreased abundance (Table 3.5). Supplementary proteome data (i.e., a list of the complete raw and filtered proteome data and a list of differentially expressed proteins) can be accessed online at compbio.ornl.gov/shewanella_metal_stress/chronic/.

**Table 3.5.** Functional Distribution of the Observed and Predicted MR-1 Proteomes

| Functional Category (no.) | Observed | Predicted | Percentage | Up-regulated | Down-regulated |
|---|---|---|---|---|---|
| Amino acid biosynthesis (1) | 72 | 91 | 79.1% | 4 | 0 |
| Biosynthesis of cofactors, prosthetic groups, and carriers (2) | 95 | 121 | 78.5% | 0 | 1 |
| Cell envelope (3) | 115 | 180 | 63.9% | 3 | 5 |
| Cellular processes (4) | 175 | 260 | 67.3% | 0 | 7 |
| Central intermediary metabolism (5) | 28 | 51 | 54.9% | 3 | 0 |
| DNA metabolism (6) | 84 | 144 | 58.3% | 5 | 0 |
| Energy metabolism (7) | 205 | 308 | 66.6% | 1 | 5 |
| Fatty acid and phospholipid metabolism(8) | 45 | 65 | 69.2% | 0 | 0 |
| Hypothetical proteins (9) | 633 | 2039 | 31.0% | 20 | 20 |
| Mobile and extrachromosomal element functions (10) | 45 | 317 | 14.2% | 14 | 0 |
| Protein fate (11) | 131 | 185 | 70.8% | 2 | 4 |
| Protein synthesis (12) | 128 | 141 | 90.8% | 0 | 1 |
| Purines, pyrimidines, nucleo-sides, and nucleotides (13) | 58 | 62 | 93.5% | 1 | 1 |
| Regulatory functions (14) | 104 | 199 | 52.3% | 0 | 2 |
| Signal transduction (15) | 34 | 61 | 55.7% | 0 | 0 |
| Transcription (16) | 45 | 54 | 83.3% | 0 | 0 |
| Transport and binding  (17) | 126 | 274 | 46.0% | 0 | 5 |
| Unknown function (18) | 190 | 379 | 50.1% | 3 | 2 |
| **Total** | **2313** | **4931** | **46.9%** | | |

Proteins identified at the two-peptide level under the two different growth

conditions were organized in Table 3.5 according to the functional categories assigned by

the J. Craig Venter Institute [formerly The Institute for Genomic Research (see

www.tigr.org; Comprehensive Microbial Resource)].  Using multidimensional HPLC-

MS/MS, we identified more than 75% of the predicted MR-1 proteins for the following

five functional categories:  amino acid biosynthesis; biosynthesis of cofactors, prosthetic

groups, and carriers; protein synthesis; purines, pyrimidines, nucleosides, and

nucleotides; and transcription.  Under 24-h Cr(VI) exposure, more proteins were

identified in the functional categories of amino acid biosynthesis; biosynthesis of

cofactors, prosthetic groups, and carriers; purines, pyrimidines, nucleosides, and

nucleotides; and protein synthesis than were found under the corresponding control

condition (Table 3.5).

Sequence analysis of the MR-1 genome revealed the presence of an integrated

lambda-like phage (LambdaSo; 51,857 bp) and two phylogenetically distinct phages

related to the *E. coli* mu (MuSo1 [34,551 bp] and MuSo2 [35,666 bp]) [69].  The

lambdalike phage genome is also present in MR-1 in a nonintegrated form [69].  There

are 75, 42, and 53 open reading frames (ORFs) annotated as LambdaSo, MuSo1, and

MuSo2 genes, respectively [69].  A previous study focusing on *S. oneidensis* MR-1

demonstrated the induction of a large number of prophage-related genes in response to

UV radiation, particularly those genes from the integrated LambdaSo genome, and the

presence of phage particles in UV-irradiated MR-1 cultures [192].  Based on

transcriptome analysis, the genomic response of MR-1 to ionizing radiation (40 Gy) was

found to be very similar to its response to UV radiation [203].  Similarly, we observed the

strong induction of numerous prophage-related genes in MR-1 cells exposed to chromate

for 24 h (see supplemental Table S1 at compbio.ornl.gov/shewanella_metal_stress/

chronic/supplemental; Table 3.6), suggesting that prolonged Cr(VI) exposure and/or the

accumulation of intracellular chromium may induce the lytic cycle of lysogenic

bacteriophage in MR-1. Overall, 16 (21%), 2 (5%), and 10 (19%) ORFs annotated as

LambdaSo, MuSo1, and MuSo2 genes, respectively, were significantly induced (more

than twofold; $P < 0.05$) in response to prolonged Cr(VI) exposure. This molecular

response was in striking contrast to the differentially expressed genes/proteins

characterizing the cellular response to a 90-min acute 1 mM chromate challenge, during

which a very small subset of predicted prophage genes (i.e., six) displayed a moderate

two- to fourfold induction [147].

Gene products for 14 ORFs with annotations corresponding to mobile and

extrachromosomal element functions were confidently identified as being up-regulated in

response to prolonged Cr(VI) treatment based on both microarray analysis and

multidimensional HPLC-MS/MS (Table 3.6). Ten of these 14 proteins were encoded in

the LambdaSo genome, whereas 1 and 2 gene products were annotated as prophage

MuSo1 and MuSo2 proteins, respectively. The majority of these genes encoded such

prophage structural proteins as minor and major tail proteins, tail assembly components,

and the major head subunit (Table 3.6). Six additional ORFs (SO2941, SO2969,

SO2973, SO2978, SO2985, and SO3006) with prophage LambdaSo-related functions

displayed significant increases (more than twofold; $P < 0.05$) in mRNA expression (see

Supplemental Table S1 at compbio.ornl.gov/shewanella_metal_stress/chronic/). The

corresponding proteins for those genes were detected only under Cr(VI) conditions by

**Table 3.6.** Relative Expression of Up-regulated mRNA and Corresponding Proteins in Response to 24-h Chromate Exposure

| Gene | Gene product (functional category no.[e]) | Transcriptomics [Cr(VI)/ Con ratio][a] | Control | | | | 24-h chromate exposure | | |
| | | | % Coverage[b] | No. of unique peptides/ protein[c] | Avg no. of spectra[d] | % Coverage | No. of unique peptides/ protein | Avg no. of spectra |
|---|---|---|---|---|---|---|---|---|---|
| SO0401 | Alcohol dehydrogenase, zinc containing (7) | 1.2 | 0.0 | 0 | 0 | 45.1 | 9 | 11 |
| SO0644 | Prophage MuSo1 DNA transposition protein (10) | 11.9 | 0.0 | 0 | 0 | 72.4 | 15 | 25 |
| SO0795 | Conserved hypothetical protein (9) | 0.80 | 5.9 | 2 | 1 | 49.7 | 16 | 16.5 |
| SO2654 | Putative transposase (10) | 1.5 | 0.0 | 0 | 0 | 36.9 | 19 | 16 |
| SO2655 | Prophage MuSo2 DNA transposition protein (10) | 3.6 | 16.0 | 3 | 2 | 57.5 | 18 | 57.5 |
| SO2660 | Conserved hypothetical protein (9) | 14.1 | 32.5 | 5 | 6 | 83.5 | 24 | 145 |
| SO2663 | Conserved hypothetical protein (9) | 8.5 | 0.0 | 0 | 0 | 58.7 | 12 | 36 |
| SO2667 | Conserved hypothetical protein (9) | 19.7 | 0.0 | 0 | 0 | 48.0 | 8 | 10 |
| SO2673 | Hypothetical protein (9) | 10.8 | 0.0 | 0 | 0 | 55.4 | 8 | 7.5 |
| SO2685 | Putative prophage MuSo2 major head subunit (10) | 24.5 | 22.7 | 3 | 2 | 51.8 | 12 | 17.5 |
| SO2688 | Hypothetical protein (9) | 24.5 | 0.0 | 0 | 0 | 42.0 | 10 | 6.5 |
| SO2834 | Anaerobic ribonucleoside-triphosphate reductase, NrdD (13) | 1.9 | 16.6 | 8 | 8 | 37.2 | 20 | 30 |
| SO2940 | Prophage LambdaSo host specificity protein J (10) | 5.4 | 3.7 | 2 | 1 | 49.6 | 49 | 107 |
| SO2942 | Hypothetical protein (9) | 42.4 | 0.0 | 0 | 0 | 59.6 | 11 | 19 |
| SO2944 | Hypothetical protein (9) | 44.7 | 0.0 | 0 | 0 | 65.3 | 26 | 96.5 |
| SO2945 | Hypothetical protein (9) | 129.8 | 26.6 | 4 | 6.5 | 67.8 | 15 | 78 |
| SO2946 | Hypothetical protein (9) | 37.2 | 0.0 | 0 | 0 | 58.0 | 10 | 65.5 |
| SO2948 | Prophage LambdaSo tail assembly protein K (10) | 23.7 | 0.0 | 0 | 0 | 51.9 | 8 | 7.5 |
| SO2949 | Prophage LambdaSo minor tail protein L (10) | 5.6 | 0.0 | 0 | 0 | 50.6 | 10 | 23.5 |
| SO2950 | Hypothetical protein (9) | 10.9 | 0.0 | 0 | 0 | 53.4 | 9 | 23.5 |

**Table 3.6.** Continued

| Gene | Gene product (functional category no.[e]) | Transcriptomics [Cr(VI)/ Con ratio][a] | Control | | | 24-h chromate exposure | | |
|---|---|---|---|---|---|---|---|---|
| | | | % Coverage[b] | No. of unique peptides/ protein[c] | Avg no. of spectra[d] | % Coverage | No. of unique peptides/ protein | Avg no. of spectra |
| SO2951 | Hypothetical protein (9) | 48.5 | 16.9 | 5 | 4 | 69.2 | 41 | 172.5 |
| SO2952 | Prophage LambdaSo minor tail protein M (10) | 28.6 | 0.0 | 0 | 0 | 82.0 | 6 | 14.5 |
| SO2953 | Prophage LambdaSo tail length tape measure protein H (10) | 33.4 | 4.2 | 2 | 2 | 58.5 | 51 | 74 |
| SO2955 | Conserved hypothetical protein (9) | 74.9 | 14.0 | 1 | 1.5 | 46.3 | 7 | 20 |
| SO2956 | Prophage LambdaSo major tail protein V (10) | 61.9 | 0.0 | 0 | 0 | 80.9 | 8 | 57.5 |
| SO2963 | Prophage LambdaSo major capsid protein, HK97 family (10) | 168.0 | 36.6 | 12 | 13 | 80.8 | 32 | 183.5 |
| SO2964 | ClpP protease family protein (11) | 57.7 | 0.0 | 0 | 0 | 23.6 | 7 | 11 |
| SO2965 | Prophage LambdaSo portal protein, HK97 family (10) | 106.0 | 14.8 | 2 | 1 | 41.9 | 14 | 16 |
| SO2979 | Hypothetical protein (9) | 11.7 | 0.0 | 0 | 0 | 45.0 | 7 | 8.5 |
| SO2980 | Hypothetical protein (9) | 8.9 | 0.0 | 0 | 0 | 69.1 | 11 | 12.5 |
| SO2988 | Conserved hypothetical protein (9) | 4.2 | 69.2 | 7 | 6.5 | 95.3 | 23 | 84.5 |
| SO2993 | Putative prophage LambdaSo type II DNA modification methyltransferase (10) | 3.3 | 0.0 | 0 | 0 | 67.2 | 19 | 42.5 |
| SO3001 | Hypothetical protein (9) | 8.4 | 0.0 | 0 | 0 | 80.6 | 9 | 28 |
| SO3004 | Putative prophage LambdaSo DNA modification methyltransferase (10) | 2.4 | 0.0 | 0 | 0 | 68.2 | 21 | 51.5 |
| SO3008 | Hypothetical protein (9) | 2.0 | 0.0 | 0 | 0 | 56.2 | 6 | 10.5 |
| SO3013 | Site-specific recombinase, phage integrase family (6) | 1.4 | 0.0 | 0 | 0 | 24.3 | 11 | 10.5 |
| SO3019 | Anthranilate synthase component I, TrpE (1) | 1.3 | 5.9 | 2 | 1 | 45.6 | 19 | 32 |
| SO3020 | Glutamine amidotransferase, TrpG (1) | 2.5 | 0.0 | 0 | 0 | 51.0 | 6 | 8.5 |
| SO3022 | Isomerase, TrpC/F (1) | 1.6 | 13.0 | 5 | 2.5 | 45.2 | 19 | 33.5 |
| SO3061 | DNA topoisomerase III, TopB (6) | 1.3 | 8.0 | 3 | 3 | 37.9 | 18 | 16 |

**Table 3.6.** Continued

| Gene | Gene product (functional category no.[e]) | Transcriptomics [Cr(VI)/ Con ratio][a] | Control | | | 24-h chromate exposure | | |
|---|---|---|---|---|---|---|---|---|
| | | | % Coverage[b] | No. of unique peptides/ protein[c] | Avg no. of spectra[d] | % Coverage | No. of unique peptides/ protein | Avg no. of spectra |
| SO3183 | Perosamine synthetase-related protein (18) | 3.3 | 12.7 | 3 | 2.5 | 46.8 | 10 | 12.5 |
| SO3185 | Polysaccharide biosynthesis protein (3) | 3.4 | 9.5 | 2 | 1 | 55.9 | 12 | 14.5 |
| SO3189 | Polysaccharide biosynthesis protein (3) | 2.2 | 28.2 | 7 | 9 | 70.9 | 18 | 34.5 |
| SO3315 | Conserved hypothetical protein TIGR00048 (9) | 0.80 | 0.0 | 0 | 0 | 36.5 | 10 | 9.5 |
| SO3726 | Sulfate adenylyltransferase, subunit I, CysN (5) | 2.4 | 14.7 | 3 | 4 | 45.9 | 14 | 22.5 |
| SO3727 | Sulfate adenylyltransferase, subunit 2, CysD (5) | 2.2 | 16.6 | 4 | 2.5 | 54.6 | 11 | 23.5 |
| SO3737 | Sulfite reductase (NADPH) hemoprotein beta component, CysI (5) | 2.1 | 3.7 | 2 | 1 | 45.7 | 17 | 27 |
| SO3797 | Peptidase, U32 family (11) | 1.0 | 0.0 | 0 | 0 | 28.1 | 10 | 10.5 |
| SO4265 | Type I restriction-modification system, M subunit, HSdM-2 (6) | 1.1 | 22.9 | 6 | 6.5 | 52.9 | 17 | 20.5 |
| SO4309 | Diaminopimelate decarboxylase, LysA (1) | 0.70 | 11.1 | 3 | 2 | 46.4 | 11 | 16.5 |
| SO4343 | Aminotransferase, class V (18) | 1.7 | 0.0 | 0 | 0 | 59.3 | 14 | 24 |
| SO4686 | NAD-dependent epimerase/dehydratase family protein (3) | 1.1 | 0.0 | 0 | 0 | 50.1 | 10 | 7 |
| SOA0003 | Putative type II restriction endonuclease (6) | 0.70 | 16.7 | 3 | 1.5 | 50.9 | 17 | 23 |
| SOA0004 | Type II DNA modification methyltransferase (6) | 0.90 | 13.0 | 4 | 4 | 54.9 | 24 | 44 |
| SOA0160 | Putative esterase (18) | 0.70 | 30.5 | 6 | 7.5 | 63.4 | 11 | 21 |

[a]Relative gene expression (induction) is presented as the mean ratio of the fluorescence intensity of Cr(VI)-exposed cells to control cells. [b]Total sequence coverage from replicate analyses. [c]Total number of unique peptides identified per protein from replicate analyses. [d]Average number of spectra identified per protein from replicate analyses. [e]From Table 3.5.

HPLC-MS/MS analysis but failed to meet our filtering criteria for determining differential expression (see Chapter 2 and Supplemental Table S2):

SO2941 (~20% sequence coverage), SO2969 (protein not detected), SO2973 (~8.2%), SO2978 (~20%), SO2985 (~15%), and SO3006 (~15%). The genes encoded a putative LambdaSo-associated lysozyme (SO2973; 1,053.6-fold), tail assembly protein I (SO2941; 366.5-fold), a putative holin (SO2969; 49.2-fold), a sitespecific recombinase (SO2978; 23.8- fold), replication protein O (SO2985; 4.4-fold), and a type II DNA modification methyltransferase (SO3006; 3.8-fold). Other up-regulated prophage-related genes had predicted functions in virion morphogenesis (SO2690; 65.6-fold), DNA transposition (SO0644; 11.9-fold; SO2655; 3.6-fold) and circulation (SO2698; 4.1-fold), positive regulation of late transcription (SO2668; 16.7-fold), baseplate (SO2700; 4.3-fold) and tail assembly (SO2699, 10.3-fold; SO2704, 17.2-fold), and assembly of the major head subunit (SO0675; 5.9-fold; SO2685; 24.5-fold), as well as assembly of other structural components (SO2681, 6.2-fold; SO2684, 6.8-fold) (see supplemental Table S1).

Of the differentially expressed proteins determined as having increased abundance under Cr(VI) conditions, 36% corresponded to hypothetical or conserved hypothetical proteins (Table 3.6). Five of these proteins (SO2660, SO2663, SO2667, SO2673, and SO2688) are encoded by genes from the prophage MuSo2 genome, which implied their potential function in prophage activation and synthesis. The majority of the up-regulated hypothetical proteins (i.e., SO2942, SO2944 to -46, SO2950, SO2951, SO2955, SO2979, SO2980, SO2982, SO2988, SO3001, and SO3008) were derived from genes located in the LambdaSo genome, while no potentially MuSo1-related hypothetical

77

or conserved hypothetical proteins were measured as being differentially expressed under our experimental conditions (Table 3.6). The gene and protein expression data strongly suggest that, similar to UV irradiation [192], prolonged exposure to chromate and its derivatives may activate the lytic cycles of some or all three of the MR-1 prophages, leading to prophage-mediated cell lysis. At this point, it is not clear whether the induction of MR-1 prophage-related genes/structural proteins is a response to an extended Cr(VI) exposure per se or to the possible intracellular accumulation of chromium, particularly reduced Cr(III). *S. oneidensis* MR-1 cells exposed to Cr(VI) have been shown to precipitate reduced chromium both extracellularly on the cell surface and as electron-dense globules inside cells [135]. Qiu et al. [192] suggested that prophage activation was the major lethal factor in *S. oneidensis* MR-1 following UV C or UV B irradiation. Our results certainly point to prophage activation as a major contributor to the toxic effects of Cr under conditions of prolonged exposure and reduction. The conditions of Cr(III) toxicity and UV irradiation may be similar mechanistically due to both being DNA damage inducing agents.

The membrane response was characterized by changes in the expression of genes encoding outer membrane structural components and polysaccharide biosynthesis proteins. Induced genes included those encoding putative outer membrane porins (encoded by SO0312 and SO1821), OmpW (encoded by SO1673), and three proteins with functions related to polysaccharide biosynthesis (encoded by SO3158, SO3181, and SO3185). Proteomic analysis indicated increases in the synthesis of SO3185 and SO3189, both annotated as polysaccharide biosynthesis proteins, while four predicted lipoproteins (SO2570, NlpD, SOA0110, and SOA0112) and an OmpA family protein

(SO3969) belonging to the functional category of cell envelope proteins showed

decreased abundance under Cr(VI) conditions (Tables 3.6 and 3.7). Located immediately

upstream of *nlpD* (SO3433) in the MR-1 chromosome is the gene *pcm*, which is

predicted to encode protein-L-isoaspartate *O*-methyltransferase, an enzyme involved in

protein modification and repair. The ORF coding for the RNA polymerase sigma factor

RpoS, which controls the expression of many stationary-phase-induced genes, is

positioned just downstream of *nlpD* and is transcribed in the same direction as *nlpD* and

*pcm*. Proteomic analysis revealed that, in addition to the lipoprotein NlpD, expression of

protein-L-isoaspartate *O*-methyltransferase was down-regulated under conditions of 24-h

Cr(VI) exposure (Table 3.7). This is of interest because, with age or under stress

conditions, proteins are susceptible to various spontaneous or deleterious covalent

modifications such as deamidation, the conversion of asparagines into aspartyls and

isoaspartyls, which can result in loss of protein function. Pcm functions in repairing

damaged proteins by selectively methylating atypical L-isoaspartyl sites and converting

them back to L-aspartyls [204, 205]. The enzyme was shown to enhance the survival of

stationary-phase *E. coli* subjected to a secondary environmental stress [206]. The

physiological significance of decreased synthesis of *S. oneidensis* Pcm under the Cr(VI)

conditions used here is unclear.

The vast majority of MR-1 proteins down-regulated under our experimental

conditions are annotated as hypotheticals (Table 3.5). Besides poorly characterized

proteins, other down-regulated proteins belonged to the functional categories of cellular

process proteins, transport and binding proteins, cell envelope proteins, and energy

metabolism proteins. The relative abundance levels of seven proteins were found to be

**Table 3.7.** Relative Expression of Down-regulated Proteins and Corresponding mRNA Levels after 24-h Exposure to Chromate

| | | | Control | | | 24-h chromate exposure | | |
|---|---|---|---|---|---|---|---|---|
| Gene | Gene product (functional category no.[f]) | Transcriptomics [Cr(VI)/Con ratio][a] | % Coverage[b] | No. of unique peptides/ protein[c] | Avg no. of spectra[d] | % Coverage | No. of unique peptides/ protein | Avg no. of spectra |
| SO0433 | Regulator of sigma D, Rsd (14) | 0.70 | 67.7 | 12 | 19 | 36.0 | 5 | 2.5 |
| SO0576 | PhoH family protein (18) | 0.80 | 58.6 | 19 | 22.5 | 0.0 | 0 | 0 |
| SO0902 | NADH:ubiquinone oxidoreductase, Na translocating, alpha subunit, NqrA-1 (7) | 1.0 | 70.6 | 24 | 24 | 16.7 | 6 | 6 |
| SO1144 | Methyl-accepting chemotaxis protein (4) | 0.60 | 86.5 | 33 | 71.5 | 31.8 | 9 | 12.5 |
| SO1425 | Hypothetical protein (9) | 1.0 | 44.8 | 17 | 13 | 0.0 | 0 | 0 |
| SO1518 | Conserved hypothetical protein (9) | 0.60 | 85.7 | 14 | 39 | 39.7 | 7 | 8 |
| SO1689 | Cation transport ATPase, E1-E2 family (4) | 0.40 | 30.0 | 13 | 10 | 0.0 | 0 | 0 |
| SO1700 | Hypothetical protein (9) | 0.90 | 61.2 | 21 | 23.5 | 0.0 | 0 | 0 |
| SO2062 | Conserved hypothetical protein (9) | 0.50 | 63.8 | 12 | 29 | 44.0 | 3 | 5.5 |
| SO2247 | Hypothetical protein (9) | 0.40 | 38.4 | 12 | 12 | 0.0 | 0 | 0 |
| SO2304 | Alanine dehydrogenase, authentic point mutation, Ald (7) | 0.40 | 67.7 | 20 | 62.5 | 23.5 | 4 | 10 |
| SO2469 | Conserved hypothetical protein (9) | 0.60 | 67.9 | 61 | 90.5 | 8.8 | 5 | 2.5 |
| SO2570 | Putative lipoprotein (3) | 1.0 | 58.3 | 34 | 54.5 | 26.5 | 10 | 6 |
| SO2682 | Hypothetical protein (9) | 2.5 | 54.5 | 3 | 7 | 0.0 | 0 | 0 |
| SO2766 | Conserved hypothetical protein (9) | 1.0 | 63.6 | 27 | 25.5 | 34.3 | 7 | 6 |
| SO2882 | Conserved hypothetical protein (9) | 0.80 | 75.2 | 51 | 91.5 | 43.8 | 21 | 23 |
| SO2893 | Conserved hypothetical protein (9) | 1.9 | 72.3 | 16 | 57 | 25.1 | 3 | 4.5 |
| SO2991 | Conserved hypothetical protein (9) | 0.70 | 55.0 | 5 | 10 | 0.0 | 0 | 0 |
| SO3030 | Siderophore biosynthesis protein, AlcA(17) | 1.6 | 60.8 | 27 | 79.5 | 36.7 | 10 | 9.5 |
| SO3069 | Conserved hypothetical protein (9) | 1.6 | 31.2 | 30 | 44.5 | 10.8 | 7 | 7 |
| SO3103 | AcrB/AcrD/AcrF family protein (4) | 1.1 | 34.1 | 23 | 22.5 | 13.7 | 9 | 6.5 |
| SO3145 | Electron transfer flavoprotein, beta subunit, EtfB (7) | 1.2 | 96.4 | 26 | 131.5 | 64.3 | 14 | 45.5 |

**Table 3.7.** Continued

| Gene | Gene product (functional category no.[f]) | Transcriptomics [Cr(VI)/Con ratio][a] | Control | | | 24-h chromate exposure | | |
|---|---|---|---|---|---|---|---|---|
| | | | % Coverage[b] | No. of unique peptides/ protein[c] | Avg no. of spectra[d] | % Coverage | No. of unique peptides/ protein | Avg no. of spectra |
| SO3207 | Chemotaxis protein (4) | 1.5 | 54.9 | 44 | 127 | 46.7 | 24 | 32 |
| SO3235 | Flagellar hook-associated protein FliD (4) | 0.80 | 71.3 | 30 | 39 | 34.6 | 7 | 5.5 |
| SO3247 | Flagellar hook protein FlgE (4) | 0.90 | 66.0 | 14 | 36.5 | 31.1 | 6 | 10 |
| SO3314 | Putative fimbrial biogenesis and twitching motility protein(4) | 1.0 | 53.1 | 10 | 10.5 | 0.0 | 0 | 0 |
| SO3343 | Conserved hypothetical protein (9) | 0.80 | 81.2 | 20 | 97.5 | 61.3 | 11 | 28 |
| SO3407 | Conserved hypothetical protein (9) | 0.80 | 28.6 | 16 | 16 | 0.0 | 0 | 0 |
| SO3422 | Ribosomal subunit interface protein, YfiA-2 (12) | 1.0 | 89.8 | 8 | 28 | 17.8 | 2 | 1 |
| SO3433 | Lipoprotein NlpD (3) | 1.2 | 46.3 | 10 | 9.5 | 0.0 | 0 | 0 |
| SO3434 | Protein-L-isoaspartate O-methyltransferase, Pcm (11) | 0.80 | 64.0 | 9 | 7.5 | 0.0 | 0 | 0 |
| SO3442 | MazG family protein (18) | 1.6 | 66.3 | 13 | 26 | 18.6 | 3 | 3.5 |
| SO3468 | Riboflavin synthase, alpha subunit, RibE-2 (2) | 1.1 | 51.8 | 10 | 11 | 0.0 | 0 | 0 |
| SO3483 | HlyD family secretion protein (17) | 1.0 | 68.6 | 19 | 18.5 | 25.5 | 5 | 3 |
| SO3516 | Transcriptional regulator, LacI family (14) | 1.2 | 49.1 | 13 | 14 | 15.7 | 3 | 3 |
| SO3539 | Peptidase, M28D family (11) | 0.50 | 72.7 | 26 | 40 | 30.7 | 9 | 5 |
| SO3550 | Hypothetical protein (9) | 1.0 | 57.4 | 9 | 7 | 0.0 | 0 | 0 |
| SO3560 | Peptidase, M16 family (11) | 0.30 | 65.0 | 45 | 50 | 33.7 | 20 | 19 |
| SO3565 | 2,3-Cyclic-nucleotide 2-phosphodiesterase, CpdB (13) | 0.70 | 79.7 | 53 | 74.5 | 28.8 | 12 | 8.5 |
| SO3597 | Conserved hypothetical protein (9) | 1.0 | 78.7 | 13 | 48 | 0.0 | 0 | 0 |
| SO3683 | Coniferyl aldehyde dehydrogenase (7) | 0.50 | 55.1 | 24 | 33 | 21.1 | 6 | 6.5 |
| SO3720 | Conserved hypothetical protein (9) | 1.6 | 80.1 | 12 | 21.5 | 29.8 | 4 | 2.5 |
| SO3800 | Serine protease, subtilase family (11) | 0.50 | 19.7 | 11 | 7 | 0.0 | 0 | 0 |

**Table 3.7.** Continued

| Gene | Gene product (functional category no.[f]) | Transcriptomics [Cr(VI)/Con ratio][a] | Control | | | 24-h chromate exposure | | |
|------|------|------|------|------|------|------|------|------|
| | | | % Coverage[b] | No. of unique peptides/ protein[c] | Avg no. of spectra[d] | % Coverage | No. of unique peptides/ protein | Avg no. of spectra |
| SO3936 | Sodium-type flagellar protein MotX (17) | 0.90 | 47.6 | 8 | 6 | 0.0 | 0 | 0 |
| SO3969 | OmpA family protein (3) | 1.5 | 61.8 | 14 | 19.5 | 20.6 | 2 | 1.5 |
| SO3980 | Cytochrome c552 nitrite reductase (7) | 1.2 | 49.5 | 27 | 49.5 | 20.3 | 7 | 6.5 |
| SO4319 | HlyD family secretion protein (17) | 1.3 | 46.7 | 20 | 23 | 25.0 | 7 | 6 |
| SO4329 | Conserved hypothetical protein (9) | 1.4 | 94.0 | 15 | 47 | 66.7 | 6 | 13 |
| SO4403 | Hypothetical protein (9) | 0.50 | 72.5 | 16 | 18 | 22.1 | 2 | 2 |
| SO4505 | Conserved hypothetical protein (9) | 0.50 | 80.7 | 8 | 11.5 | 0.0 | 0 | 0 |
| SO4523 | Iron-regulated outer membrane virulence protein IrgA (17) | 0.60 | 78.9 | 55 | 169 | 48.3 | 23 | 36.5 |
| SOA0110 | Putative lipoprotein (3) | 0.50 | 43.3 | 10 (46) | 147 | 16.4 | 1 (16) | 32 |
| SOA0112 | Putative lipoprotein (3) | NA[e] | 47.7 | 0 (52) | 224.5 | 20.4 | 0 (17) | 31.5 |

[a]Relative gene expression (induction) is presented as the mean ratio of the fluorescence intensity of Cr(VI)-exposed cells experimental) to control (nonexposed) cells. [b]Total sequence coverage from replicate analyses. [c]Total number of unique peptides identified per protein from replicate analyses. [d]Average number of spectra identified per protein from replicate analyses. [e]NA, not applicable; gene was not represented on the microarray. [f]From Table 1.

decreased under 24-h Cr(VI) exposure compared to the control conditions: two

chemotaxis proteins (SO1144 and SO3207), a cation transport ATPase (SO1689), an

AcrB/AcrD/AcrF family protein (SO3103), and three proteins involved in motility (FliD,

FlgE, and SO3314) (Table 3.7). Down-regulation of proteins involved in motility and

chemotaxis was consistent with confocal laser scanning microscopy observations, which

indicated a prevalence of nonmotile cells under prolonged Cr(VI) exposure (data not

shown).

## **Conclusions**

The proteome datasets presented here provide an in-depth analysis of the global

response of *S. oneidensis* MR-1 to both acute toxic shock and chronic exposure to

chromate. Using 2D LC-MS/MS for the proteome measurements, the proteome datasets

provided protein expression information on previously unknown proteins and revealed

possible post-transcriptional regulation that is not possible to predict by microarray

analysis. Under the acute shock conditions, the proteins most commonly up-regulated

were involved in iron uptake and assimilation as well as sulfate transport. Many

hypothetical proteins were also found to be up-regulated in response to acute chromate

exposure. In addition, a putative operon containing three genes (*so3585-so3587*) was

detected up-regulated at the transcript and protein levels primarily under chromate

exposure.

After growth in the presence of chromium for 24-h, the differentially expressed

proteins identified from *S. oneidensis* were unique in comparison to the acute shock

proteome dataset. Proteins corresponding to structural components and hypothetical

proteins encoded within an integrated lambda-like phage (LamdaSo) were identified not

83

only highly up-regulated at the time of proteome sampling in the chromate-exposed culture, but was primarily detected in this sample as well. A similar response was found for this bacterium after exposure to UV radiation [192]. The activation of the lysogenic phage may be the lethal cause that results from prolonged exposure to chromate. Since the phage is activated after transformation of Cr(VI) to presumably Cr(III), understanding cellular response in the presence of Cr(III) may not be possible without first inactivating the lysogenic phage. The use of global whole-cell proteomics allowed the detection and characterization of these novel responses to chromate by *S. oneidensis*.

Since the purpose is to gain global understanding of the Cr(VI) response of *S. oneidensis* for the purposes of bioremediation, we chose to explore environmentally relevant scenarios of Cr(VI) exposure. Therefore, the two proteome studies discussed here provided insight into both initial and long-term Cr(VI) exposure. We found that after prolonged exposure, a high level of stress was demonstrated that may be detrimental to a long-term bioremediation strategy; however, without this work, the stress response would have remained unknown. Due to the large differences between the two proteomes, this resulted in the exploration of the Cr(VI) dosage response, another relevant scenario.

# Chapter 4

## Dosage-Dependent Proteome Response of *Shewanella oneidensis* MR-1 to Acute

## Chromate Challenge

*The work presented below has been published as the following journal article and adapted from the text*

Melissa R. Thompson, Nathan C. VerBerkmoes, Karuna Chourey, Manesh Shah, Dorothea K. Thompson, and Robert L. Hettich. Dosage-Dependent Proteome Response of *Shewanella oneidensis* MR-1 to Acute Chromate Challenge. *Journal of Proteome Research*, 2007; 6, 1745-1757. *Sample preparation, LCQ and LTQ measurements, and data analysis were performed by Melissa R. Thompson. Supplemental material can be found at http://compbio.ornl.gov/shewanella_metal_stress/dosage.*

## Introduction

Technological advances in the area of "shotgun" proteomics have substantially improved the quantity and quality of data that can be derived from diverse biological systems, including molecular descriptions of the growth states and environmental perturbation responses of various bacteria. Areas that have seen improvements include liquid chromatographic separation of peptides, new and improved mass spectrometers, and advanced computational approaches to search the resulting raw data files. In 2001, Washburn *et al* [33] first introduced multidimensional protein identification technology (MudPIT), in which a multiphasic column was used to separate peptides online for direct injection into a nanoelectrospray mass spectrometer. This design was enhanced by McDonald *et al* [166] in 2002 to incorporate a biphasic column that is connected upstream from the reverse-phase analytical column. These improvements in online peptide separation have greatly improved sample handling and have reduced the

complexity of ions injected into the mass spectrometer, thereby allowing more ions to be analyzed by MS/MS. Over the past 8 years, the quadrupole ion trap (QIT) has been the prevalent choice of instrumentation for this experimental approach [207-212]. However, advancements in instrumentation (i.e., the new linear trapping quadrupole LTQ [30] and LTQ-Orbitrap [23]) that can acquire full MS and MS/MS scans on greatly reduced time scales have become just as critical as improvements made in peptide separation. All of these improvements have led to a greater depth of proteome coverage, yielding confident identifications of greater than 2000 proteins on a routine basis for many microbes [63, 65, 147, 158].

These technological advances have enabled the elucidation of microbial proteomes at a remarkably deep level, now permitting the extensive examination of how a microbe adjusts its molecular machinery as a function of diverse growth conditions. The focus of this work was to exploit this technology to characterize for the first time the dosage-dependent changes in the proteome of the metal-reducing bacterium *Shewanella oneidensis* MR-1 in response to chromate insult. Hexavalent chromium [Cr(VI)], in the form of chromate ($CrO_4^{2-}$) or dichromate ($Cr_2O_7^{2-}$), is a widespread environmental contaminant due to its prevalent use in industrial and defense applications [121]. Microbially-catalyzed transformation of soluble Cr(VI) to less soluble Cr(III) hydroxides has been proposed as a potentially economical and environmentally friendly remediation strategy for Cr(VI)-contaminated subsurface environments [213]. Cr(III) hydroxides, for example, exhibit substantially reduced mobility, bioavailability, and toxicity. Species of the γ-proteobacterial genus *Shewanella* have been shown to mediate both direct

enzymatic [86, 133, 134] and indirect chemical reduction [131, 136, 137] of Cr(VI) to

Cr(III), thus suggesting their potential utility for *in situ* bioremediation.

Under aerobic growth conditions, *S. oneidensis* preferentially uses oxygen as its

terminal electron acceptor.  Chromate reduction by this organism has been reported under

both aerobic and anaerobic conditions [86, 133, 135, 146].  Although *S. oneidensis* can

generate energy for growth via respiratory-linked reduction of certain metals such as

Fe(III) and Mn(IV) [79], there have been no reports to date of this organism's growth on

Cr(VI) as a sole terminal electron acceptor under anaerobic respiratory conditions [132].

Current evidence suggests that Cr(VI) reduction mechanisms are likely associated with

bacterial electron transport systems [86, 214] and that Cr(VI) reduction is dependent on

the physiological state of the culture [133], but the details describing these mechanisms

and the molecular components themselves are not understood in *S. oneidensis*.  Thus, the

prediction and assessment of bioremediation performance is a difficult task, being

compounded by the lack of fundamental knowledge on the molecular basis and regulation

underlying bacterial metal reduction and cellular responses to metal toxicity.

This research builds upon and extends a previous study probing the temporal

transcriptomic and proteomic changes in *S. oneidensis* MR-1 in response to a single

exposure concentration of chromate [147] by investigating the Cr(VI) dosage-dependent

proteome response.  The lethal dosage for *S. oneidensis* was found to be around 2 mM;

however, at dosages of 0.5 and 1 mM, a reduction in growth of 50 and 66%, respectively,

was observed [147].  Given the likelihood that microorganisms are exposed to a gradient

of chromate levels in contaminated local microenvironments, it is important to

understand changes in the molecular response to metal toxicity that are dependent on

concentration and, thus, gain insight into the level of resistance exhibited by a bacterium. The objective of this study was to identify proteins involved in the general response of the bacterium to chromate insult, in particular those that demonstrate a dosage-dependent alteration in abundance level.  The majority of proteins synthesized by this bacterium were expected to remain relatively constant, as this is the vital machinery that enables cell growth and survival maintenance.  However, a subset of proteins that play important roles in the stress response to chromate and ultimately for cell survival in toxic metal environments should exhibit measurable changes in abundance as a function of growth conditions, and thus would provide key information about how this microbe copes with acute exposures to Cr(VI).

A component of this study was to evaluate the differences between the datasets obtained by the more conventional quadrupole ion trap technology as compared to the newer linear trapping quadrupole instrumentation.  These technological advances have the potential to increase the correlation of transcriptome and proteome measurements due to improved dynamic range of the LTQ, thereby increasing proteome coverage and improving confidence in the interpretation of results derived from global high-throughput techniques.

## **Materials and Methods**

*Preparation of Proteomes for HPLC-MS/MS Analysis*

Four independent cultures of wild-type *S. oneidensis* MR-1 were grown in Luria-Bertani (LB) medium (pH 7.2) at 30 °C under aerobic conditions.  Aerobic conditions were maintained in batch cultures by growing both the control and treatment samples in the same-size Erlenmeyer flasks (with sufficient headspace) and continuously aerating

each culture by vigorous shaking at 200 rpm. Batch cultures were grown to mid-

exponential phase ($OD_{600}$, 0.5), followed by the addition of potassium chromate ($K_2CrO_4$)

to three of the four cultures at a final concentration of 0.3, 0.5, or 1.0 mM, respectively.

The culture receiving no chromate served as the control and was grown in parallel with

the treated cultures. Thirty minutes after chromate exposure, cells were harvested and

lysed by sonication, as described in Chapter 2. On the basis of initial growth response

studies of *S. oneidensis* [147], the decision for sampling the resulting proteome during the

initial transitional period (i.e., 30 min post-shock) from logarithmic growth toward

adaptation to chromate would reveal key regulatory and structural proteins that respond

to the initial signs of chromate stress (see Supplemental Figure S1, Supporting

Information). Each of the resulting four proteome samples (control and 0.3, 0.5, or 1.0

mM chromate) were separated into two fractions (crude/soluble and membrane) by high-

speed centrifugation [147] and quantitated using BCA analysis (Pierce, Rockford, IL).

Approximately 2 mg of each proteome fraction was then digested with trypsin, desalted

and stored at –80 °C until HPLC-MS/MS analysis.

*LC-MS/MS Analysis*

Both the soluble and membrane-associated proteome fractions of the chromate-

treated and control samples were analyzed in duplicate via two-dimensional LC-MS/MS

on both an LCQ Deca XP three-dimensional ion trap mass spectrometer (Thermo

Electron) and an LTQ linear trapping quadrupole mass spectrometer (Thermo Electron).

Both instruments were coupled to identical Ultimate HPLC pumps (LC Packings; a

division of Dionex, San Francisco, CA), each of which had an initial flow rate of ~100

µL/min that was split precolumn to obtain a flow rate of ~300 nL/min at the nanospray

tip.  See Chapter 2 for details of chromatographic separation of peptides for both

instrument platforms.  The multiphasic column system was positioned in front of the

LCQ or LTQ on a nanospray source (ThermoFinnigan).  All samples were analyzed via a

24-h 12-step 2-D analysis consisting of increasing concentration (0–500 mM) salt pulses

of ammonium acetate followed by 2-h reverse phase gradients (see Chapter 2 for details).

During the entire chromatographic process, the LTQ was operated in a data-dependent

MS/MS mode detailed in Chapter 2.

*Proteome Bioinformatics*

The *S. oneidensis* MR-1 protein database used in all MS/MS spectra searches

consisted of 4798 open reading frames that were downloaded from the J. Craig Venter

Institute (formerly TIGR) and concatenated with a list of common contaminants (trypsin,

keratin, etc.); the entire database can be downloaded from compbio.ornl.gov/shewanella_

metal_stress/dosage/databases/.  All MS/MS spectra were searched with the Sequest

Algorithm [39] as detailed in Chapter 2.  The Xcorr values used here have been tested in

a rigorous manner in the laboratory and typically give a maximum false positive rate of

1-2% for both bacterial isolates [184] and simple microbial communities [65].  All of the

resulting DTASelect files from both the LCQ and LTQ datasets are available on the

analysis page under the corresponding dataset and are filtered at 1 and 2 peptides per

protein.  Spectra identified for each peptide can be viewed by clicking on the spectral

number (first column, labeled by filename) in the html version of the DTASelect files.

The algorithm Contrast [171] was used to compare DTASelect files and all of the

resulting files are also found on the page.  Although there are a number of large-scale

proteomics studies describing *S. oneidensis* MR-1 under various growth conditions, the

availability and accessibility of the entire datasets to the general scientific community is limited. This precludes direct comparisons of the large dataset across different laboratories. To assess the false identification rate of both the LTQ and LCQ, a database consisting of the protein sequences in reverse as well as the forward protein sequences was created. The 1 mM Cr(VI) dataset from each instrument was searched against this database to determine the false identification rate of the filtering levels used in this study and the results were calculated according to the equation presented in ref [185]. Results in the form of DTASelect files are found at compbio.ornl.gov/shewanella_ metal_stress/dosage/.

Label free quantitation was employed to determine those proteins that were differentially expressed under the Cr(VI) dosage conditions with respect to the control sample. The criteria for label free quantitation are measurable differences in percent sequence coverage, peptide count, and spectral count, where two of the three criteria must be met for a protein to be identified as differentially expressed (see Chapter 2 for details). The use of more stringent criteria for the LTQ is due to the increase in the number of spectra acquired and identified. The increase in the amount of spectra acquired results in a plateauing effect in determination of differential expression for many moderately abundant proteins as discussed below.

**Results and Discussion**

*Global Proteome Analysis*

A total of 2406 out of 4931 predicted proteins from the annotation provided by TIGR were confidently identified with at least 2 peptides per protein using both the LCQ and the LTQ. This constitutes 49% of the predicted *S. oneidensis* MR-1 proteome and

represents the largest proteome characterization of this bacterium to date from this laboratory [147, 158]. The total number of proteins identified under each treatment condition for the two instruments is provided in Table 4.1. With the LTQ mass spectrometer, confident identification of 1883 proteins in the control sample was achieved, with a comparable amount detected in each of the dosage samples. By contrast, the LCQ instrument yielded 809 proteins for the control, with comparable amounts in each of the dosage samples.

The total number of proteins identified was organized into a list with the percent sequence coverage, functional category, isoelectric point (pI), and molecular weight information included (Supplemental Table S1, Supporting Information). With respect to these factors, there were no major biases between the observed proteome in this study and the predicted proteome. Proteins detected under the three chromate dosage conditions and the control condition are organized according to percent sequence coverage (the total percent of the protein sequence identified), peptide count (the number of peptides identified for a given protein), and spectral count (the number of spectra identified for a given protein) and can be accessed in Supplemental Table S2 (see Supporting Information) for both the LCQ and LTQ datasets.

Biological variability with respect to independently grown cultures of *S. oneidensis* was assessed to determine if perturbation of the environment reveals a culture-dependent proteome. Independently grown cultures that were sampled previously [147] for proteomics analysis were compared to the current datasets, and gave very similar results. These previous samples included a wild-type control culture grown in parallel with a 1 mM chromate-treated culture that was harvested 45 min after addition of the

92

**Table 4.1.** Global Proteome Analysis of Chromate Dose on *S. oneidensis*.

| Condition | Instrument[a] | Proteins Identified 1 pep[b] | Proteins Identified 2 pep[c] | Average Sequence Coverage[d] | Reproducibility |
|---|---|---|---|---|---|
| Control | LCQ | 1263 | 809 | 26.0% | 78.6% |
| 0.3 mM K₂CrO₄ | LCQ | 1267 | 835 | 30.0% | 72.8% |
| 0.5 mM K₂CrO₄ | LCQ | 1185 | 716 | 31.1% | 81.8% |
| 1.0 mM K₂CrO₄ | LCQ | 1307 | 879 | 29.1% | 74.5% |
| Control | LTQ | 2574 | 1883 | 35.6% | 79.4% |
| 0.3 mM K₂CrO₄ | LTQ | 2423 | 1752 | 38.9% | 70.5% |
| 0.5 mM K₂CrO₄ | LTQ | 2425 | 1807 | 37.2% | 81.3% |
| 1.0 mM K₂CrO₄ | LTQ | 2624 | 1953 | 37.5% | 79.8% |
| **Total Non-Redundant** | | 3239 | 2406 | | |

[a]LCQ:  Thermo Electron NanoES-quadrupole ion trap.  LTQ:  Thermo Electron NanoES-linear trapping quadrupole.  [b]Non-redundant proteins identified with at least 1 peptide per protein.  [c]Non-redundant proteins identified with at least 2 peptides per protein.  [d]Average sequence coverage with at least 2 peptides per protein.

metal. For these two datasets, the reproducibility at the protein level is 79% for the 30 min 1 mM chromate dosage versus the 45 min 1 mM acute shock sample and for the control samples is 78%. Therefore, even though there were differences in the chromate exposure times for these two separate experiments, there does not seem to be much measurable biological variation, even when chromate exposure is 15 min longer in the case of Brown *et al* [147].

*Instrumentation Effects on Proteome Characterization*

The LCQ and LTQ datasets were compared to determine differences in informational content. Under the control conditions (no added chromate), both instruments together identified a common list of 803 proteins at the two-peptide level. A total of 5 proteins were uniquely identified using the LCQ, whereas approximately 1100 proteins were only identified using the LTQ. The 5 unique LCQ proteins were identified with only 2 peptides per protein. For the 0.3 and 0.5 mM chromate-treated cells, 876 and 777 proteins identified were common between the two instruments, respectively; a total of 5 proteins were unique to the LCQ under these dosages, whereas almost 1000 proteins were unique to the LTQ. The unique LCQ proteins were investigated further and once again were found to consist of only 2 or 3 peptides identified for a given protein. A different trend was observed for the 1 mM chromate-treated samples compared to the other three conditions. All proteins identified using the LCQ were also identified using the LTQ mass spectrometer. Overall, 99% of the proteins identified using the LCQ were also identified when using the LTQ, indicating that the performance of these two instruments is similar; however, the LTQ offered enhanced performance in terms of dynamic range, as discussed below.

Because of its slower scanning speed, the LCQ mass spectrometer is not able to trigger and conduct MS/MS on as many ions as the LTQ during a chromatographic run, thus resulting in a lower number of total protein identifications (for the LCQ). The capability of the LTQ [30] instrument for not only higher duty cycle experiments but also a greater trapping efficiency of ions relative to the LCQ is already making LTQ technology the high-throughput instrumentation of choice for shotgun proteomics measurements. This improved ion trapping efficiency is evident by the triggering of MS/MS acquisitions on low abundance ions that would be missed by the LCQ measurements (see Figure 4.1). Due to the increased dynamic range offered by the LTQ linear trapping quadrupole, only the LTQ-derived datasets were analyzed in greater detail in the following discussion sections.

The increased scanning speed of the LTQ instrument results in a concomitant increase in the number of false positive identifications of peptides [32, 215]. Generation of more spectra with the LTQ leads to more possible matches to spectra that are considered "noise" and do not contain "true" peptide spectra. The 1 mM Cr(VI) dataset was used to test the false identification rate of both the LCQ and LTQ. These datasets were searched against a protein database containing the forward sequence of the protein database (see above) with the reverse sequences of the database concatenated to the end. The false identification rates for the duplicate analyses were averaged to give the following results for each instrumentation platform (Table 4.2). For the LCQ, a protein false identification rate of 10.5% was identified for the 1 peptide dataset and 0.1% for a 2 peptide requirement. For the LTQ, the protein false identification rate at the 1 peptide level was 33.3% and at the 2 peptide level 4.8%. As expected, the false identification
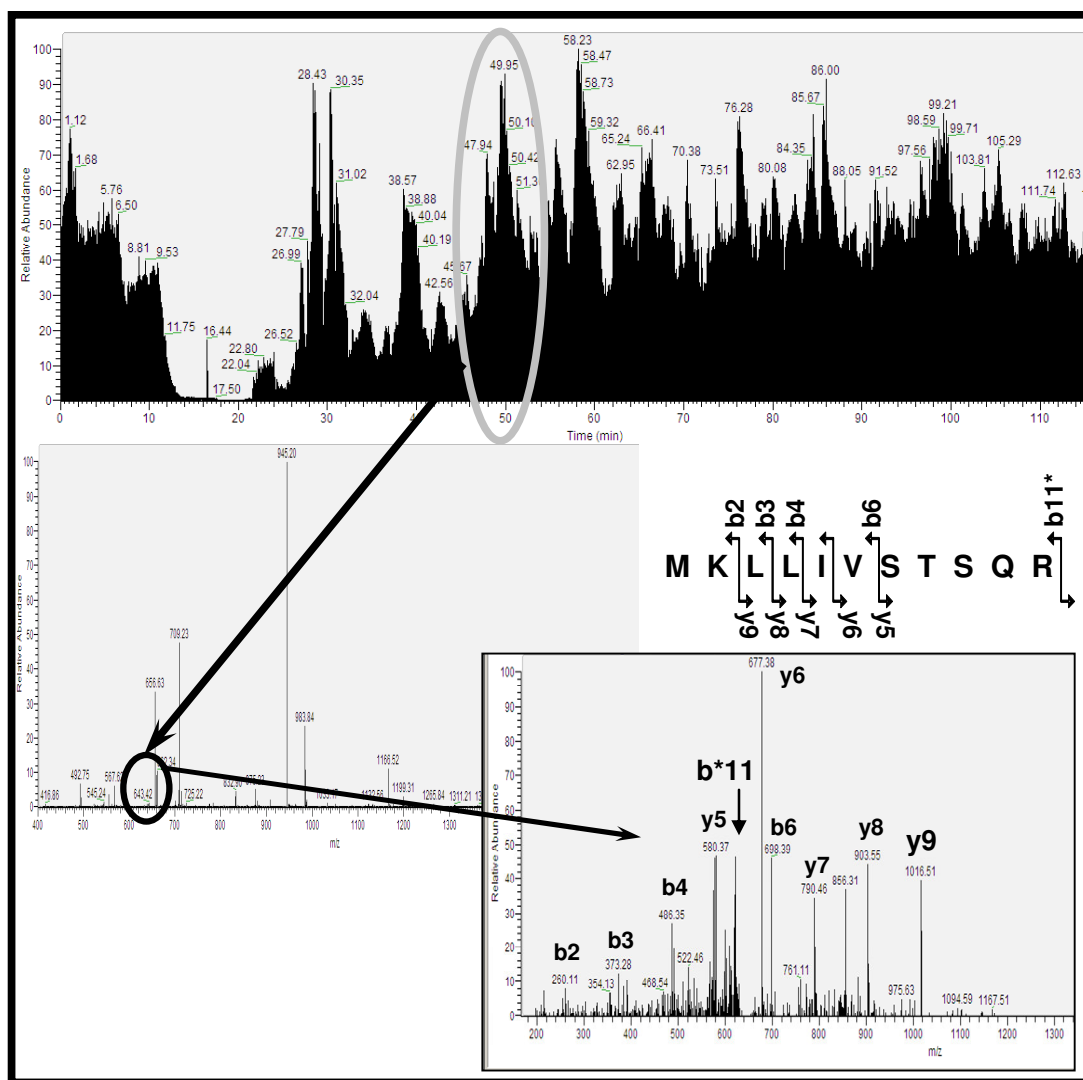
**Figure 4.1.** MS/MS from a peptide of SO3585 identified in the 0.3 mM dosage.

The base peak chromatogram (top figure) from the 0.3 mM chromate sample illustrates a

full LTQ-MS scan that contains a peak of very low abundance at m/z 638.8 (middle

figure). This peak was subsequently isolated and fragmented, giving the MS/MS

spectrum (bottom figure), which contains the sequence of a peptide from the SO3585

protein that was identified as up-regulated at the 0.5 mM and 1 mM chromate levels.

**Table 4.2.** False Peptide and Protein Identifications with a LCQ and a LTQ

| Instrument [a] | Requirement [b] | Peptide Identifications | | | Protein Identifications | | |
|---|---|---|---|---|---|---|---|
| | | False Unique [c] | Total Unique [d] | Rate | False Ids [e] | Total Ids [f] | Rate |
| LCQ Run 1 | 1 peptide | 55 | 6041 | 1.8% | 58 | 1157 | 10.0% |
| | 2 peptide | 0 | 5666 | 0.0% | 0 | 736 | 0.0% |
| LCQ Run 2 | 1 peptide | 65 | 6912 | 1.8% | 73 | 1311 | 11.1% |
| | 2 peptide | 2 | 6492 | 0.0% | 1 | 835 | 0.2% |
| LTQ Run 1 | 1 peptide | 428 | 22014 | 3.8% | 434 | 2735 | 31.7% |
| | 2 peptide | 68 | 21153 | 0.6% | 34 | 1808 | 3.8% |
| LTQ Run 2 | 1 peptide | 527 | 22476 | 4.6% | 498 | 2854 | 34.8% |
| | 2 peptide | 108 | 21531 | 1.0% | 53 | 1844 | 5.8% |

[a]1 mM Cr(VI) Dataset on both instrumentation plaforms. [b]Requirement is whether protein identification is based on at least 1 or 2 peptides for a given protein, respectively. [c]False Unique refers to the number of peptides identified using protein sequences from the reversed protein database. [d]Total Unique is the total number of unique peptides identified using the forward and reverse protein databases. [e]False IDs are the proteins identified from the reverse database. [f]Total IDs are the total number of proteins identified from the forward and reverse databases.

LTQ is higher, especially at the 1 peptide level; however, with the 2 peptide requirement, rate for the the rate is low enough that ~95% of the protein identifications are true. Therefore, using more stringent filtering criteria is necessary in order to minimize the false positive identifications for LTQ datasets.

In order to further understand the contribution of peptides to protein identification, the distribution of the number of peptides identified for a given protein was calculated using the Control dataset as an example. A protein identification is only made if at least two peptides are identified from one of the two MS analyses. The protein identifications made by two replicate MS experiments comprising two peptides were 260 (~14%) of the total proteins identified for the Control condition. A total of 75 (~4%) proteins were identified with three peptides, while 103 (~5%) proteins were found with four peptides. A majority of the protein identifications comprise at least five peptide identifications with 1443 (~77%) proteins. Therefore, the false identification rate is maintained at a reasonable level with less than 30% of the protein identifications arising from proteins identified with four or fewer peptides.

Semiquantitation was investigated on both the LCQ and LTQ datasets using the criteria described in Chapter 2. Results generated using both instrumentation platforms are provided in Supplemental Table S3 (see Supporting Information). Although there is reasonable agreement between the LCQ and LTQ datasets for the three chromate dosage samples, the different depth of protein identification and the distinct semiquantification criteria for each MS platform made it difficult to directly compare the results. This is especially noticeable for the lower abundance proteins; they are likely to be missed in the LCQ analysis but found in the LTQ measurements. Thus, whereas the LCQ

measurements often reveal large semiquantitative differences, such as proteins completely absent in the control sample but present with a few peptides detected in a dosage sample, the deeper measurements possible with the LTQ often reveal the presence of the protein in both control and sample, with some differences in abundance. At the other end of the spectrum, the most abundant proteins also are somewhat difficult to quantify with the LTQ, because at least one of the semiquantitative metrics (percent sequence coverage) "saturates", precluding determination of differences between samples. Thus, these factors must be taken into account in the determination of semiquantification criteria for LTQ measurements. Because of the increased depth of proteome coverage achieved by the LTQ measurements, all semiquantification determinations and discussions were focused on these datasets.

*LTQ-Based Global Proteome Characterization*

The proteins identified using the LTQ were organized according to the 18 functional category assignments given in TIGR's Comprehensive Microbial Resource (www.tigr.org) and were grouped according to their annotated subcellular roles (Figure 4.2). The functional categories of amino acid biosynthesis; biosynthesis of cofactors, prosthetic groups, and carriers; cell envelope; cellular processes; DNA metabolism; energy metabolism; fatty acid and phospholipid metabolism; and protein fate were all represented with at least 50% of their protein members identified for the control and each Cr(VI) dosage sample. Furthermore, in the functional categories of protein synthesis, purines/pyrimidines/nucleosides/nucleotides, and transcription, identification of 80% of the members was achieved for each condition.
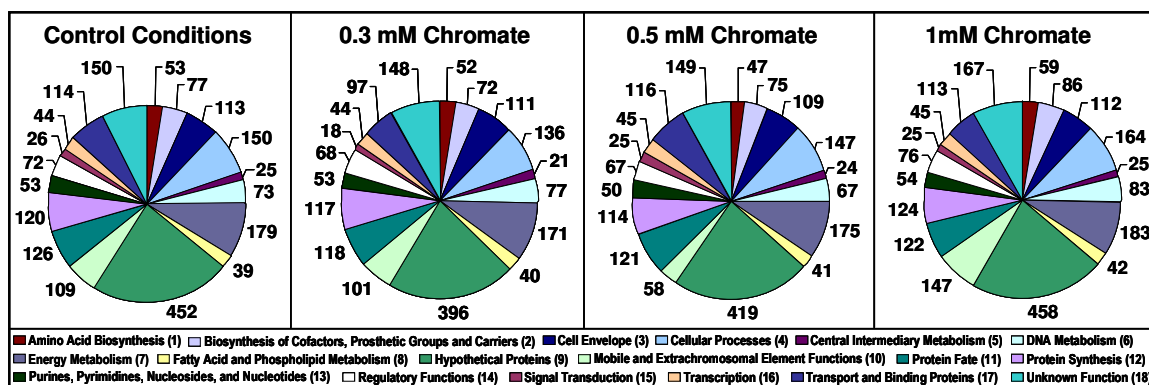
**Figure 4.2** Functional category distribution for each dosage and the control condition. The total number of non-redundant proteins identified by LTQ-MS under each dosage condition organized according to the functional categories assigned by TIGR. The number of proteins identified under each functional category is labeled by the slice in which it is located for each dosage condition and the control condition.

The Venn diagrams in Figure 4.3 illustrate the relationship of the proteins identified under the three chromate dosage treatments: 1,535 proteins were identified under all three dosage conditions using the LTQ, while relatively few proteins (ranging from 86 to 231 proteins) were unique to a specific chromate dose, thus suggesting that the majority of the proteome was not measurably altered in response to acute chromate exposure. However, as detailed below, the actual proteins identified as putatively differentially expressed for each dosage treatment are for the most part unique for a given dose.

*Dosage-Dependent Response of Differentially Expressed Proteins*

The information in Supplemental Table S2 (see Supporting Information) (percent sequence coverage, peptide count, and spectral count) was used to identify those proteins from the three dosage treatments that were differentially expressed when compared to the control sample (as detailed in Chapter 2). Semiquantitation or label free quantitation is important in cases where other methods of quantitation like isotope coded affinity tags (ICAT) [177] or metabolic labeling [175] are not feasible for sample characterization. The semi-quantitation method relies on the average between the two sample runs of the same treatment, as compared to the average of the two runs of the control condition (Table 4.1). A previous study by our group [184] sampled technical replications from six growth conditions in *Rhodopseudomonas palustris* and established that a 70% measurement of reproducibility in replicated samples was readily and consistently achievable under optimal instrument conditions. This criterion has become the standard in our laboratories and has enabled more confident determination of proteins changing
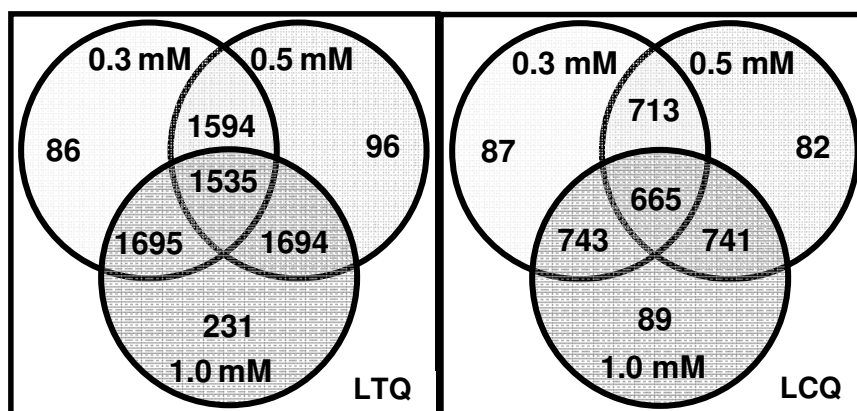
**Figure 4.3.** Venn diagrams of dosage response proteome samples for the LTQ and LCQ. The Venn diagram on the left represents results using an LTQ mass spectrometer. The Venn diagram on the right depicts results using an LCQ mass spectrometer. The top left circle is the 0.3 mM $K_2CrO_4$ dosage sample, the top right circle is the 0.5 mM, and the bottom circle is the 1 mM. The numbers are the total number of proteins shared between the respective samples. For instance, 1535 proteins were identified in all three dosage conditions for the LTQ and 665 proteins were identified in all three doses with the LCQ.

abundance as a result of different biological growth conditions as opposed to varying sampling and instrumentation fluctuations.

A total of 14 proteins were up-regulated under all three Cr(VI) doses, with representatives in bold in Tables 4.3-4.5. In comparing the 0.3 mM chromate dosage with the other two higher doses, this dose appears to elicit a unique response in which only three up-regulated proteins are shared with the 0.5 mM dose. The datasets for the 0.5 and 1 mM chromate doses had 13 up-regulated proteins in common, suggesting that the toxicity of these two Cr(VI) concentrations may induce a similar molecular response. A greater variation was observed among the various chromate doses for proteins identified as down-regulated in contrast to up-regulated proteins, with only 10 proteins exhibiting decreased abundance levels across all three metal treatments. The 0.3 and 0.5 mM dosages were most similar with seven additional down-regulated proteins shared between the two; the down-regulated protein profiles for doses 0.3 and 1 mM shared a total of five proteins. No additional proteins were identified that were solely shared between the 0.5 and 1 mM doses, indicating that down-regulated protein expression in response to chromate was much more varied and showed more instances of dose-dependent patterns compared to the up-regulated expression profiles. Proteins with annotated functions in transport and binding consistently displayed up-regulation, with representatives over the treatments remaining constant as well. For down-regulated proteins, all three dosages displayed a large number of energy metabolism proteins with reduced expression. However, only a small subset of the proteins are conserved over the three treatment conditions. In addition, the representative functional categories varied

103

**Table 4.3.** Selected Differentially Expressed Proteins Identified after Treatment with 0.3 mM K$_2$CrO$_4$[a]

| Locus | 0.3 mM Average | | | Control Condition Average | | | Functional Category[b] | Description |
|---|---|---|---|---|---|---|---|---|
| | % coverage | peptide | spectra | % coverage | peptide | spectra | | |
| **Up-regulated** | | | | | | | | |
| **SO1190** | **69.3%** | **25** | **75.5** | **48.0%** | **11** | **19** | **9** | **conserved hypothetical protein** |
| **SO1482** | **78.2%** | **63.5** | **192.5** | **52.2%** | **26** | **35** | **17** | **TonB-dependent receptor, putative** |
| **SO2426** | **31.7%** | **6.5** | **11** | **0.0%** | **0** | **0** | **15** | **DNA-binding response regulator** |
| **SO3030** | **43.3%** | **20.5** | **37.5** | **4.0%** | **1** | **2** | **17** | **siderophore biosynthesis protein (alcA)** |
| **SO3033** | **50.6%** | **38** | **112** | **27.5%** | **15** | **23** | **17** | **ferric alcaligin siderophore receptor** |
| **SO3599** | **56.2%** | **16.5** | **49.5** | **42.0%** | **13** | **18** | **17** | **sulfate ABC transporter, periplasmic sulfate-binding protein (cysP)** |
| **SO3667** | **91.4%** | **24** | **88** | **25.2%** | **4** | **4** | **9** | **conserved hypothetical protein** |
| **SO3669** | **76.1%** | **47** | **269** | **11.2%** | **4** | **4** | **17** | **heme transport protein (hugA)** |
| **SO3914** | **62.4%** | **56** | **199.5** | **44.5%** | **30** | **49** | **17** | **TonB-dependent receptor, putative** |
| SO4077 | 42.9% | 23.5 | 39.5 | 9.6% | 4 | 5 | 17 | TonB-dependent receptor, putative |
| SO4523 | 65.8% | 42.5 | 146.5 | 51.4% | 28 | 47 | 17 | iron-regulated outer membrane virulence protein (irgA) |
| **SO4743** | **74.6%** | **66** | **223** | **67.1%** | **46** | **100** | **17** | **TonB-dependent receptor, putative** |
| **Down-regulated** | | | | | | | | |
| SO0004 | 24.6% | 9 | 13.5 | 36.5% | 23 | 42.5 | 3 | inner membrane protein, 60 kDa |
| SO0848 | 45.2% | 31.5 | 45.5 | 66.4% | 68 | 186.5 | 7 | periplasmic nitrate reductase (napA) |
| SO0947 | 28.0% | 9.5 | 12.5 | 46.1% | 19.5 | 32 | 16 | ATP-dependent RNA helicase SrmB (srmB) |
| SO0970 | 52.5% | 27.5 | 51.5 | 55.9% | 38.5 | 99 | 7 | fumarate reductase flavoprotein subunit precursor |
| SO1425 | 36.9% | 9 | 12 | 45.2% | 22 | 30.5 | 9 | hypothetical protein |
| SO1490 | 30.7% | 7.5 | 11.5 | 55.9% | 21 | 43 | 7 | alcohol dehydrogenase II (adhB) |
| SO1519 | 34.7% | 18.5 | 23.5 | 54.9% | 33.5 | 63 | 7 | iron-sulfur cluster-binding protein |
| SO1677 | 68.6% | 22 | 60.5 | 74.9% | 32.5 | 112.5 | 8 | acetyl-CoA acetyltransferase (atoB) |
| SO1679 | 36.6% | 9.5 | 15.5 | 53.2% | 18.5 | 28.5 | 8 | acyl-CoA dehydrogenase family protein |
| SO1776 | 11.6% | 4 | 4 | 40.0% | 30.5 | 35.5 | 3 | outer membrane protein precursor MtrB (mtrB) |

**Table 4.3.** Continued

| Locus | 0.3 mM Average | | | Control Condition Average | | | Functional Category[b] | Description |
|---|---|---|---|---|---|---|---|---|
| | % coverage | peptide | spectra | % coverage | peptide | spectra | | |
| **SO1778** | **28.2%** | **15** | **25.5** | **43.3%** | **39** | **80** | **7** | **decaheme cytochrome c (omcB)** |
| SO2606 | 19.0% | 11 | 16 | 51.6% | 36.5 | 55 | 18 | PqiB family protein |
| **SO2929** | **45.6%** | **15.5** | **22** | **69.6%** | **31** | **56** | **9** | **hypothetical protein** |
| **SO4053** | **9.5%** | **3.5** | **3.5** | **38.4%** | **18** | **25.5** | **4** | **methyl-accepting chemotaxis protein** |
| **SO4513** | **8.0%** | **6** | **8** | **42.3%** | **42** | **67.5** | **7** | **formate dehydrogenase, alpha subunit** |

[a]Proteins in bold are found differentially expressed under all three Cr(VI) dosage conditions. [b]Functional Category number

refers to the designation in Figure 4.2 as an abbreviation for the category name.

**Table 4.4.** Selected Differentially Expressed Proteins Identified after Treatment with 0.5 mM $K_2CrO_4{}^a$

| Locus | 0.5 mM Average | | | Control Condition Average | | | Functional Category[b] | Description |
|---|---|---|---|---|---|---|---|---|
| | % coverage | peptide | spectra | % coverage | peptide | spectra | | |
| **Up-regulated** | | | | | | | | |
| SO0578 | 51.7% | 36 | 68 | 29.6% | 17 | 29 | 9 | hypothetical protein |
| **SO1190** | **84.0%** | **33** | **118** | **48.0%** | **11** | **19** | **9** | **conserved hypothetical protein** |
| **SO1482** | **71.0%** | **69** | **201** | **52.2%** | **26** | **35** | **17** | **TonB-dependent receptor, putative** |
| **SO2426** | **22.4%** | **5** | **18** | **0.0%** | **0** | **0** | **15** | **DNA-binding response regulator** |
| SO2903 | 92.1% | 59 | 617 | 82.3% | 36 | 150 | 1 | cysteine synthase A (cysK) |
| **SO3030** | **48.7%** | **22** | **47** | **4.0%** | **1** | **2** | **17** | **siderophore biosynthesis protein (alcA)** |
| **SO3033** | **55.0%** | **38** | **100** | **27.5%** | **15** | **23** | **17** | **ferric alcaligin siderophore receptor** |
| SO3420 | 75.8% | 23 | 100 | 54.4% | 9 | 30 | 7 | cytochrome c |
| SO3509 | 40.5% | 23 | 28 | 17.1% | 10 | 12 | 7 | beta-hexosaminidase b precursor (hex) |
| SO3585 | 28.0% | 6 | 8 | 0.0% | 0 | 0 | 4 | azoreductase, putative |
| **SO3599** | **80.0%** | **34** | **170** | **42.0%** | **13** | **18** | **17** | **sulfate ABC transporter, periplasmic sulfate-binding protein (cysP)** |
| **SO3667** | **91.4%** | **33** | **281** | **25.2%** | **4** | **4** | **9** | **conserved hypothetical protein** |
| **SO3669** | **69.6%** | **49** | **228** | **11.2%** | **4** | **4** | **17** | **heme transport protein (hugA)** |
| **SO3914** | **61.7%** | **52** | **126** | **44.5%** | **30** | **49** | **17** | **TonB-dependent receptor, putative** |
| **SO4743** | **74.1%** | **64** | **271** | **67.1%** | **46** | **100** | **17** | **TonB-dependent receptor, putative** |
| **Down-regulated** | | | | | | | | |
| SO0441 | 41.3% | 13 | 27 | 49.8% | 23 | 70.5 | 13 | phosphoribosylamine--glycine ligase (purD) |
| SO1185 | 43.5% | 10 | 28 | 38.2% | 15 | 57 | 9 | conserved hypothetical protein TIGR00092 |
| SO1519 | 30.9% | 19 | 27 | 54.9% | 33.5 | 63 | 7 | iron-sulfur cluster-binding protein |
| **SO1776** | **17.0%** | **6** | **6** | **40.0%** | **30.5** | **35.5** | **3** | **outer membrane protein precursor MtrB (mtrB)** |
| SO1853 | 16.5% | 6 | 7 | 29.3% | 16 | 25.5 | 17 | ABC transporter, ATP-binding protein |
| SO2304 | 38.2% | 10 | 27 | 47.6% | 20.5 | 45 | 7 | alanine dehydrogenase, authentic point mutation (ald) |
| SO2590 | 16.3% | 3 | 3 | 44.6% | 16 | 23.5 | 18 | GTP-binding protein |
| **SO2929** | **39.3%** | **10** | **14** | **69.6%** | **31** | **56** | **9** | **hypothetical protein** |

**Table 4.4.** Continued

| Locus | 0.5 mM Average | | | Control Condition Average | | | Functional Category | Description |
|---|---|---|---|---|---|---|---|---|
| | % coverage | peptide | spectra | % coverage | peptide | spectra | | |
| SO3783 | 12.6% | 5 | 7 | 36.7% | 14 | 16 | 16 | ATP-dependent RNA helicase, DEAD box family |
| SO3863 | 50.4% | 11 | 22 | 68.6% | 16.5 | 41 | 17 | molybdenum ABC transporter, periplasmic molybdenum-binding protein (modA) |
| **SO4053** | **13.5%** | **7** | **8** | **38.4%** | **18** | **25.5** | **4** | **methyl-accepting chemotaxis protein** |
| SO4066 | 20.3% | 5 | 7 | 64.1% | 25 | 36.5 | 13 | phosphoribosylaminoimidazole-succinocarboxamide synthase, putative |
| SO4249 | 37.1% | 6 | 11 | 47.1% | 13.5 | 21.5 | 2 | DNA/pantothenate metabolism flavoprotein (dfp) |
| **SO4513** | **19.8%** | **16** | **22** | **42.3%** | **42** | **67.5** | **7** | **formate dehydrogenase, alpha subunit** |

[a]Proteins in bold are found differentially expressed under all three Cr(VI) dosage conditions.  [b]Functional Category number refers to the designation in Figure 4.2 as an abbreviation for the category name.

**Table 4.5.** Selected Differentially Expressed Proteins Identified after Treatment with 1 mM $K_2CrO_4$[a]

| Locus | 1.0 mM Average | | | Control Condition Average | | | Functional Category[b] | Description |
|---|---|---|---|---|---|---|---|---|
| | % coverage | peptide | spectra | % coverage | peptide | spectra | | |
| **Up-regulated** | | | | | | | | |
| SO0343 | 48.0% | 38 | 69 | 20.5% | 14 | 14 | 7 | aconitate hydratase 1 (acnA) |
| SO0578 | 49.2% | 32 | 68 | 29.6% | 17 | 29 | 9 | hypothetical protein |
| **SO1482** | **70.5%** | **60** | **160** | **52.2%** | **26** | **35** | **17** | **TonB-dependent receptor, putative** |
| **SO2426** | **32.3%** | **8** | **13** | **0.0%** | **0** | **0** | **15** | **DNA-binding response regulator** |
| SO2903 | 90.7% | 54 | 482 | 82.3% | 36 | 150 | 1 | cysteine synthase A (cysK) |
| SO2912 | 62.7% | 54 | 115 | 36.8% | 26 | 57 | 7 | formate acetyltransferase (pflB) |
| **SO3030** | **62.0%** | **30** | **77** | **4.0%** | **1** | **2** | **17** | **siderophore biosynthesis protein (alcA)** |
| **SO3033** | **57.0%** | **46** | **158** | **27.5%** | **15** | **23** | **17** | **ferric alcaligin siderophore receptor** |
| SO3420 | 69.8% | 20 | 66 | 54.4% | 9 | 30 | 7 | cytochrome c |
| SO3577 | 58.8% | 66 | 167 | 34.0% | 33 | 51 | 11 | clpB protein (clpB) |
| SO3585 | 42.7% | 10 | 32 | 0.0% | 0 | 0 | 4 | azoreductase, putative |
| **SO3599** | **81.4%** | **35** | **107** | **42.0%** | **13** | **18** | **17** | **sulfate ABC transporter, periplasmic sulfate-binding protein (cysP)** |
| **SO3667** | **92.2%** | **34** | **241** | **25.2%** | **4** | **4** | **9** | **conserved hypothetical protein** |
| **SO3669** | **72.4%** | **54** | **199** | **11.2%** | **4** | **4** | **17** | **heme transport protein (hugA)** |
| SO3737 | 67.0% | 52 | 112 | 26.1% | 14 | 20 | 5 | sulfite reductase (NADPH) hemoprotein beta-component (cysI) |
| SO3914 | 56.5% | 53 | 217 | 44.5% | 30 | 49 | 17 | TonB-dependent receptor, putative |
| SO4215 | 65.9% | 31 | 90 | 45.9% | 16 | 36 | 4 | cell division protein FtsZ (ftsZ) |
| **SO4743** | **69.6%** | **61** | **266** | **67.1%** | **46** | **100** | **17** | **TonB-dependent receptor, putative** |
| SOA0048 | 45.6% | 31 | 49 | 23.7% | 12 | 15 | 11 | prolyl oligopeptidase family protein |
| **Down-regulated** | | | | | | | | |
| SO0029 | 20.8% | 7 | 12.5 | 30.7% | 15 | 29 | 17 | potassium uptake protein TrkA (trkA) |
| SO0130 | 10.0% | 4.5 | 5 | 29.6% | 21 | 28 | 11 | protease, putative |
| **SO0398** | **6.1%** | **2** | **2** | **28.5%** | **13** | **19** | **7** | **fumarate reductase flavoprotein subunit (frdA)** |

**Table 4.5.** Continued

| Locus | 1.0 mM Average | | | Control Condition Average | | | Functional Category[b] | Description |
|---|---|---|---|---|---|---|---|---|
| | % coverage | peptide | spectra | % coverage | peptide | spectra | | |
| SO0848 | 41.9% | 34 | 60.5 | 66.4% | 68 | 187 | 7 | periplasmic nitrate reductase (napA) |
| SO0902 | 24.8% | 6.5 | 8 | 39.0% | 16 | 22 | 7 | NADH:ubiquinone oxidoreductase, Na translocating, alpha subunit (nqrA-1) |
| SO0907 | 11.9% | 5 | 7 | 32.3% | 15 | 19 | 7 | NADH:ubiquinone oxidoreductase, Na translocating, beta subunit (nqrF-1) |
| SO1066 | 14.6% | 6.5 | 6.5 | 36.9% | 20 | 24 | 6 | extracellular nuclease |
| SO1424 | 23.0% | 10 | 12.5 | 37.3% | 25 | 29 | 9 | hypothetical protein |
| **SO1776** | **9.0%** | **4** | **6** | **40.0%** | **31** | **36** | **3** | **outer membrane protein precursor (mtrB)** |
| SO1779 | 29.4% | 17.5 | 25.5 | 47.9% | 43 | 81 | 7 | decaheme cytochrome c (omcA) |
| **SO2929** | **47.8%** | **14.5** | **21.5** | **69.6%** | **31** | **56** | **9** | **hypothetical protein** |
| SO3175 | 23.9% | 13.5 | 15 | 40.3% | 27 | 38 | 1 | asparagine synthetase, glutamine-hydrolyzing (asnB-2) |
| **SO4053** | **9.3%** | **6** | **19.5** | **38.4%** | **18** | **26** | **4** | **methyl-accepting chemotaxis protein** |
| **SO4513** | **11.7%** | **8** | **11** | **42.3%** | **42** | **68** | **7** | **formate dehydrogenase, alpha subunit** |
| SOA0141 | 2.3% | 1 | 1 | 51.4% | 24 | 36 | 9 | hypothetical protein |

[a]Proteins in bold are found differentially expressed under all three Cr(VI) dosage conditions. [b]Functional Category number refers to the designation in Figure 4.2 as an abbreviation for the category name.

over the treatment conditions, with 9-13 categories represented in each dose (as outlined in detail below).

*Differentially Expressed Proteins After Exposure to 0.3 mM Cr(VI)*

Comparison of the 0.3 mM chromate-treated sample with the control sample revealed a total of 90 proteins that were differentially expressed in response to Cr(VI), with 26 proteins up-regulated and 64 proteins down-regulated. Overall, representatives of transport and binding proteins dominated this subset of up-regulated proteins, with 54% of the proteins annotated as such. Twelve of the proteins had increased abundance levels only in response to the 0.3 mM chromate dose (dose-dependent expression) and included a conserved hypothetical protein (SO0564), TopB (DNA topoisomerase III), and a sigma-54 dependent response regulator (SO4718) (Supplemental Table S3). The functional categories of hypothetical proteins with 12 members and cellular processes (in particular, chemotaxis and motility) with 8 members dominated the list of down-regulated proteins with a total of 12 hypothetical and conserved hypothetical proteins unique to this dose. The large proportion of repressed genes with unassigned cellular functions under this Cr(VI) dosage suggests that much remains to be explored in terms of the molecular response of MR-1 to chromate toxicity. Of the two predominant functional classes, only SO2929 (a hypothetical protein) and SO4053 (a methyl accepting chemotaxis protein) consistently displayed down-regulated expression across all three chromate doses (Tables 4.3-4.5), indicating the greater influence of chromate concentration on the down-regulated protein expression profiles. Interestingly, SO2929 was also identified in another study [216] investigating *S. oneidensis* growth under aerobic conditions.

110

In response to the 0.3 mM chromate dose, 14 annotated transport and binding proteins were up-regulated and comprised members of a hemin ABC transporter complex (HmuT, HmuV), a heme transport protein (HugA), a TonB-dependent heme receptor (SO1580), siderophore biosynthesis proteins (AlcA, SO3032), a member (CysP) of a sulfate ABC transporter system, four putative TonB-dependent receptors (SO1482, SO3914, SO4077, SO4743), a ferric alcaligin siderophore receptor (SO3033), a ferric vibriobactin receptor (ViuA), and an iron-regulated outer membrane virulence protein (IrgA) (Tables 4.3-4.5 and Supplemental Table S3, Supporting Information). The vast majority of these proteins have predicted functions in iron sequestration and transport, suggesting a possible linkage between chromate stress and iron transport and/or metabolism at the molecular level. Interestingly, 10 of these putative transport and binding proteins (SO1482, AlcA, SO3032, SO3033, CysP, HugA, HmuT, HmuV, SO3914, SO4743) were identified as being up-regulated in response to all three chromate doses (Tables 4.3-4.5 and Supplemental Table S3, Supporting Information), whereas 2 proteins (SO4077, IrgA) displayed a 0.3 mM chromate dose-dependent up-regulation (Table 4.3). These proteomic results suggest that the 10 shared transport and binding proteins constitute part of the initial core molecular response to chromate exposure that is induced irrespective of environmental metal concentration. Many of these initial responders are localized to the *S. oneidensis* outer membrane, periplasmic space, and the cytoplasmic membrane, all of which constitute cellular structures or compartments that are immediately impacted by metal stress and other environmental perturbations. SO1580 (a TonB-dependent heme receptor) and SO4516 (ViuA) were identified as being up- regulated upon challenge to 0.3 and 0.5 mM Cr(VI) but not to 1 mM, whereas,

111

SO3671 (TonB system transport protein ExbB1), SO4652 (sulfate ABC transporter Sbp), and SO4655 (sulfate ABC transporter CysA-2) were basically detected only in cells exposed to the higher Cr(VI) concentrations of 0.5 and 1 mM (Supplemental Table S3, Supporting Information).

Certain proteins with predicted functions in transport and binding demonstrated dose-dependent down-regulation in response to 0.3 mM chromate. For example, two proteins displaying decreased abundance in response to the 0.3 mM chromate dose were SO4598 and SOA0153, members of the CzcA family of heavy metal efflux pumps. However, these proteins were identified at low levels (2-4 peptides) under the other two Cr(VI) doses. In bacteria, efflux pumps are a commonly employed metal resistance mechanism, with both plasmid and chromosomal systems, and active efflux of chromate appears to be a resistance strategy used by some microorganisms [126, 217]. At this point, it is not known what role, if any, efflux pumps are playing in the resistance of *S. oneidensis* MR-1 to chromate stress. Perhaps these efflux pumps (SO4598 and SOA0153) are expressed only at higher (i.e., ≥ 0.3 mM) chromate concentrations or their expression is temporally controlled depending on the levels of accumulated intracellular chromate. More detailed studies are needed to explore this efflux pump system down-regulation, but are beyond the scope of this present study.

In addition, proteins with annotated functions in chemotaxis (SO3052, SO3282, SO4454) and in the assembly of the flagellum (PomB, FliS, FlgH) were down-regulated only in response to the 0.3 mM chromate dose based on our analysis. Two of these down-regulated proteins are the flagellar protein FliS, a chaperone, and the L-ring protein FlgH, a structural component of the flagellar apparatus. FliS is the molecular chaperone

that ensures the prevention of premature folding and polymerization of FliC in the cytosol of the cell [218].  Under low dosage conditions as exhibited here, this protein is down-regulated, but the two encoded flagellum proteins (SO3237 and SO3238), which are structural components of the helical filament that extends outward from the cell, are not differentially expressed under the dosage conditions investigated here.  The down-regulation of FliS has implications with respect to cytosolic polymerization of flagellin upon exposure to low doses of chromate.  Because assembly and rotation of the flagellum require an investment of energy by the cell, it seems reasonable that MR-1 might divert energy away from the production of flagella to cellular processes more directly involved in metal stress protection and detoxification.

*Differentially Expressed Proteins After Exposure to 0.5 mM Cr(VI)*

A total of 79 (40 up-regulated and 39 down-regulated) proteins were identified as being differentially expressed in cells challenged with 0.5 mM chromate for 30 min (Table 4.4 and Supplemental Table S3, Supporting Information) relative to the control culture.  The up-regulated proteins were distributed over 10 of the 18 functional categories, where the most prevalent category was transport and binding proteins (category 17).  Other notable functional categories represented in this dose response included hypothetical and conserved hypothetical proteins, with five members, and the unknown function category, with four members.  Transport and binding proteins were represented by 17 members, constituting 44% of the up-regulated proteins in this dose response group, with 12 also showing up-regulated expression under the 0.3 mM dose.  Particularly noteworthy was the observation from these global proteomic studies that proteins involved in sulfate transport and sulfur metabolism displayed dose-dependent

113

up-regulation. In *Pseudomonas fluorescens*, it was shown that chromate acts as a competitive inhibitor for sulfate uptake via sulfate active transport systems [149]. Similarly, it is likely that chromate enters *S. oneidensis* MR-1 cells by energy-dependent sulfate transport mechanisms because of chromate's structural similarity to sulfate [126, 149, 154]. We hypothesize that the increase in abundance of sulfate transporters observed following chromate exposure could be the result of a decrease in intracellular sulfur due to competitive inhibition by chromate.

Challenge with the 0.3 mM dose resulted in the up-regulation of only one protein involved in sulfate transport: SO3599 (sulfate ABC transporter CysP) (Table 4.3). However, in response to 0.5 mM Cr(VI), we observed a substantial expansion in the number of up-regulated proteins detected for the sulfate transport and sulfur metabolism categories, with five additional representatives identified under this higher dose (Table 4.4, Supplemental Table S3). With the exception of SO3602 (CysA-1), six proteins with annotated roles in transport and binding and central intermediary metabolism [SO3599 (CysP), SO3726 (CysN), SO3727 (CysD), SO3738 (CysJ), SO4652 (Sbp), and SO4655 (CysA-2)] were also found to be up-regulated in response to 1 mM Cr(VI) treatment (Tables 4.4, 4.5 and Supplemental Table S3). Three of these proteins are annotated as members of one of two sulfate ABC transporter systems found in the MR-1 genome and are located in gene clusters *so3599-3602* or *so4652-5*. With the reduction of Cr(VI) in the bacterium, the generation of an unstable Cr(V) intermediate produces reactive oxygen species and leads to an increase in oxidative stress[126, 140]. A possible explanation for the increased level of expression observed for certain sulfur metabolism proteins in this proteomic study may be due, in part, to an increased demand for protective thiol-

containing compounds needed for coping with the oxidative stress imposed by sub-toxic metal exposures.

In addition, proteins with predicted functions in stress responses (IbpA) and DNA repair (RecN) showed increased abundance under 0.5 mM Cr(VI), indicating that this dose creates increased oxidative stress in the bacterium.  Two proteins, IbpA and RecN, were up-regulated at both the 0.5 and 1 mM Cr(VI) doses, while these same proteins were identified with a comparable number of peptides as the control under 0.3 mM Cr(VI) treatment, indicating no detectable differential expression.  However, both proteins exhibited greater abundance levels in response to the 1 mM dose compared to the 0.5 mM dose demonstrating an increased response with dosage.  With the 1 mM Cr(VI) dose, both the ATP-dependent ClpA and ClpB proteases were up-regulated.  Clp proteases have been shown to be involved in the cellular response to thermal and other types of cellular stress by working in conjunction with the chaperone proteins in *Escherichia coli* [219].  In addition, the expression of a putative azoreductase (SO3585), which is predicted to play a role in cellular detoxification, was detected only under chromate treatment, with an average of ~28 and ~43 % sequence coverage identified under 0.5 and 1 mM Cr(VI), respectively.  However, under the 0.3 mM dose treatment, SO3585 was not identified in the two analytical replicates, suggesting that this protein may be involved in metal detoxification mechanisms at higher chromate doses.  The increased expression levels of general stress response proteins, DNA repair proteins, and SO3585 indicate an elevated response to the Cr(VI) stress in the growth medium.

Three members of the functional category energy metabolism were identified at a higher abundance level after exposing *S. oneidensis* to chromate.  One member (PrpC, a

115

methylcitrate synthase) is up-regulated under all three dosage conditions. The second protein, Hex beta-hexosaminidase b precursor (SO3509), is only up-regulated in response to 0.5 mM chromate (Table 4.4), whereas SO3420 (cytochrome c) demonstrates increased abundance at higher levels of chromate (Tables 4.4 and 4.5). This cytochrome c does not change its abundance under the control condition or the 0.3 mM chromate dose, with approximately 10 peptides identified under both conditions. However, at a dosage of 0.5 and 1 mM chromate, the number of peptides identified increased to at least 20 on average for cytochrome c, suggesting the possible involvement of SO3420 in the MR-1 response to chromate stress. Previously, Myers *et al* [86] localized Cr(VI) reductase activity to the cytoplasmic membrane of anaerobically grown *S. oneidensis*. They concluded the possible involvement of cytochromes in the reductase activity, which is not inhibited by $O_2$. Also, Lovely and Phillips [214] identified the cytochrome $c_3$ of *Desulfovibrio vulgaris*, a member of the δ-proteobacteria, as a Cr(VI) reductase. Both studies either implicated or demonstrated the involvement of cytochromes in Cr(VI) reduction.

Proteins identified as down-regulated using our semi-quantitation method totaled 39 after cells were exposed to 0.5 mM chromate for 30 min, with the most dominant functional category being energy metabolism (category 7). Most of the 8 energy metabolism proteins showing differential expression were annotated as dehydrogenases and reductases (Table 4.4 and Supplemental Table S3, Supporting Information). With the exception of SO2304 and SO3546, all of the energy metabolism proteins down-regulated under the 0.5 mM chromate condition were also down-regulated under the 0.3

116

mM dose.  There appears to be a similar response between the 0.3 mM and 0.5 mM doses; however, as detailed below, this is not the case with the 1 mM dose.

*Differentially Expressed Proteins After Exposure to 1 mM Cr(VI)*

The proteome response to 1 mM Cr(VI) was characterized by a total of 92 proteins showing differential expression, with 66 proteins up-regulated and 26 proteins down-regulated (Table 4.5 and Supplemental Table S3, Supporting Information).  The functional category distribution among up-regulated proteins was similar to that elicited by the 0.3 mM dosage, where all but 5 of the functional categories were represented.  There was a bias toward the following functional categories for up-regulated proteins: amino acid biosynthesis (7 proteins); cellular processes (4); central intermediary metabolism (6); energy metabolism (12); hypothetical and conserved hypothetical proteins (5); protein fate (7); and transport and binding proteins (14).  The subsets of transport and binding proteins that were differentially expressed in response to the three different chromate doses were very similar.  Exposure of cells to 1 mM Cr(VI) resulted in only one protein (SO1072, a putative chitin-binding protein) being unique to this dosage.  The other two doses (0.3 and 0.5 mM) have a repertoire of more unique proteins in each category and share two other proteins not identified as up-regulated under the 1 mM chromate dosage (SO1580, a TonB-dependent heme receptor, and ViuA, ferric vibriobactin receptor).  The category of amino acid biosynthesis is only represented by 1 member in the other two doses, whereas for the 1 mM Cr(VI)-treated sample there are 7 members identified as up-regulated.  This same trend holds true for the functional classes of central intermediary metabolism, energy metabolism, and protein fate, which only

117

have up to 3 members in the other two Cr(VI) doses and 6-12 members with the 1 mM dose.

Four proteins were identified as being up-regulated under the 1 mM dose and annotated as involved in cellular processes. One of these proteins, a putative azoreductase (SO3585), demonstrates a dose-dependent response to Cr(VI) challenge and is up-regulated under the higher doses, as mentioned above. There are two other proteins (SO3586 and SO3587) located immediately downstream of SO3585 in the MR-1 chromosome that may be organized in an operon, as suggested previously [147]. SO3586 encodes a glyoxalase family protein and SO3587 encodes a hypothetical protein with a putative transmembrane domain [147]. SO3586-87 have a reduced level of detection with respect to SO3585, where SO3587 is identified with no more than 4 peptides. The predicted transmembrane-spanning domain of SO3587 comprises a 36 amino acid tryptic peptide (residues 38-73), which may contribute to the lack of identified peptides. SO3585-87 were identified solely under Cr(VI) challenge (Supplemental Table S2, Supporting Information). This is in contrast to a previous study [147] in which SO3587 was identified under both control and chromate treatments in response to temporally longer growth exposures (up to 90 min). It is important to note that the previous study involved control samples grown for 45 and 90 min. prior to harvesting, whereas the current study has a control sample grown for 30 min. After 90 min of exposure to Cr(VI), the growth rate of *S. oneidensis* decreased significantly with respect to control growth conditions (Supplemental Figure S1), whereas 30 and 45 min after introduction of Cr(VI) the difference is reduced. Also, in the previous study, only two peptides were identified for SO3587 in the 45 min control with the difference between them resulting

118

from a missed cleavage. However, since SO3587 is located at the terminus of the operon, the gene may not be translated to the same frequency perhaps due to the ribosome disengaging from the mRNA. This has been shown in *E. coli* using a fabricated betagalactosidase hexamer transcriptional unit,where the ribosome disengaged from the mRNA approximately half of the time when reaching the third copy of the gene [220].

An important class of proteins involved in signal transduction consists of relatively low-abundance proteins, with only about 50 members confidently identified at the two-peptide level under a given growth condition using the LTQ (Figure 4.2). Of these 50 proteins, two (SO2426, a DNA-binding response regulator, and SO4003, a response regulator) were up-regulated under doses 0.3 and 1 mM Cr(VI). Only one of these proteins (SO2426) was also up-regulated in response to the 0.5 mM dose (Tables 4.3-4.5). SO2426 was exclusively detected in the Cr(VI) treatment samples under all three doses relative to the control sample, and a MR-1 strain harboring an in-frame deletion of this gene showed impaired growth and Cr(VI) reduction activity under metal conditions (K. Chourey and D. K. Thompson, unpublished data), thus suggesting that this signal transduction component plays a regulatory role in the cellular response to chromate stress. Also, in a previous study [147], the transcript data for the *so2426* gene was consistent with the corresponding protein expression data.

As described above, proteins involved in general stress responses in the bacterium were up-regulated at a dosage of 0.5 mM Cr(VI). In addition to IbpA (16 kDa heat shock protein A) and RecN (a DNA repair protein), ClpB, PrlC, and a prolyl oligopeptidase family protein (SOA0048) were all up-regulated in response to 1 mM Cr(VI) exposure. Jiang *et al* concluded that PrlC in *E. coli* functions as a molecular chaperone [221]. The

119

greater degree of oxidative stress imposed by the higher 1 mM Cr(VI) dose correlates with the increased number of up-regulated proteins involved in stress protective response and protein fate determination.

A subset of 26 down-regulated proteins identified following acute exposure to 1 mM Cr(VI) displayed a trend similar to the other two dosages.  The two dominant categories, energy metabolism and hypothetical/conserved hypothetical proteins, constituted almost 60% of the proteins down-regulated under 1 mM Cr(VI) challenge.  Interestingly, a large number of energy metabolism proteins were also up-regulated under this dosage.  Of the 308 energy metabolism proteins predicted from the MR-1 genome annotation, 183 were identified confidently at the two-peptide level representing almost 60% of the predicted total.  Of these 183 proteins, 21 were differentially expressed under the 1 mM dose.  Five of the down-regulated energy metabolism proteins are annotated as reductases, which might suggest a reconfiguration of energy metabolism by the cell in response to chromate stress.

## Conclusions

The work presented here provides the first large-scale description (2406 proteins) of the proteome response of *S. oneidensis* to three different acute chromate concentration challenges.  Approximately 90% of the proteins identified did not demonstrate a change in abundance in response to chromate treatment; however, the remaining 10% of proteins that demonstrated an abundance change in response to chromate (Tables 4.3-4.5 and Supplemental Table S3, Supporting Information) provided some important insights into the molecular response to chromate insult.  Among the up-regulated proteins identified, it was clear that both general stress and specific chromate responses were elicited.

Primarily, the up-regulated sulfate transport system under higher chromate dosage (> 0.5 mM) suggested an increased demand for sulfate uptake and metabolism, possibly due to sulfate deficiency in the bacterium. Sulfate deficiency could lead to a decrease in the biosynthesis of certain amino acids (i.e., methionine and cysteine) and this may explain why there are many more down-regulated proteins under the lower dosage. Out of the 11 genes encoding members of the cytsteine metabolism pathway, none were differentially expressed subsequent to 0.3 mM Cr(VI) exposure and one protein (SO2903, CysK) was up-regulated at a dosage of 0.5 mM. However, subsequent to the addition of a 1 mM Cr(VI) dosage, four proteins in the cysteine metabolism pathway were up-regulated: SO1095 (a putative O-acetylhomoserine (thiol)-lyase), SO2406 (AspC-2), SO2903 (cysK), and SO3598 (cysM). The methionine metabolism pathway is composed of 12 protein-encoding genes and demonstrated one up-regulated protein in the 1 mM dosage sample, SO1095. The 0.5 and 1 mM dosages may be compensating for sulfate deficiency with the apparent increased production of sulfate transport and metabolism protein complexes. Also, during Cr(VI) reduction, a one-electron transfer can occur producing the metastable, highly reactive Cr(V) species [140] leading to an increase in reactive oxygen species within the cytoplasm. Both of these response systems may lead to the observed proteome changes detected in this study. There were at least three general stress response proteins up-regulated as well under each dose, indicating a general stress response similar to other stress-response proteome studies [222].

The functional category of transport and binding proteins, particularly those with annotated functions in iron and sulfate transport, dominated the group of up-regulated proteins. Most of the TonB-dependent receptors and other iron acquisition proteins (e.g.,

siderophore biosynthesis proteins) showed increased abundance under all three Cr(VI)

dose exposures, indicating that expression of these proteins is an important feature of the

MR-1 response to chromate.  In addition, proteins whose abundance levels behaved in a

dose-dependent manner were identified by examining differences in peptide spectral

counts between samples.  Most of these transport and binding proteins were highly up-

regulated under the 0.3 mM dosage and increased when challenged with 0.5 mM but

remained the same between the 0.5 and 1 mM dosages (Figure 4.4).  This was the general

trend for most differentially expressed proteins, as illustrated by binning and comparing

the relative abundances of the proteins between the 0.3 mM and 1.0 mM dosage

conditions (Figure 4.5).  Whereas more than 50 proteins that were up-regulated relative to

the control showed no change between the 0.3 mM and 1.0 mM conditions, there were

almost twice as many proteins up-regulated in the 1.0 mM condition relative to the 0.3

mM condition.  Many fewer proteins were down-regulated in the 1.0 mM condition.

Table S4 (see Supporting Information) lists the identities and spectral count ratios of all

proteins shown in the bins of Figure 4.5.

Recently, an instrumentation comparison study between the linear trapping

quadrupole and the three-dimensional ion trap for the proteome of the model organism *E.

coli* OP50 was published [31], with similar results to the study discussed here.  Our

present study confirmed the substantially increased ability to identify proteins at the two

peptide level in the LTQ datasets, with many proteins identified as being differentially

expressed that were not even detected with the three-dimensional quadrupole ion trap.

One objective of our present study was to conduct a direct comparison between the QIT

and LTQ technology platforms to determine which would provide more extensive insight
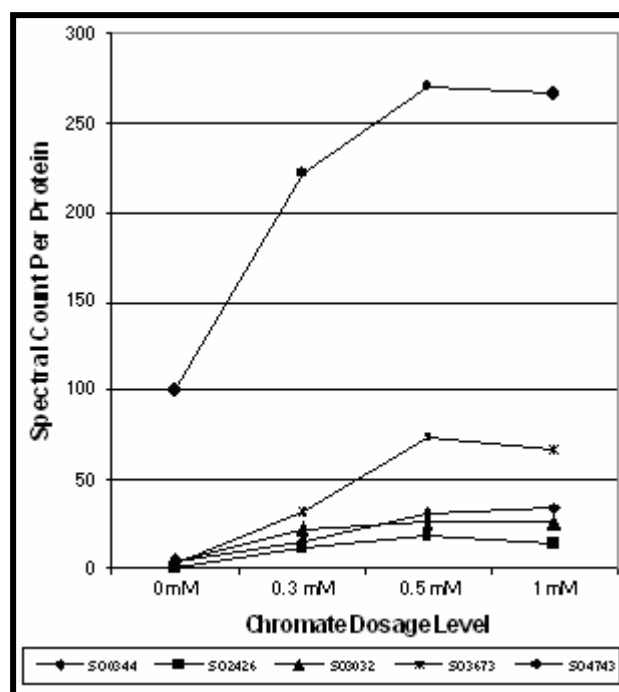
**Figure 4.4.** Proteins demonstrating dosage-dependent up-regulation.

Proteins up-regulated under all dosages tested, but were at similar levels of up-regulation

under 0.5 and 1 mM dosages. SO0344: methylcitrate synthase (prpC), SO2426: a DNA-

binding response regulator, SO3032: a putative siderophore biosynthesis protein,

SO3673: the periplasmic hemin-binding protein of the hemin ABC transporter (hmuT),

and SO4743: a putative TonB-dependent receptor.
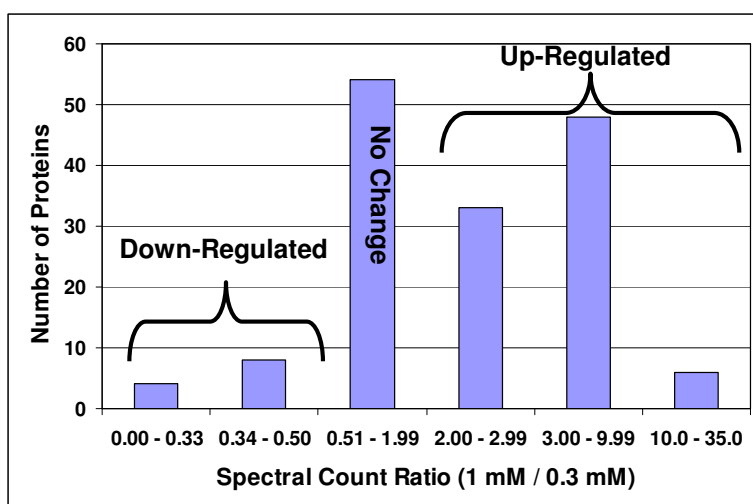
**Figure 4.5.** Dosage dependent abundance of proteins between 0.3 and 1 mM dosages.
A comparison of the relative abundances of proteins identified as differentially expressed
between the 0.3 and 1 mM Cr(VI) dosage datasets (Supplemental Table S3). The x-axis
represents the relative change in protein expression (1.0 mM / 0.3 mM), where the first
two bins represent reduced abundance of proteins in the 1 mM dose, the third bin are the
proteins that have the same concentration level in both dosages, and the last three bins
represent proteins with increased abundance in the 1 mM dosage. The y-axis represents
the number of proteins found in each bin out of a total of 153 differentially expressed
proteins. Supplemental Table S4 contains the proteins identified under each bin.

into the molecular response of *S. oneidensis* to dosage-dependent chromate insult. In addition to a substantially deeper level of proteome characterization, the LTQ datasets should provide a much better degree of correlation with differentially expressed transcriptome data determined with microarray measurements. Furthermore, the LTQ datasets provided information about low-abundance proteins not identified by the three-dimensional ion trap. Proteins involved in transcription regulation are generally of low abundance, such as SO2426, a DNA-binding response regulator. With LTQ technology, the detection of SO2426 and other proteins with regulatory functions was achievable. In fact, SO2426 was identified only under chromate exposure and correlated with the results from a previous study [147], where the transcript levels were also up-regulated in response to acute chromate challenge.

# Chapter 5

## Systematic Assessment of the Benefits and Caveats in Mining Microbial Post-Translational Modifications from Shotgun Proteomic Data; Response of *Shewanella oneidensis* to Chromate Exposure

## Introduction

During the course of examining the global proteomic response of chromate exposure in *Shewanella oneidensis*, the need to obtain information on post-translational modifications (PTMs) involved in chromate response arose. PTMs are enzymatically-driven chemical changes to amino acids and constitute one mechanism used by microbes to regulate both protein synthesis and activity. Therefore, we expect that PTMs will contribute to the response of *S. oneidensis* to Cr(VI) exposure. This chapter outlines three algorithms evaluated for the purpose of providing a foundation for global PTM searches, which has to date been unexplored.

Protein activity can be modulated via the addition and/or removal of PTMs, which act to repress or stimulate target protein function [223]. Some common PTMs include phosphorylations, methylations, acetylations, and oxidations. Methylations and acetylations in bacteria typically occur on lysine and arginine residues [14, 223].

126

Methylation modifications have been implicated in regulation of gene expression as well as protein-protein and protein-DNA interactions (i.e. the histone code) [14]. In bacterial chemotaxis, which is the response to a chemical gradient of attractant or repellent molecules, methylation is involved in the adaptive mechanism and short-term memory abilities of methyl-accepting chemotaxis receptor proteins (chemoreceptors) [54, 224]. Two enzymes, a methyltransferase (CheR) and a methylesterase (CheB), function to covalently attach and remove a methyl group, respectively, from the chemoreceptor. Acetylations have also been shown to be involved in bacterial chemotaxis [53] and provide protein terminus stability in eukaryotes [50, 225] as well. For example, the ability of the chemotactic response regulator CheY to activate the flagellar switch is enhanced upon acetylation [53]. In contrast to the above PTMs, modification via the oxidation of a protein can be due to either chemical [46, 47, 226] or in some cases biological [227, 228] mechanisms. Free radicals formed during aerobic cellular metabolism are highly reactive and often damaging proteins abundant within the cytosol through oxidation [46, 47]. Detecting the level of chemically-induced oxidation of the proteome has been hypothesized as a way to determine oxidative stress levels in organisms [46]. In eukaryotes, a well studied enzymatically driven oxidation conversion of the protein ubiquinol to ubiquinone, a vital component of the electron transport chain in mitochondria, is mediated by the cytochrome bc1 complex [228].

To date, most studies detecting PTMs in both prokaryotes and eukaryotes have been accomplished using targeted approaches [48, 50, 229, 230] by purifying protein(s) or enrichment of a specific PTM. A study by MacCoss *et al* [51] was one of the first attempts at a more global approach by searching shotgun proteomic data for PTMs. In

this study, the authors purified the *Schizosacchromyces pombe* Cdc2p complex as well as

proteins from human lens tissue. Using a three-dimensional ion trap, they were able to

characterize a number of PTMs, including phosphorylations, methylations, and

acetylations. The tandem mass spectra (MS/MS) searches were accomplished with an

initial search where no modifications were specified, followed by the creation of a sub-

database to re-search the data for PTMs [51]. The sub-database was created taking

identifications made by an initial search specifying no modifications and the -98 Da loss

common to phosphorylated peptides using the collision induced dissociation (CID)

fragmentation method. The building of a sub-database was necessary for this work in

order to reduce the false discovery rate (FDR) of identified PTM-containing peptides due

to the combinatorial increase in database search space generated by including PTMs.

In 2004, Strader *et al* [48] published a comprehensive characterization of the 70S

ribosome from *Rhodopseudomonas palustris*. The 70S ribosome was characterized using

the top-down bottom-up technique, wherein PTM information was acquired from both

the intact protein mass spectral data and peptide MS/MS data. The term top-down refers

to the measurement of the intact protein, while bottom-up corresponds to first digesting

proteins followed by acquisition of fragmentation MS leading to sequencing information

on the proteolytic peptides. The authors were able to confirm and identify novel PTMs

present on the constituent proteins of the complex. One of the primary advantages of the

top-down approach is the use of a high-resolution mass spectrometer (i.e. FT-ICR) to

discern the difference between a trimethylation and an acetylation, which are isobaric on

a lower resolution instrument (i.e. a 3-D or linear ion trap). Nielsen *et al* [231] further

demonstrated the utility of a high-resolution instrument (LTQ-FT hybrid instrument) for reducing the FDR of PTM peptides.

Phosphorylations are prevalent PTMs in both eukaryotes and prokaryotes. These PTMs play important roles in cell signaling cascades, specifically two-component signal transduction [232, 233], and in the regulation of cell cycle progression and cellular differentiation [223]. Phosphorylation sites play a critical role in both normal and diseased cell function, making them important targets for global studies. Phosphorylation in prokaryotes usually targets histidine and aspartic acid residues, producing modified species which are known to be acid labile and often do not survive the sample handling process for subsequent MS analysis. This usually precludes searching for phosphorylation in prokaryotic proteome studies. In contrast, phosphorylation in eukaryotes usually targets threonine and serine residues; these modified species are much more stable and survive the sample handling process. Unfortunately, most of these modified residues readily display phosphate neutral loss as the dominant fragmentation pathway when a low-energy, multiple collision fragmentation method is used [51, 233], thereby precluding information about the modification site. Nevertheless, Olsen *et al* [55] were able to identify and quantitate 6600 phosphorylation sites in HeLa cells using an LTQ-FTICR. To address the neutral loss problem, the authors performed MS/MS/MS to obtain more sequencing information on the modified peptide.

All of the aforementioned studies were targeted, whether that target was a protein complex (Cdc2p [51] or 70S ribosome [48]), a subset of proteins [230], or an enrichment due to the presence of a specific modification [55, 56, 58, 234]. As such, none of these targeted studies provides guidance on performing a truly non-biased global analysis of

129

PTMs in eukaryotes or prokaryotes. In addition, most PTM studies to date have focused on eukaryotic systems [46, 181, 223, 231, 234-236]. Since protein post-translational modifications also play a critical role in bacterial cellular processes, it is important to evaluate and optimize informatic methods which can accurately unravel this level of proteome detail. The goal of the present work is to identify a robust and accurate method for performing global PTM searches of shotgun proteomics data for microbial isolates. To this end, we searched a previously acquired shotgun proteomics dataset [147] derived from control (non-stressed) and chromate-stressed *S. oneidensis* cells in order to examine the potential roles of PTMs in the cellular response to acute metal challenge. Three search algorithms (DBDigger [44], InsPecT [40], and Sequest [39]) were employed to test the feasibility of performing a PTM search on a proteome and determine the feasibility of each method to provide detailed PTM information at an acceptable false discovery rate. Since the false discovery rate will increase with increasing search space, this work is a systematic attempt to compare different algorithmic tools on the same dataset for their efficacies in managing false positive rates. In addition, the search results for each algorithm were compared for overlap in PTM identifications as a method of validation, the ease of conducting searches using a standard desktop computer, and the flexibility in specifying PTMs in the search parameters. Note that this study was focused on data mining of peptide MS/MS data, and did not attempt to provide PTM heterogeneity at the protein level.

## **Materials and Methods**

*Experimental dataset*

The shotgun proteomics dataset used in the PTM searches was acquired

previously [147]. Briefly, four separate cultures of *S. oneidensis* MR-1 were grown

aerobically to mid-log phase followed by addition of a sub-lethal dose (1 mM) of $K_2CrO_4$

to two of the cultures, which were exposed to chromate for 45 or 90 min and then

harvested for proteome characterization. The other two cultures served as controls and

were grown in parallel under identical conditions excluding metal addition. The goal of

the previous work was to identify differentially expressed proteins implicated in the

response to a sub-lethal, acute chromate exposure [147]. Cultures were lysed by

sonication and fractionated into a crude/soluble and membrane-associated fraction by

high-speed centrifugation. Lysates were then digested with trypsin and analyzed via

online two-dimensional liquid chromatography (strong cation exchange and reverse

phase) coupled to a LTQ mass spectrometer (Thermo Scientific, Waltham, MA). The

LTQ was operated in a data-dependent mode for 24 h, and digested lysates were analyzed

in duplicate.

*Computational searches and algorithms*

The above data (a total of 8 datasets) was searched using three separate search

algorithms: DBDigger [44], InsPecT [40], and Sequest [39]. The database(s) used for all

searches performed can be downloaded from the project website compbio.ornl.gov/

shewanella_metal_stress/databases/. A reverse database was used for estimating the FDR

in this study by applying the equation from Peng *et al* [185]. The use of a reverse

database is one method to estimate the FDR, where peptide(s) and/or protein(s) matching

nonsense sequences from the reverse database are considered incorrect identifications. The false negative rate is a more difficult number to quantify in shotgun proteomics datasets, because the exact number and identity of all true peptide identifications are unknown. This number was qualitatively determined in the datasets presented here by comparing the total number of peptides or proteins identified in the unmodified searches to the searches specifying PTMs.

In our lab, DBDigger has been the algorithm of choice for PTM searches, because it allows simultaneous evaluation of sequentially modified forms of individual amino acid residues in the same peptide. DBDigger was used initially to develop proper filtering thresholds. The following PTMs were specified in separate DBDigger searches: (1)mono-, (2)di-, and (3)trimethylations on lysines and arginines; (4)monomethylations on glutamate; (5)mono- and (6)diacetylations on lysines and arginines; and (7)monooxidations on methionines, cysteines, tyrosines, and tryptophan; dioxidation on methionines and cysteines; and trioxidation on cysteine residues. Table 5.1 depicts the modifications and corresponding residue mass shifts that were targeted in the present study. Due to the use of an LTQ mass spectrometer in this study, phosphorylations were not searched in the resulting datasets. Results from DBDigger were stored in the form of .sqt files that were then filtered and sorted using DTASelect [171] with the following options: -m 0 (extracts only modified peptides) and --DB (exports results in database format). The resulting DTASelect output files were imported into Microsoft Access for storage and data analysis.

Sequest searches were performed on a subset of mass spectra initially identified as PTM-containing by DBDigger as follows: selected DTA files were extracted from one

**Table 5.1.** Amino Acid Residues and Corresponding Post-translational Modifications

| Amino Acid | PTM[a] | Amino Acid Mass | Amino Acid + PTM mass (Da) |
|---|---|---|---|
| Lysine | $+CH_2$ | 128.09 | 142.11 |
| Lysine | $+C_2H_4$ | 128.09 | 156.12 |
| Lysine | $+C_3H_6$ | 128.09 | 170.14 |
| Lysine | $+COCH_2$ | 128.09 | 170.10 |
| Lysine | $+C_2O_2C_2H_4$ | 128.09 | 210.11 |
| Arginine | $+CH_2$ | 156.10 | 170.12 |
| Arginine | $+C_2H_4$ | 156.10 | 184.13 |
| Arginine | $+C_3H_6$ | 156.10 | 198.15 |
| Arginine | $+COCH_2$ | 156.10 | 198.11 |
| Arginine | $+C_2O_2C_2H_4$ | 156.10 | 238.12 |
| Glutamic Acid | $+CH_2$ | 129.04 | 143.06 |
| Methionine | $+O$ | 131.04 | 147.03 |
| Methionine | $+O_2$ | 131.04 | 163.02 |
| Cysteine | $+O$ | 103.01 | 119.00 |
| Cysteine | $+O_2$ | 103.01 | 134.99 |
| Cysteine | $+O_3$ | 103.01 | 150.99 |
| Tryptophan | $+O$ | 186.08 | 202.07 |
| Tyrosine | $+O$ | 163.06 | 179.05 |

[a]Empirical formula for post-translational modification

directory and placed into another directory using DTACopy (Vilmos Kertesz, Oak Ridge

National Laboratory) and results were filtered and sorted using DTASelect as described

above. Since Sequest is only able to search three modifications at a time, the above DTA

files were searched repeatedly until all of the above modifications had been specified in

the sequest.params file (total of 8 searches for each dataset). InsPecT searches were

performed in one search where all 8 modification events on 7 amino acids were specified

(Table 5.1). A total of 3 modification events were allowed for each peptide. The

following filter levels were used for the InsPecT dataset: a minimum of 2 peptides

identified per protein and a p-value score that gives a FDR of 2% for the total peptide

dataset (see *InsPecT scoring optimization for post-translationally modified peptides* for

details). The results were then imported into Access in order to compare the PTM

identifications with the DBDigger results.

## **Results and Discussion**

*DBDigger scoring optimization for post-translationally modified peptides*

        Initially, searches were performed using the DBDigger algorithm on the 45 min

Control dataset in order to determine the correct filter thresholds for minimizing the FDR

and false negative rate. The 45 min Control Run1 dataset was searched using two

methods: (1) raising the scoring threshold levels and (2) creating a sub-database prior to

PTM searches applying loose filtering criteria and no modifications. As a comparison,

all searches were performed either with no modifications specified or with

monomethylations on lysines and arginines. This particular PTM was chosen first due to

the small mass shift in the amino acid residue by the addition of a methyl group, and

preliminary searches demonstrated a high FDR through the use of a reverse database with

this modification. Table 5.2 depicts the FDRs for a number of filter threshold levels tested and demonstrates a basic premise of data filtering in shotgun proteomics at present. This premise is that care must be taken in order to maximize the number of resulting peptide identifications made while minimizing the FDR of the filter threshold levels chosen. As a control, the dataset was searched with no modifications using "normal" filtering levels that give a FDR of 2.6% at the two peptide threshold level corresponding to the minimum peptide score of 25 for +1 peptides, 30 for +2 peptides, and 45 for +3 peptides using DBDigger. Using these same scoring minima, the addition of a differential modification (monomethylation) to lysines and arginines raises the FDR to 7.6%. The increased FDR is expected and is caused by the combinatorial increase in database search space due to the modification. Increasing the minimum peptide level for protein identification to four decreased the FDR to 1.7%; however, this also led to a dramatic decrease in the number of proteins identified. Therefore, the decision was made to pursue increasing the minimum peptide score while retaining the minimum two peptides for protein identification. Increasing the minimum scores to 29 (+1), 34 (+2), and 49 (+3) leads to a FDR of 1.9%, while maintaining a protein identification level similar to the non-modification search. There was less of a reduction in protein identifications using stringent peptide scoring filters (283 proteins) than by increasing the minimum peptide count to four (585 proteins), demonstrating a relatively lower false negative rate for the stringent peptide scoring filters.

The other option explored for retaining a FDR of 2%, while identifying at least a similar number of proteins compared to a non-modification search, is to construct a sub-database for subsequent PTM searching. This was performed by MacCoss *et al* [51] on a

135

**Table 5.2.** False Discovery Rates for 45min Control Run1 Dataset at Various Filter Threshold Levels

| No PTM | | 2 peptide filter level | | | | | | 4 peptide filter level | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Protein | | | Peptide | | | Protein | | | Peptide | |
| 25, 30, 45[a] | 156[b] | 2111[c] | **14.8%**[d] | 333[e] | 25943[f] | **2.6%**[g] | 9 | 1443 | **1.2%** | 17 | 24390 | **0.1%** |
| 27, 32, 47 | 68 | 1933 | **7.0%** | 136 | 24111 | **1.1%** | 7 | 1383 | **1.0%** | 8 | 22817 | **0.1%** |
| 29, 34, 49 | 20 | 1816 | **2.2%** | 37 | 22643 | **0.3%** | 3 | 1326 | **0.5%** | 3 | 21482 | **0.0%** |

| Monomethylation | | 2 peptide filter level | | | | | | 4 peptide filter level | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Protein | | | Peptide | | | Protein | | | Peptide | |
| 25, 30, 45 | 427 | 2560 | **33.4%** | 1069 | 28112 | **7.6%** | 51 | 1526 | **6.7%** | 224 | 25690 | **1.7%** |
| 27, 32, 47 | 215 | 2146 | **20.0%** | 503 | 25388 | **4.0%** | 21 | 1420 | **3.0%** | 75 | 23685 | **0.6%** |
| 29, 34, 49 | 96 | 1924 | **10.0%** | 217 | 23330 | **1.9%** | 9 | 1343 | **1.3%** | 25 | 21952 | **0.2%** |
| 31, 36, 51 | 50 | 1818 | **5.5%** | 106 | 21764 | **1.0%** | | | | | | |
| 35, 40, 55 | 15 | 1689 | **1.8%** | 32 | 19089 | **0.3%** | | | | | | |

[a]Minimum peptide scores for +1, +2, and +3 charged peptides, respectively. Number of proteins[b] or peptides[e] identified matching to reverse database sequences. Total number of proteins[c] or peptides[f] identified from both the forward and reverse databases. [d,g]False discovery rate value calculated as described in Peng *et al* [185].

yeast protein complex and human lens tissue.  The sub-database in this study was

constructed by using loose filtering criteria in an initial non-modification search.  The

filtering criteria were as follows:  1 peptide minimum for protein identification and

peptide scores of 25 (+1), 30 (+2), and 45 (+3).  These criteria were chosen since the

minimum score will filter out 92.4% of the false peptide identifications, but retain a

reasonable database size so that when the modification search is done, the identification

of a protein with one modified and one unmodified peptide is possible.  The sub-database

consists of 2,751 forward protein sequences plus their corresponding reversed sequences,

thus yielding a total database size of 5,502 protein sequence entries.  The normal *S.*

*oneidensis* proteome database size is 4,798 forward protein sequences; therefore, the sub-

database reduced the number of protein entries by approximately half.  After the sub-

database was created, a modification search for monomethylations on lysines and

arginines was performed with the following threshold filters:  "normal" [2 peptides per

protein, 25 (+1), 30 (+2), and 45 (+3)] and "stringent" [2 peptides per protein, 29 (+1), 34

(+2), and 49 (+3)].  The "normal" filters yielded a FDR of 7.1% and the "stringent" filters

1.8%.  These FDRs are similar to those above for the full-size database, and comparing

the peptides identified with both strategies leads to an overlap in identification of 90%

(Figure 5.1).  The use of a sub-database for PTM searches on bacterial species does not

appear to be necessary based on the results obtained here.  The caveats of first creating a

sub-database include the added complexity of performing two searches for each dataset,

creating a separate database for each dataset, and a relatively insignificant decrease in

overall search time.  Therefore, the decision was made for the remaining searches to use a

Monomethylation Peptide IDs of 45 min Control Run1
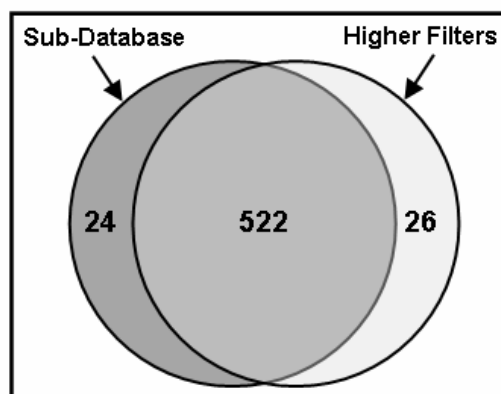
**Figure 5.1.** A Venn diagram comparing peptide identifications.

The results are from a search for monomethylations on lysines and arginines by two different search methods. The circle on the left represents the modification search with the sub-database and the circle on the right represents the modification search with the full-size database and stringent filters. The value in the middle (522) is the number of modified peptide identifications shared between the two methods and corresponds to a 90% overlap between the two methods.

full-size database with the stringent filters [2 peptides per protein, 29 (+1), 34 (+2), and 49 (+3)].

Using the above threshold filters, many peptides with putative PTMs were identified with the search algorithm DBDigger [44]. DBDigger allows more flexibility in searching for differential PTMs than the search algorithm Sequest, because it generates a candidate sequence once for the collection of spectra rather than once for every spectrum and is much faster in searching complex datasets [44]. Therefore, initial searches with DBDigger were conducted in order to obtain a preliminary list of putative modified peptides. There were a total of 8494, 7203, 9358, and 6588 non-redundant peptides, respectively, containing at least one post-translationally modified amino acid residue for the 45 min Control, 45 min Cr, 90 min Control, and 90 min Cr datasets. Table 5.3 displays the number of peptides identified in each dataset for the PTMs specified, and all modified peptides are listed in Supplemental Tables S1 and S2 (Supporting Information). However, as also depicted in Table 5.3, the FDR for modified searches is high (~50%) compared to the overall FDR described in Table 5.2. This high FDR led to the testing of two other algorithms (InSPecT [40] and Sequest [39]) to determine if the FDR of a PTM search could be improved.

*Sequest scoring optimization for post-translationally modified peptides*

MS/MS in the form of DTA files (tab-delimited flat-files for each spectrum) identified as PTM-containing peptides from the above DBDigger searches were extracted using DTACopy (Vilmos Kertesz, ORNL) and interrogated with the algorithm Sequest. This algorithm was chosen in order to help validate the DBDigger results since Sequest is a widespread algorithm for searching shotgun proteomics data. Only selected MS/MS

139

**Table 5.3.** Number of Post-Translationally Modified Peptides Identified in Each Dataset Searched using DBDigger

| Modification Type | 45 min Control | FDR[a] | 45 min Cr | FDR | 90 min Control | FDR | 90 min Cr | FDR |
|---|---|---|---|---|---|---|---|---|
| Monomethylation | 1147[b] | 53.9% | 1167 | 62.4% | 1356 | 45.7% | 902 | 53.8% |
| Dimethylation | 1058 | N/A[c] | 852 | N/A | 951 | N/A | 890 | N/A |
| Trimethylation | 889 | N/A | 769 | N/A | 798 | N/A | 720 | N/A |
| Monoacetylation | 709 | 62.1% | 594 | 65.3% | 619 | 59.0% | 557 | 65.5% |
| Diacetylation | 779 | N/A | 669 | N/A | 667 | N/A | 620 | N/A |
| Oxidation | 5567 | 8.0% | 3940 | 8.6% | 7068 | 4.2% | 3209 | 9.3% |

[a]FDR: false discovery rate. [b]Number of PTM peptides identified. [c]N/A: not available

first identified by DBDigger were searched due to the slower speed of Sequest on desktop computers. The resulting OUT files were filtered and sorted with DTASelect [171]. The filtering thresholds used in the Sequest searches were a one-peptide requirement for protein identification and 1.8 for +1 charged peptides, 2.5 (+2), and 3.5 (+3). Table 5.4 gives the number of peptides identified using Sequest as well as the number of MS/MS searched (DTA files extracted).

As apparent in Table 5.4, the number of MS/MS identified in the Sequest searches compared to the total number searched is much lower for all PTMs considered. There are two possible reasons for this low identification rate using Sequest: one is due to the high FDR of PTM-containing peptides obtained with DBDigger, and the other may be the result of the search algorithm itself. Concerning the high FDR, the average for monomethylated peptides was 54% using DBDigger. In addition, the Sequest searches were performed on all MS/MS spectra identified by DBDigger as containing a PTM including MS/MS identified from relevant reverse database searches and matching protein contaminant sequences. Therefore, the MS/MS count is somewhat inflated with erroneous identifications (false-positives from the reverse database) and protein contaminant identifications. The inflated MS/MS count may be an underlying cause of the low identification rate from Sequest.

The second possible reason for the resulting low identification rate may lie in the fact that Sequest was not originally designed to search for differential PTMs. The original design of the algorithm required that the PTM mass (i.e. 14 Da for monomethylation) be added to the amino acid mass [i.e. 14 Da + lysine (128.09 Da)] with the modification considered static or always present [39]. This search method is limiting,

141

**Table 5.4.** Sequest Peptide Identifications from a Subset of MS/MS Spectra First Identified by DBDigger

| Dataset | Monomethylation | | Dimethylation | | Trimethylation | | Monoacetylation | | Diacetylation | | Oxidation | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Spectra | Identified | Spectra | Identified | Spectra | Identified | Spectra | Identified | Spectra | Identified | Spectra | Identified |
| 45 min Control | 1363 | 314 | 967 | 172 | 776 | 123 | 887 | 123 | 749 | 73 | 6044 | 3480 |
| 45 min Cr | 1357 | 406 | 790 | 136 | 669 | 116 | 719 | 95 | 647 | 73 | 4371 | 2352 |
| 90 min Control | 1558 | 456 | 885 | 159 | 682 | 111 | 716 | 78 | 628 | 74 | 7388 | 4770 |
| 90 min Cr | 1059 | 248 | 808 | 140 | 628 | 89 | 680 | 74 | 580 | 70 | 3549 | 1896 |

resulting in both a decrease in peptide identifications and an increase in incorrect

identifications.  A subsequent version of Sequest will only allow 3 differential

modifications to be specified, while DBDigger allows an unlimited number of PTMs to

be specified in the search parameters.  This limitation leads to reduced identifications,

since different modifications can occur on the same peptide and result in a total of 32

separate searches performed on the datasets.  This is a significant issue with the oxidation

dataset (total of 7 modifications on 4 amino acids), where DBDigger completed the

search in one round while Sequest required 3 separate rounds.  Therefore, the Sequest

results were noted and can be found in Supplemental Tables S3-S6 but were not used for

validation.

*InsPecT scoring optimization for post-translationally modified peptides*

Due to difficulties with Sequest and the high FDR of DBDigger when considering

PTM-containing peptides only, the decision was made to search the datasets using the

algorithm InsPecT [40].  Compared to DBDigger and Sequest, InsPecT uses a different

approach to searching MS/MS for PTMs.  The first stage in filtering for InsPecT is to

create a set of tags (25 were specified in the searches performed here) three amino acids

in length using fragment ions in the experimental MS/MS.  This enables a relatively short

list of peptides to be searched from the forward and reverse sequence database used here.

The resulting candidate peptides are scored based on seven different criteria:  the number

of predicted (1) b and (2) y ions that match to the MS/MS, how well the intensity of the

identified (3) b and (4) y ions match the predicted intensity, (5) trypsin specificity, (6) the

length of the candidate matching peptide where the presence of PTM(s) indicates a

shorter peptide, and (7) the fragment ion profile of the spectrum (presence of an isotope,

143

higher fragment ion intensity in the middle of the spectrum, and properties of the neighboring residue) [40]. The resulting top score (MQScore) is then compared to the distribution of the lower scores (DeltaScore) and a p-value is calculated. The p-value used by InsPecT is based on the p-value devised for the algorithm Peptide Prophet [174]. A lower p-value and higher match score are indicative of a better match, which is in contrast to the limited evaluative parameters employed by the other two algorithms used in this study.

To determine filtering parameters that would give at least an equivalent if not lower FDR for modified peptides, various filtering options were applied to the 45 min Control Run1 dataset. The default parameters for InsPecT are a p-value of 0.1 and the requirement of at least one peptide for protein identification. Using the default parameters yields a FDR of 0.7% (125 reverse peptides and 37,120 forward peptides). Therefore, the p-value was raised to 0.35, giving a FDR of 2.0% (428 reverse peptides and 42,200 forward peptides). We then evaluated a p-value giving a 2% FDR and requiring two peptides for positive protein identification. A p-value of 0.75 was chosen for the two-peptide dataset since this gave a FDR of 2.1% (499 reverse peptides and 46,943 forward peptides). A total of 97% of the proteins identified in the two-peptide dataset were identified in the one-peptide dataset as well. However, the one-peptide dataset had only 77% overlap with an additional 472 proteins identified with only one peptide. This observation raises an important issue of proteins identified solely by one peptide with the possibility of the identification coming from a modified peptide. The FDRs when considering only peptides containing a putative PTM was 11.9% (420 reverse peptides and 6668 forward peptides) for the one-peptide dataset and 10.8% (492

reverse peptides and 8634 forward peptides) for the two-peptide dataset. When comparing the reverse peptide identifications for the PTM peptides and the total reverse peptide identifications for the datasets, ~99% of the reverse identifications came from a PTM peptide. Therefore, all InsPecT datasets presented below had the following filtering criteria: two peptides required for protein identification and a p-value cutoff that provides a 2±0.1% FDR for the total peptide dataset. These criteria are more conservative and reduce the FDR of modified peptides by an average of 30% when compared to DBDigger's modified peptide FDR.

The result files from the InsPecT search were filtered differently according to the p-value distribution for that dataset. For the 45 min Control dataset, p-values of 0.75 (Run1) and 0.73 (Run2) were used as cutoff scores, because this value gave a FDR of 2.1% and 2.0%, respectively, for the total peptide dataset (both unmodified and modified peptide identifications) with 10.8% (Run1) and 9.1% (Run2) FDRs when considering only modified peptide identifications. Table 5.5 depicts the FDRs and number of peptides (total and modified only) identified for the remaining datasets evaluated in this study. All PTM peptides identified by InsPecT are located in Supplemental Tables S7 and S8. The reason for choosing a p-value that reflects a FDR value instead of one p-value to be applied across all the datasets is due to the large variance in FDRs across the datasets for a given p-value. For instance, if a p-value of 0.75 were chosen as the cutoff for all datasets, FDRs would range from 1.8-4.0% with an average of 2.3%. This wide range skews the identifications between datasets and leads to some having a great deal more false or true identifications than others. This variance was only pronounced in one dataset from the DBDigger searches (90 min Control oxidation versus 90 min Cr

145

**Table 5.5.** FDRs for Total and PTM Peptide Identifications with InsPecT

| Dataset | p-value | FDR[a] Total Peptide Identifications | | | FDR PTM Peptide Identifications | | |
|---|---|---|---|---|---|---|---|
| | | Reverse IDs | Total IDs | % FDR | Reverse IDs | Total IDs | % FDR |
| 45 min Control Run 1 | 0.75 | 499 | 47442 | 2.1 | 492 | 9126 | 10.8 |
| 45 min Control Run 2 | 0.73 | 921 | 91769 | 2.0 | 857 | 18888 | 9.1 |
| 45 min Cr Run 1 | 0.75 | 458 | 48057 | 1.9 | 449 | 8928 | 10.0 |
| 45 min Cr Run 2 | 0.75 | 444 | 46475 | 1.9 | 378 | 8771 | 8.6 |
| 90 min Control Run 1 | 0.75 | 482 | 48375 | 2.0 | 476 | 8671 | 10.4 |
| 90 min Control Run 2 | 0.55 | 585 | 56904 | 2.1 | 530 | 18040 | 5.7 |
| 90 min Cr Run 1 | 0.70 | 454 | 44784 | 2.0 | 444 | 9053 | 9.4 |
| 90 min Cr Run 2 | 0.80 | 440 | 45736 | 1.9 | 381 | 7466 | 9.7 |

[a]FDR: false discovery rate

oxidation); and in general, DBDigger and Sequest do not have the FDR variation between datasets that was observed with InsPecT. This variance is most likely due to differences in the p-value distribution in each dataset, which is in contrast to the charge-state dependant scoring of the other two algorithms.

*Post-translational modification results with DBDigger and InsPecT*

To illustrate the benefit of performing a search for PTMs, all modified proteins identified were compared according to their functional category. Annotated genes of the *S. oneidensis* MR-1 genome were assigned to one of 18 functional categories by The Institute for Genomic Research (now the J. Craig Venter Institute). Figure 5.2 depicts the functional distribution of the modified peptides identified using DBDigger and InsPecT according to the role categories. As a comparison to the previously reported unmodified dataset [147], the categories of hypothetical proteins, energy metabolism, unknown function, cellular processes, and protein fate were the top five functional categories representing at least 55% of the unmodified protein identifications. By contrast, the top five modified categories based on DBDigger analysis (Figure 5.2A) consisted of hypothetical proteins, energy metabolism, protein synthesis, protein fate, and cellular processes. The category of protein synthesis is composed primarily of genes encoding components of the ribosome, which has been shown previously to contain a large number of modifications on the constituent proteins [48]. The cellular processes category also comprises a large number of proteins with putative PTMs (average of 570 PTM peptides) and some of these proteins are involved in the chemotactic response of *S. oneidensis* to environmental cues. The identification of a number of modified proteins involved in chemotaxis is discussed below. The functional category distribution of PTM peptides

147
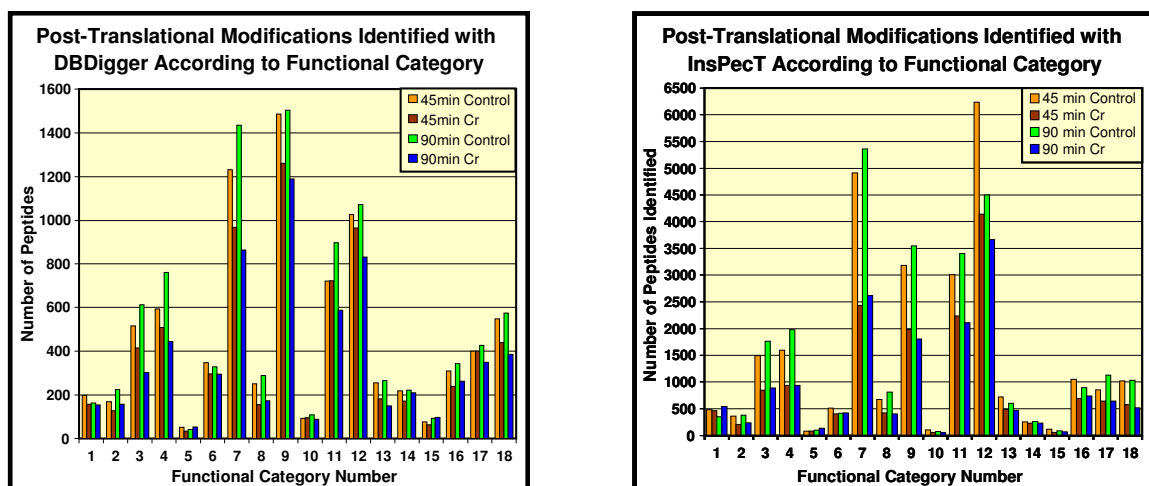
**Figure 5.2.** Identification of PTM peptides using (A) DBDigger and (B) InsPecT.
Assignment of a peptide into a category was according to the parent protein's
membership. The functional category assignments are as follows: (1) Amino acid
biosynthesis; (2) Biosynthesis of cofactors, prosthetic groups, and carriers; (3) Cell
envelope; (4) Cellular processes; (5) Central intermediary metabolism; (6) DNA
metabolism; (7) Energy metabolism; (8) Fatty acid and phospholipid metabolism; (9)
Hypothetical proteins; (10) Mobile and extrachromosomal element functions; (11)
Protein fate; (12) Protein synthesis; (13) Purines, pyrimidines, nucleosides, and
nucleotides; (14) Regulatory functions; (15) Signal transduction; (16) Transcription; (17)
Transport and binding proteins; and (18) Unknown function.

148

based upon InsPecT analysis was similar to that observed with DBDigger. The predominant categories containing PTM peptides were protein synthesis (~4600), energy metabolism (~3800), protein fate (~2700), hypothetical proteins (~2600), and cellular processes with ~1300 peptides on average (Figure 5.2B). Some functional categories demonstrate a putative bias in modifications in the control versus chromate-shocked datasets (energy metabolism and hypothetical proteins). At this time, however, the detailed understanding of this putative difference in control versus chromate-shocked PTM peptides is not known and is beyond the scope of the current work.

A majority of the proteins comprising the functional categories of cellular processes and signal transduction are annotated as part of signaling cascades first activated by a PTM on the initiating protein, which responds to various environmental stimuli. Signal transduction proteins are generally of low abundance, and identifications are usually confirmed by observing approximately two peptides per protein. In a search that specifies no modifications on the proteins, less than 0.2% of the total spectra count are attributed to signal transduction proteins; where spectral count was used as a measure of protein abundance previously [182, 183, 237]. Out of the 61 proteins predicted to be involved in signal transduction in the *S. oneidensis* proteome, 53 of the proteins are annotated as members of two-component regulatory systems. Two-component systems contain a sensor histidine kinase, which autophosphorylates in response to a specific environmental cue and then transfers the high energy phosphate group to a cognate response regulator, which effects a change in gene expression upon phosphorylation [232, 238, 239]. In our previous study [147], 39 members of two-component systems were identified without specifying PTMs in the search. By including the option of

methylation, acetylation, or oxidation to seven different amino acids, there was an increase in the rate of identification for some members of this category using identifications from both DBDigger and InsPecT. DBDigger was able to identify a total of 269 peptides containing at least one PTM from 41 proteins. The signal transduction functional category yielded 47 protein identifications with at least two peptides using InsPecT. A total of 253 unique PTM peptides annotated as involved in signal transduction were identified in the InsPecT dataset, with 101 of these peptides containing at least one methyl group. Both algorithms identified more proteins than in our previous study solely by including PTMs in the search parameters. All but two of the proteins identified previously [147] were identified by either DBDigger or InsPecT. Both algorithms identified 13 proteins as being post-translationally modified in this study. Interestingly, the control datasets yielded these shared peptides, while the chromate-exposed cultures did not yield any shared modified peptides.

Methylation has been implicated previously in changing DNA-protein interactions [238, 239]. The signal transduction group of proteins in bacteria may be phosphorylated in the activation domain, while DNA interaction occurs in the other domain [52]. The methylation could either perturb the domain preventing interaction with the phosphate backbone of the DNA or the phosphorylation event could cause a conformational change that leads to the methylation event. The signal transduction proteins identified above were cross-referenced between DBDigger and InsPecT. SO2541, a response regulator, was identified with various PTMs including methylation of various degrees up to trimethylation on multiple residues by both algorithms, but was identified in the Control dataset with InsPecT whereas DBDigger identified this protein in the Cr dataset for the

150

45 min time-point.  While DBDigger identified SO3196 and SO3688 in the 90 min

dataset, InsPecT identified them in the 45 min Control dataset.  Both algorithms

confirmed only one protein, SO2544, from the 90 min dataset.  However, each algorithm

identified a different MS/MS and peptide.  This is not unexpected in complex proteome

datasets, and illustrates the need to validate MS/MS from PTM identifications (especially

since many modified peptides are of low abundance).

Chemotaxis, part of the cellular processes functional category, plays an important

role in enabling bacteria to identify and respond to small molecules (nutrients or toxins)

in the surrounding environment.  The end result is a mechanical response that moves the

organism either toward or away from the molecule concentration gradient.  Bacterial

chemotaxis has been well-studied and largely elucidated mechanistically in the model

bacterium *E. coli* [53, 54, 224, 240, 241].  In *E. coli*, methyl-accepting chemotaxis

protein receptors (chemoreceptors) are encoded by five genes that detect amino acids

such as serine (*tsr*) and aspartate (*tar*); dipeptides (*tap*); ribose and galactose (*trg*); and

the redox potential (*aer*) [224] of the surrounding environment.  The level of stimulation

of a particular chemoreceptor is controlled by two proteins CheR, a methyltransferase,

and CheB, a methylesterase.  Therefore, the presence of a methyl group on the

chemoreceptor is indicative of stimulation due to the binding of a specific molecule, this

initiates a signaling cascade through the autophosphorylation of CheA and results in the

activation of FliM (the protein subunit comprising the flagellar motor).  Activation of

FliM results in the bacterium swimming in a tumbling motion either toward or away from

the chemical gradient.  Methylation of the chemoreceptors is important in controlling the

level of stimulation so that the bacterium is not constantly tumbling [54, 224].

Chemotaxis is not only indicative of stress (i.e. from chromate shock as described here), but rather is a response to the surrounding environment, so the presence of methylation on chemoreceptors from both the control and chromate-shocked datasets [88, 92, 242-244] is not surprising.

The chemotaxis system of *S. oneidensis* is more complicated and includes pathways responding to a number of anaerobic electron acceptors [88, 92], three separate signal cascades initiated by CheA, and 29 chemoreceptors. The *E. coli* and *S. oneidensis* chemoreceptors were aligned and a dendogram (or sequence tree) was constructed to assess sequence similarity between *S. oneidensis* and *E. coli* chemoreceptors. Figure 5.3 depicts the relationship between the five chemoreceptors from *E. coli* along with 29 from *S. oneidensis.* As shown, four of the *E. coli* chemoreceptors cluster tightly together in the middle of the tree and do not demonstrate sequence similarity with chemoreceptors from *S. oneidensis.* Aer, which responds to redox potential, shows sequence similarity with three *S. oneidensis* chemoreceptors (SO1385, SO0584, and SO3404). The lack of sequence similarity between the chemoreceptors in the two organisms is expected due to the differences of the natural habitats in which the bacteria are found. *S. oneidensis* is a fresh water microbe, which must be ready to adapt to rapidly changing concentrations in nutrients and metals [88, 92, 242, 243]. A large number of chemoreceptors were identified with and without PTM peptides in the 45 and 90 min datasets both under control and chromate-shocked growth conditions (Table 5.6). DBDigger was able to identify all but one (SO3510) of the chemoreceptors predicted. While InsPecT identified SO3510 in the 90 min Control dataset, this algorithm did not identify a couple of low abundant receptors identified by DBDigger (SO1434 and SO2117).

152

**Figure 5.3.** A sequence tree (or dendogram) depicting sequence similarity.

The sequence similarity is between *E. coli* and *S. oneidensis* chemoreceptors to determine

if there is any sequence conservation between the two organisms. MotA (motility protein

A) and MotB (motility protein B) from both organisms served as outgroups for the

alignment. Circled are the four *E. coli* chemoreceptors that are not recognizably

conserved with *S. oneidensis*. The protein sequences were imported into and aligned

using Clustal X (version 1.8, [245]). A bootstrap Neighbor-Joining tree was created

using Clustal X (bootstrap value of 10,000). The resulting dendogram was visualized

with TreeView (version 1.6.6, taxonomy.zoology.gla.ac.uk/rod/treeview.html).

**Table 5.6:** Identification of Chemoreceptors using DBDigger and InsPecT from the 45 and 90 min Datasets

| Locus | DBDigger Identifications | | | | InsPecT Identifications | | | |
|---|---|---|---|---|---|---|---|---|
| | 45 min Control | 45 min Cr | 90 min Control | 90 min Cr | 45 min Control | 45 min Cr | 90 min Control | 90 min Cr |
| SO0500 | 2[a] | 4 | 2 | 3 | 3 | 4 | 2 | 3 |
| SO0584 | 2 | 3 | 4 (2)[b] | 2 | 4 (1) | 2 (1) | 5 (1) | 0 |
| SO0987 | 32 | 24 (2) | 39 (2) | 28 (1) | 29 (2) | 20 (2) | 36 (1) | 25 (2) |
| SO1056 | 44 (1) | 37 (1) | 45 (5) | 37 (1) | 57 | 35 (1) | 74 (2) | 26 |
| SO1144 | 31 (2) | 29 | 24 (2) | 24 | 19 | 18 | 19 (2) | 15 |
| SO1278 | 5 | 10 | 11 (1) | 9 | 2 | 6 | 3 | 3 |
| SO1385 | 11 | 13 (2) | 8 | 12 | 9 (1) | 7 | 5 | 8 |
| SO1434 | 4 | 3 | 4 | 4 | 0 | 0 | 0 | 0 |
| SO2083 | 7 | 8 | 7 | 8 | 6 | 2 | 3 | 2 |
| SO2117 | 4 | 3 | 4 | 4 | 0 | 0 | 0 | 0 |
| SO2123 | 0 | 3 | 2 | 0 | 6 (1) | 4 (1) | 4 (1) | 2 |
| SO2240 | 17 (2) | 20 (2) | 25 (1) | 17 | 19 (1) | 17 | 27 | 9 |
| SO2317 | 2 | 2 | 3 | 3 | 0 | 2 (1) | 3 (1) | 5 (2) |
| SO2323 | 11 | 11 | 6 | 6 | 8 | 13 | 7 | 5 |
| SO3052 | 13 (1) | 14 | 13 | 21 | 8 | 8 | 7 | 16 |
| SO3207 | 38 | 36 | 40 (1) | 39 | 43 (3) | 27 | 27 | 33 |
| SO3282 | 16 | 8 | 25 | 14 | 14 (1) | 0 | 24 | 6 |
| SO3396 | 4 | 4 | 4 | 5 (2) | 2 (2) | 3 (1) | 2 (1) | 0 |
| SO3404 | 4 | 2 | 5 | 3 | 0 | 3 | 2 | 0 |
| SO3510 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| SO3582 | 17 (1) | 15 (1) | 14 | 20 (2) | 9 | 8 (2) | 8 | 6 |
| SO3642 | 43 (1) | 38 (2) | 38 (2) | 42 (3) | 45 (3) | 34 (2) | 41 (2) | 35 (2) |
| SO3838 | 21 (1) | 32 (1) | 24 (1) | 25 (2) | 19 | 25 (1) | 31 (1) | 15 (1) |
| SO3890 | 7 (1) | 7 | 7 | 9 | 8 | 7 | 4 | 9 (1) |
| SO4053 | 20 | 11 | 24 | 12 | 12 | 6 | 25 | 9 |
| SO4454 | 22 (2) | 24 (1) | 24 | 24 (1) | 12 | 15 (1) | 22 (1) | 14 |

**Table 5.6.** Continued.

| Locus | DBDigger Identifications | | | | InsPecT Identifications | | | |
|---|---|---|---|---|---|---|---|---|
| | 45 min Control | 45 min Cr | 90 min Control | 90 min Cr | 45 min Control | 45 min Cr | 90 min Control | 90 min Cr |
| SO4466 | 5 (1) | 2 | 3 | 4 | 2 | 0 | 4 | 0 |
| SO4557 | 30 | 37 | 30 | 33 | 40 (1) | 37 (1) | 52 (4) | 31 (2) |
| SO4635 | 4 | 5 | 4 | 4 | 0 | 0 | 2 | 2 |
| SOA0106 | 36 (2) | 39 (2) | 37 (3) | 34 | 42 (2) | 48 (3) | 61 (1) | 33 (2) |

[a]Number of unique peptides both unmodified and containing a PTM identified for a given chemoreceptor. [b]Number of monomethylated glutamate residues found for a given chemoreceptor.

The *S. oneidensis* chemoreceptor SO3642 was identified consistently as having a methylated glutamate residue at the C-terminus of the protein. A total of 9 MS/MS were identified using InsPecT and 16 were identified using DBDigger (Table 5.6). An example MS/MS found in common between the two algorithms is depicted in Figure 5.4. The mass shift due to the addition of 14 Da is readily present and causes the fragment ions with the glutamate residue to be 14 Da heavier. Using InsPecT, two spectra were identified and scored very well (p-values of $< 4e^{-5}$) in the 45 min Control dataset. DBDigger did not score these MS/MS high enough to pass the filter thresholds. In addition, InsPecT did not identify any modified MS/MS in the 45 min Cr datasets, but DBDigger found 3 spectra. The 90 min Control Run 2 dataset was more consistent with 4 MS/MS found by InsPecT and 9 by DBDigger. Overall, the 90 min Cr dataset was the most consistent as both algorithms identified the MS/MS depicted in Figure 5.4. As mentioned, chemoreceptors are a well-studied class of proteins, and information in the literature [54, 224, 246] was used to help confirm the physiological significance of this modification. The SO3642 peptide was identified solely in the membrane fraction of cell lysates, which is consistent with a cytoplasmic membrane location of known chemoreceptors [224, 240]. Chemoreceptors have also been shown to be methylated on conserved C-terminal glutamate residues, as is the case with SO3642 demonstrated here [224, 240]. Therefore, based on our MS data and the extant literature, we conclude that SO3642 is methylated on residue E515. However, the implication of this modification in terms of a response to chromate is not known at this time and was not evaluated in detail within the scope of this work. While the present study delineates the computational data mining approach for identifying the range of protein post-translational modifications at
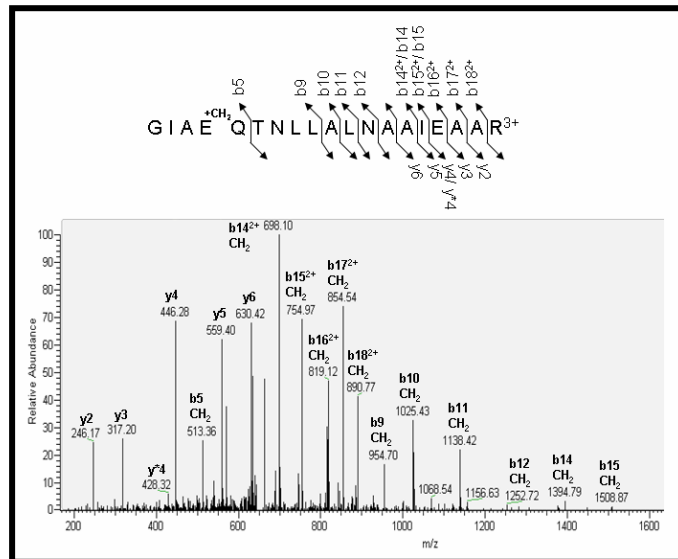
**Figure 5.4.** A MS/MS for SO3642, a chemoreceptor, confirms a C-terminal peptide. The C-terminal peptide contains a monomethylated glutamate residue. All major peaks of the MS/MS are labeled, and 12 fragment ions demonstrate a mass shift of 14 Da corresponding to the addition of a methyl group. The MS/MS scored high for both algorithms: DBDigger score of 75.73, DelCN of 0.46 and InsPecT MQScore 3.268, DeltaScore 0.65, p-value 0.151.

the global level, thereby providing detailed information about the types of proteins modified and the range of modifications, future detailed studies will be required to fully unravel the biological details of PTM correlation with chromate exposure for *S. oneidensis*.

**<u>Conclusions</u>**

InsPecT and DBDigger are complementary algorithms that both identified a large number of putative post-translationally modified peptides. However, the FDR for DBDigger was more than twice that for PTM peptides, and its search capabilities for large datasets was reduced, thereby limiting combinatorial PTMs. Therefore, InsPecT was tested and resulted in a much lower FDR for PTM peptides. If an identification resulting from the chemical addition of a methyl, acetyl, or hydroxyl group was made by both DBDigger and InsPecT for a particular peptide, this was taken as a positive confirmation of that modification. The PTM peptide identifications between DBDigger and InsPecT were compared in this study. For the 45 min Control Run1 dataset, a total of 3219 modified peptides were shared between the two algorithms (Figure 5.5). This corresponds to 69% of the peptides identified using DBDigger (a total of 4651 PTM peptides) and 37% from InsPecT. The InsPecT overlap is much smaller due to the scoring scheme of the algorithm: InsPecT allows high scoring MS/MS from non-tryptic and semi-tryptic peptides to pass the filtering thresholds. The option of specifying a non-tryptic cleavage was not used with DBDigger due to the increased time and search space, and thus higher FDRs. Removing the non-tryptic and semi-tryptic identifications from the InsPecT dataset yields 4005 fewer peptides; this brings the overlap between

158
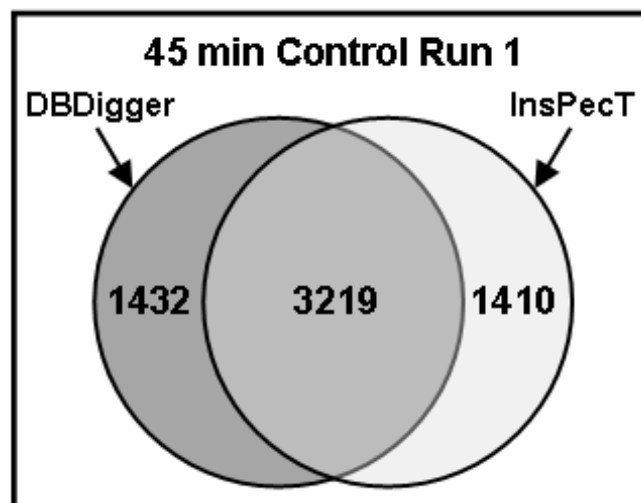
**Figure 5.5.** A Venn diagram comparing DBDigger and InsPecT PTM peptides.

The results represent the 45min Control Run1 dataset. A total of 3219 peptide

identifications were shared between the two algorithms. DBDigger identified a further

1432 unique tryptic peptides. InsPecT identified 3996 semi-tryptic, 9 non-tryptic, and a

further 1410 unique tryptic peptide identifications.

DBDigger and InsPecT to 70%.  Comparable results were found for the remaining datasets (Supplemental Figures S1-S7, pubs.acs.org/journals/jprobs/index.html).

Even though an average of 3000 post-translationally modified peptides for each dataset were shared between the two algorithms, there is still room for improvement in performing PTM searches.  DBDigger and InsPecT search the MS/MS very differently.  DBDigger uses the MASPIC [173] scoring algorithm to determine peptide identifications for the MS/MS.  MASPIC separates the MS/MS spectra into different intensity classes where the most intense class is restricted to a few fragment ions.  The score is weighted based on identification of intense peaks.  InsPecT is similar to MASPIC in the respect that there is a scoring factor for intensity.  However, InsPecT first discriminates whether an intense peak is identified as a b or y ion and then weights that intensity into the score, where the y ions derived from the lower energy CID as used in this study are more intense than their counterpart b ions.  This is more discriminatory than MASPIC and leads to more stringent filtering of false identifications.  The more stringent scoring of InsPecT is the reason for the 30% reduction in false identifications of modified peptides.  Another advantage to using InsPecT is the incredible speed with which the algorithm searches shotgun proteomic data.  InsPecT was able to search over 240,000 MS/MS in approximately 84 hours specifying all PTMs in one search on a desktop computer for each dataset, while DBDigger took twice as long with seven different searches for each dataset.  The authors initially attempted to specify all methylations in one search, but due to the combinatorial increase in search space the computational memory was overwhelmed.  The ability to search for all modifications at one time is physiologically relevant, since a peptide can be methylated and oxidized at the same time.  However,

160

using both algorithms together increases the confidence and accuracy of identifications with the overlap between them composed of the top-scoring PTM peptides, which was a major goal of this work.

As evident from other studies, the use of a higher performance mass spectrometer will undoubtedly provide more confident identifications of peptide PTMs, since improved mass accuracy will reduce ambiguities.  However, at present, many readily available algorithms, such as the three tested in this study, do not contain a scoring element which exploits mass accuracy of the parent peptide or resulting fragment ions as an additional metric.  Since medium resolution mass spectrometers are common in many research laboratories, and since many existing bioinformatic approaches do not utilize higher mass accuracies, we chose to focus on evaluating methodologies which would be readily available to the current field.  There are certainly efforts underway in many laboratories to incorporate high performance mass spectral datasets from Orbitrap, FTICRMS, and QTOF platforms into data mining schema.  This is already providing increased confidence in identification of peptide modifications, including the ability to discriminate between phosphorylation (79.980) and sulfonation (80.064) as well as trimethylation (42.03 Da) versus monoacetylation (42.01 Da).

# Chapter 6

## Proteomic Comparison of *Shewanella oneidensis* MR-1 Wild-Type and a Response Regulator Deletion Strain under Conditions of Chromate Transformation

## Introduction

Our previous work with Cr(VI) exposure and *S. oneidensis* MR-1 identified over 2400 proteins expressed from the MR-1 genome [147, 156, 158], with a subset demonstrating differential expression in response to Cr(VI). Many of the differentially expressed proteins had annotated functions in metal ion transport and sulfate transport/metabolism. Interestingly, a putative response regulator, SO2426, was also identified as being up-regulated under Cr(VI) conditions (5, 39). Response regulators are part of two-component signal transduction systems that serve as a basic stimulus-response coupling mechanism, allowing organisms to sense and respond to changes in many different environmental conditions [238]. For example, the redox sensing system pathway involves the ArcAB complex in *Escherichia coli* [247]. In fact, due to the extensive respiratory versatility in *S. oneidensis*, Gralnick *et al* [247] hypothesized that this bacterium must have an extensive regulatory system.

In our acute shock study [147], *S. oneidensis* was exposed to a sub-lethal concentration of 1 mM Cr(VI) for 45 and 90 min. When compared to the control cultures (no Cr(VI) added), SO2426 was highly up-regulated. The corresponding transcript was found to follow a trend of increasing expression level with increased exposure time [147]. A Cr(VI) dosage study [156] also found SO2426 to be up-regulated under three different sub-lethal concentrations (0.3, 0.5, and 1 mM) of Cr(VI) exposure for 30 min. Therefore, SO2426 appears to be activated under growth conditions where Cr(VI) is present, irrespective of the dosage. In another study, *so2426* was found to also be up-regulated after exposure to acute Strontium stress [248]. These observations led to the conclusion that SO2426 may be a response regulator for transition metal redox state. This prompted further investigation of the role SO2426 may be playing in response to sub-lethal concentrations of Cr(VI) by "knocking out" the gene through an in-frame deletion, as described by Chourey *et al* [157]. The Δ2426 mutant strain demonstrates sensitivity to other transition metals besides Cr(VI) (Chourey *et al*, unpublished results).

The use of shotgun proteomics of biological replicates to probe the functionality of global regulator proteins provides information not found with other methodologies. Also, the resulting variability between cultures consisting of a gene deletion mutant has not been explored in great detail. Information about gene function can be probed utilizing shotgun proteomics, thus allowing for increased understanding of the global response [249] and possible compensatory pathways encoded in the bacterium's genome. Most shotgun proteomics studies have been performed using technical replicates [19, 20, 147, 211, 250] instead of biological replicates. A technical replicate is defined as multiple LC-MS experiments on one sample (i.e. a cell culture, tissue sample, etc.) [251].

163

However, recently there has been a number of shotgun proteomics studies [210, 252, 253] and a review article [251] on both the informational value and necessity of biological replicates for quantitation of the measured proteins. Most of the focus for the discussion between the use of technical replicates or biological replicates has been on the resulting statistical treatment of the dataset [182, 183, 251, 253]. The primary argument for the use of biological replicates is to reduce the background biological noise of the sample [251]. Available nutrients, temperature fluctuations, and light availability (relevant for phototropic bacteria) are factors considered biological noise and may play a significant role in the growth variability of a culture. The resulting protein complement of the culture may vary based on these factors.

The possible role of SO2426 as a global regulator of Cr(VI) response pathways is explored further through the comparison of a Δ2426 mutant to the wild-type strain of MR-1. In this study, differentially expressed proteins solely from biological replicates of wild-type and a Δ2426 mutant in response to a sub-lethal concentration of Cr(VI) are identified. Differential regulation of protein function through the identification of post-translational modification sites on proteins present in the global dataset is also investigated.

## **Materials and Methods**

*Chemical reagents and culture growth*

Luria-Bertani (LB) medium, Guanidine HCl, Tris, EDTA, Ammonium Acetate, and $CaCl_2$ were purchased from Sigma Chemical Co. (St. Louis, MO) and were used without further purification. Modified sequencing grade trypsin was purchased from Promega (Madison, WI) and used for all protein digestions. The 99% formic acid was

obtained from EM Science (Darmstadt, Germany) and HPLC-grade water and acetonitrile from Burdick & Jackson (Muskegon, MI).

Three independent cultures of wild-type *S. oneidensis* MR-1 and three independent cultures of strain Δ2426 (obtained from D. Thompson, Purdue University) were grown in a 250 mL flask with 100 mL of LB medium (pH 7.2) at 30°C under aerobic conditions with constant agitation at 200 rpm. When the OD at 600 nm reached 0.5 (mid-log phase), a 10-ml aliquot of each culture was taken as the 0 h time point and processed as described in Brown *et al* [147]. Afterwards, Cr(VI) in the form of $K_2CrO_4$ was added to yield a final concentration of 0.3 mM in all six cultures. Cr(VI) concentration in the media was monitored using the 1,5-diphenylcarbazide (DPC) method as described previously [138]. When the approximate concentration of Cr(VI) in the media of the control cultures (wild-type MR-1 strain) reached 0.25 (1 h), 0.15 (3 h), and 0 (4 h) mM, 10 mL aliquots of culture were harvested. *Cell lysis and digestion*

The resulting frozen cell pellets from the 10 mL aliquots were weighed and then further distributed into 5 mg wet cell pellets for lysis. The lysis protocol used is essentially as described in Thompson *et al* [160]. Briefly, 6 M Guanidine with 10 mM DTT were added to each sample and incubated at 37°C overnight after which the guanidine was diluted 6-fold with 50 mM Tris/ 10 mM $CaCl_2$. Then, 5 μg of trypsin was added, followed by a 6 h incubation period after which another aliquot of trypsin was added with an overnight incubation at 37°C. A final reduction step for 2 h was performed with 20 mM DTT. Protein digests were then spun at 10,000 x g for 10 min to remove cellular debris and then stored at -80 °C until the LC/LC-MS/MS experiment.

*Global proteome LC/LC-MS/MS analysis*

The LC/LC-MS/MS experiments were performed as described in Brown *et al* [147]. Each lysis/digestion from each culture was a separate 24 h LC/LC-MS/MS experiment. Samples were loaded onto a split phase column consisting of reverse phase (C18) and strong cation exchange (SCX) separation materials. This column was placed behind a 15 cm C18 analytical column located directly in front of the mass spectrometer (LTQ, Thermo Scientific, San Jose, CA). The LTQ was coupled with an Ultimate HPLC pump (LC Packings, a division of Dionex, San Francisco, CA). The samples were analyzed with a 12-step 2D HPLC analysis by adding increasing concentrations (0 to 500 mM) of ammonium acetate salt pulses followed by an aqueous (95% $H_2O$, 5% ACN, 0.1% formic acid) to organic (30% $H_2O$, 70% ACN, 0.1% formic acid) gradient. The LTQ was operated in the data-dependent mode during the chromatographic separations as detailed in Brown *et al* [147].

*Proteome bioinformatics*

The protein database used for all MS searches consisted of Version 8 (www.tigr.org) of the *S. oneidensis* MR-1 proteome (4,798 proteins) as well as 36 common contaminant sequences (trypsin, keratin, etc.). The database is available for download at the project website compbio.ornl.gov/shewanella_metal_stress/databases/. Initially, Sequest was used to search the resulting MS raw files for peptide/protein identifications. Sequest searches were performed as detailed in Brown *et al* [147] with the following scoring cutoffs. A minimum of 2 peptides were required for a positive identification of a protein with peptide charge state score minimums of 1.8 (+1), 2.5 (+2), and 3.5 (+3). A delCN value of 0.08 was required for peptide identifications. Sequest

results were then filtered and sorted according to the above criteria with DTASelect and Contrast [171], and all resulting data is available on the project website (compbio.ornl.gov/ shewanella_metal_stress/reduction).

For performing searches for post-translational modifications, the search algorithm InsPecT [40] was used. For optimization of PTM searches with InsPecT on shotgun proteomics data, see reference [59]. The proteome database used in these searches consisted of the forward database (mentioned above) with the reversed sequences of all proteins concatenated to the end. This database was used in order to estimate the false discovery rate of each individual search. MS/MS spectra were extracted from the raw files using ReAdW.exe to create a .mzXML file. The mzXML files were then searched with InsPecT [40] specifying the following PTMs. An optional mass of 14 Da (monomethylation) was added to lysines, arginines, and glutamates; optional masses of 28 (dimethylation), 42 (trimethylation/monoacetylation), 84 (diacetylation) Da were added to lysines and arginines; 16 Da (monooxidation) to methionines, cysteines, tryrosines, and tryptophans; 32 Da (dioxidation) to methionines and cysteines; and 48 Da (trioxidation) on cysteines. The scripts Summary.py and Pvalue.py were used to filter the resulting output tab-delimited text files by p-value and peptide count. A peptide count of two was required for all proteins, and the p-value was adjusted to give a false discovery rate of 2±0.2% for the total peptide dataset.

## Results and Discussion

The gene *so2426* encodes a putative DNA-binding response regulator and may regulate the transcriptional activation of genes involved in response to Cr(VI) in the bacterium. *so2426* contains two putative domains, a CheY-like domain and a putative

167

DNA-binding domain. The CheY-like domain is ~50% similar to the other three CheY

proteins predicted in the *S. oneidensis* genome. However, the similarity extends only to

residue 131, leaving 106 residues not aligning to CheY. The 106 remaining residues are

on the C-terminal end and constitute a putative DNA-binding domain. SO2426 is also

highly conserved at the amino acid level to a putative ortholog from fifteen other

*Shewanella* genomes sequenced to date. The C-terminal DNA-binding domain portion of

the protein is annotated as a member of the Trans_reg_C family (PF00486) as described

using the Pfam database [254]. The top-aligning proteins from the other *Shewanella*

genomes are also annotated as part of the Trans_reg_c family. Figure 6.1 is a dendogram

(or sequence tree) tree depicting SO2426 clustering with a protein from the Sargasso Sea

*Shewanella* genome [74]. CheY from *S. oneidensis* and *Escherichia coli* were used as an

outgroup, further demonstrating SO2426 is not CheY, but rather a fusion protein with a

CheY-like domain, that is conserved across the genus *Shewanella*. Even though the top

blast hit from *Geobacter metallireducins* GS-15 did not cluster with these *Shewanella*

proteins, this protein is also identified in the same family with SO2426.

*Microbial growth during Cr(VI) transformation*

   *S. oneidensis* strains MR-1 and Δ2426 were cultivated aerobically in LB medium

(see Materials and Methods for details). The strains were cultured for 4 h in the presence

of Cr(VI) in the form of 0.3 mM $K_2CrO_4$. Figure 6.2 depicts the transformation of Cr(VI)

over the exposure period. The arrows point to the time points of harvest for global

proteome analysis. The standard error bars for the strains are shown for triplicate

cultures and the abiotic control, which demonstrated undetectable transformation of

Cr(VI) during the 4-h period. The standard deviation for the abiotic control ranged from
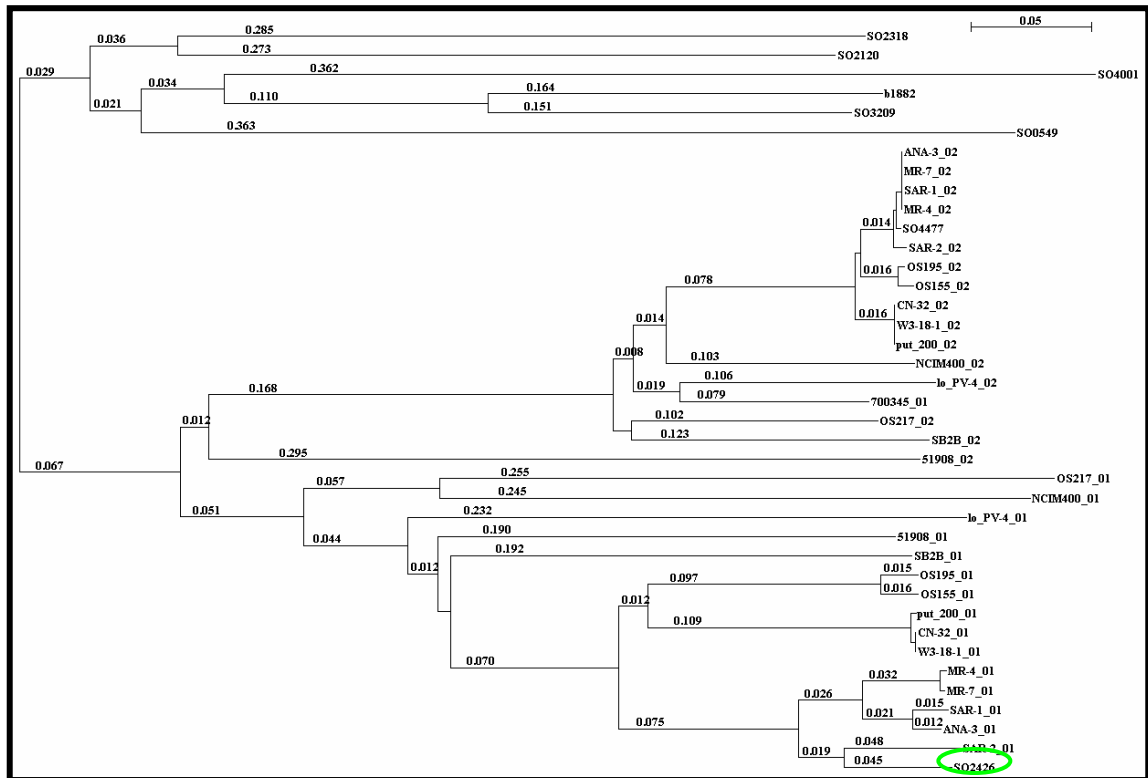
**Figure 6.1.** A sequence tree aligning the protein sequences of SO2426 and SO4477. SO2426 and SO4477 (CpxR transcriptional regulator) are aligned against the top Blast hit for each species/strain of *Shewanella* for genomes sequenced to date. SO2426 is at the bottom of the tree and most closely aligns with a protein sequence from a *Shewanella* strain from the Sargasso Sea[74]. The following proteins were used as an outgroup: from *S. oneidensis* MR-1 SO2318 (CheY-2); SO2120 (CheY-1); SO3209 (CheY-3); and SO0549 and SO4001 (chemotaxis protein CheY/response regulator receiver domain proteins) and b1882 (CheY) from *E. coli.* The species/strains of *Shewanella* are as follows: *S. oneidensis* MR-1 (SO4477), *S. putrefaciens* 200 (put_200_01 and put_200_02), *Shewanella* W3-18-1 (W3-18-1_01 and W3-18-1_02), *Shewanella* ANA-3 (ANA-3_01 and ANA-3_02), *S. woodyi* ATCC 51908 (51908_01 and 51908_02),

169

**Figure 6.1.**  Continued.

*S. frigidimarina* NCIMB400 (NCIM400_01 and NCIM400_02), *Shewanella* MR-4 (MR-4_01 and MR-4_02), *Shewanella* MR-7 (MR-7_01 and MR-7_02), *S. baltica* OS195 (OS195_01 and OS195_02), *S. denitrificans* OS217 (OS217_01 and OS217_02), *S. putrefaciens* CN-32 (CN-32_01 and CN-32_02), *Shewanella* SAR-1 environmental sequence (SAR-1_01 and SAR-1_02), *Shewanella* SAR-2 environmental sequence (SAR-2_01 and SAR-2_02), *S. loihica* PV-4 (lo_PV-4_01 and lo_PV-4_02), *S. baltica* OS155 (OS155_01 and OS155_02), *S. amazonensis* SB2B (SB2B_01 and SB2B_02), and *S. pealeana* ATCC 700345 (700345_01).  There was no significant conservation found between proteins from *Geobacter sulfurreducins* GS-15 and *S. oneidensis* MR-1 (data not shown).  Alignment and Neighbor-Joining tree (bootstrap value of 1000) was created by ClustalX [version 1.8, [245]].  The dendogram was visualized with njplot (pbil.univ-lyon1.fr/software/njplot.html).
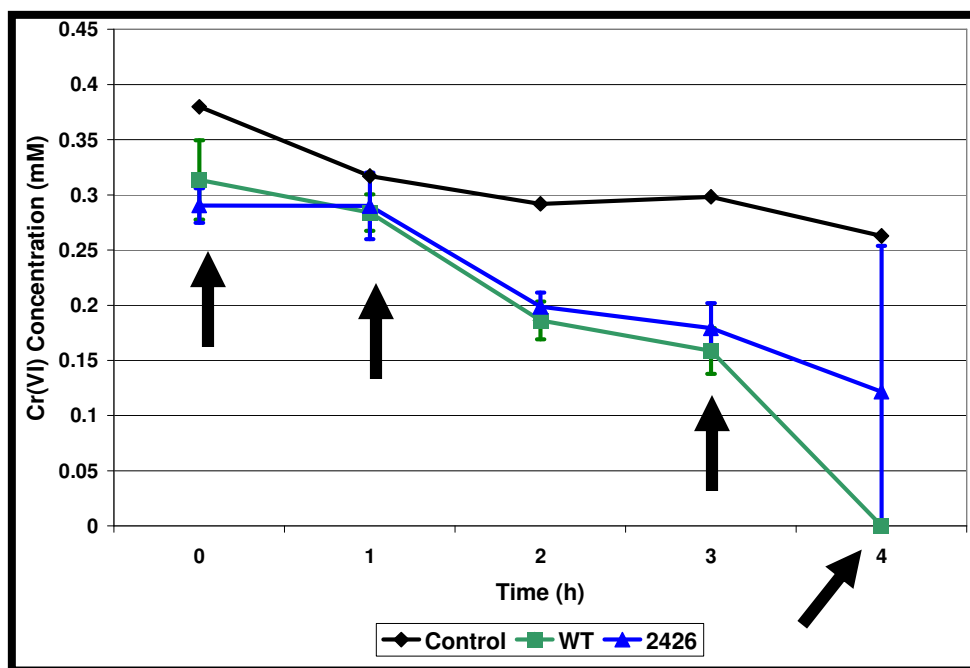
**Figure 6.2.** Transformation of 0.3 mM Cr(VI) in the form of $K_2CrO_4$.

The Cr(VI) transformation is depicted for strains MR-1 and Δ2426 of *S. oneidensis*.

There was no detectable reduction of Cr(VI) by an abiotic control (sterile LB medium)

shown as black diamonds and the standard deviation of the triplicate measurements

ranged between $9 \times 10^{-5}$ to 0.02 mM. This standard deviation is too small to be

visualized in the figure. Strain MR-1 (WT and squares in the figure) completely reduced

the Cr(VI) during the 4 hour time period and only 1 culture of strain Δ2426 (2426 and

triangles in the figure) completely reduced Cr(VI). Error bars represent triplicate cultures

for stains MR-1 and Δ2426.

9 x $10^{-5}$ to 0.02 mM.  Strain MR-1 was able to completely transform the 0.3 mM

Cr(VI)by the fourth hour, with 0.15 mM left by the third hour of exposure.  Strain Δ2426

demonstrated comparable reduction as MR-1 during the first 3 hours of exposure.  In

contrast, between 3 and 4 hours of Cr(VI) exposure, strain Δ2426 transformed the Cr(VI)

incompletely for two of the three cultures.  The incompletely reduced cultures had 0.10

and 0.26 mM Cr(VI) remaining after 4 hours of cultivation.  The incomplete

transformation of Cr(VI) by Δ2426 for two of the three cultures indicates the possibility

that SO2426 may be involved in response to Cr(VI) toxicity.  The reason as to why one

culture from strain Δ2426 was able to transform Cr(VI) completely is not understood at

this time and is beyond the scope of this work.

*Global proteomics profile of strains MR-1 and Δ2426*

   A total of 5 mg of cell pellet were lysed, and the resulting protein content was

digested with trypsin and analyzed by 2D-HPLC-MS/MS.  The resulting MS/MS spectra

were searched with Sequest [39] for each culture during the 4 harvested time points.  A

total of 2,121 proteins comprising 44% of the predicted proteome were confidently

identified at the two peptide level for strains MR-1 and Δ2426.  Table 6.1 depicts the

total proteins identified for strains MR-1 and Δ2426 for each time point harvested.  The

molecular weight and isoelectric point distribution for the total dataset from each time

point was plotted and demonstrated no discrepancy when compared to the predicted

distribution of the proteome (data not shown).  The reproducibility at the protein level for

each biological replicate was at least 60%.  Previously, the aim was to obtain a

reproducibility of 70% for technical replicates [65, 147, 158, 184].  Achieving a

**Table 6.1.** Total Proteins Identified by Strains MR-1 and Δ2426 at Each Time Point

| Time of harvest | Amount of Cr(VI) present | Strain MR-1[a] | Strain Δ2426[a] |
|---|---|---|---|
| 0 hr | 0.3 mM | 1537 | 1398 |
| 1 hr | 0.3 mM | 1583 | 1540 |
| 3 hr | ~0.2 mM | 1535 | 1599 |
| 4 hr | 0.0 – 0.2 mM | 1468 | 1687 |

[a]Total number of proteins identified by the three biological replicates.

reproducibility of at least 60% as demonstrated here is comparable since there is the added complexity of an additional proteome dataset represented by the third culture.

All three replicates of each strain during each time point were grouped according to their functional category distribution annotated by the J. Craig Venter Institute (formerly The Institute for Genomic Research) and the results are presented in Figure 6.3. Overall, most of the functional categories did not fluctuate significantly between the first time point and the last or between the two strains. However, proteins annotated in the functional categories biosynthesis of cofactors, prosthetic groups, and carriers; hypothetical proteins; transport and binding proteins; and proteins of unknown function demonstrated large fluctuations in protein expression between the two strains or over time within a particular strain. Strain Δ2426 exhibited the most variation during the time-course. The number of proteins identified in a particular category was lower for the sample obtained at the 0 time point versus the last sample that was harvested at 4 h. This may be due to the strain in general demonstrating deficient growth, where the starter culture required a longer time of incubation prior to inoculation for strain Δ2426 (data not shown). The effects of removing a transcriptional regulator are known to cause growth deficiencies in bacteria [94, 255, 256].

The large fluctuations of biosynthesis of cofactors, prosthetic groups, and carriers; hypothetical proteins; central intermediary metabolism; transport and binding proteins; and proteins of unknown function may be the result of perturbing regulation of the expressed genome through the knockout of *so2426*. However, the hypothetical proteins category appears to fluctuate due to the presence of Cr(VI) in the growth media. Prior to the addition to Cr(VI), there are comparable numbers of proteins identified by each

174
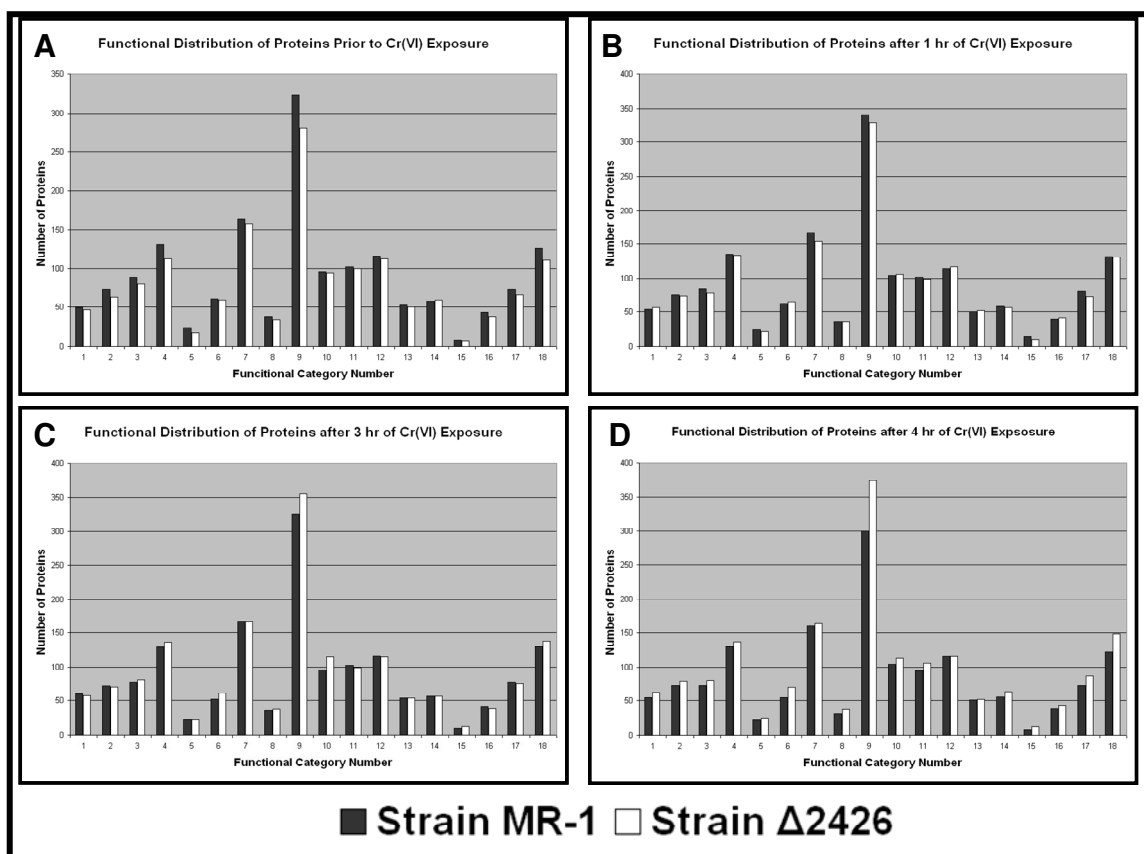
**Figure 6.3.** Functional category distribution of proteins from strains MR-1 and Δ2426. Distribution for each strain (A) prior to Cr(VI) exposure, (B) after 1 hr of exposure, (C) 3 hr exposure, and (D) 4 hr exposure.  The functional category numbers along the x axis are:  (1) Amino acid biosynthesis; (2) Biosynthesis of cofactors, prosthetic groups, and carriers; (3) Cell envelope; (4) Cellular processes; (5) Central intermediary metabolism; (6) DNA metabolism; (7) Energy metabolism; (8) Fatty acid and phospholipid metabolism; (9) Hypothetical proteins; (10) Mobile and extrachromosomal element functions; (11) Protein fate; (12) Protein synthesis; (13) Purines, pyrimidines, nucleosides, and nucleotides; (14) Regulatory functions; (15) Signal transduction; (16) Transcription; (17) Transport and binding proteins; and (18) Unknown function.

175

strain. However, at the 1 h time point (equivalent to acute shock) where no significant Cr(VI) transformation has been detected, there is an increase by at least 20 proteins between the two strains. The effect is much greater in strain Δ2426, with 48 proteins. In strain MR-1, the resulting number counts for hypothetical proteins declines during the 3 and 4 h time points. In contrast, strain Δ2426 consistently continues to increase the number of hypothetical proteins detected to the 4 h time point. The protein expression level implications of a number of hypothetical proteins are discussed further below.

*Protein expression effects due to deletion of so2426*

Within the *S. oneidensis* genome, *so2426* is organized into a cluster with five other genes (*so2422-2425* and *so2427*), which constitutes an approximately 5 kb region of the genome [69]. The gene cluster organization with respect to localization and transcriptional regulation is described in detail in Chourey *et al* [157]. Prior to Cr(VI) introduction into both strains, SO2424 and SO2427 were identified with the same number of peptides indicating there was no polar effect on expression of these proteins (Supplemental Table S1). Following 1 and 3 h of Cr(VI) exposure, SO2424 and SO2427 did not demonstrate a difference in protein expression level (Supplemental Table S2-S3). However, SO2422 was identified with 2 peptides in the mutant strain and not identified in the wild-type strain following Cr(VI) exposure for 1 h. After 3 h of Cr(VI) exposure, SO2422 was identified in both the wild-type and mutant Δ2426 strains. Following complete reduction of Cr(VI) in strain MR-1 at 4 h post-exposure, SO2422 was not identified as expressed, while SO2424 and SO2427 maintained the same relative protein expression level as prior to Cr(VI) introduction in both strains. SO2423 and SO2425 were not found in the MS/MS data presented here, however they were identified at

baseline levels (2-5 peptides) in previous work [147, 156]. Therefore, the expression of these two proteins may be repressed due to the deletion of *so2426*.

*Proteins differentially expressed between strains MR-1 and Δ2426*

Semiquantitation was employed as the method for determining proteins differentially expressed between the two strains. This method takes into account reproducible differences of percent sequence coverage, peptide count, and spectral count between the knockout mutant strain Δ2426 and the wild-type strain MR-1. Usually, semiquantitation compares the differences between technical replicates of a HPLC-MS/MS experiment having a reproducibility of at least 70% between them. In this study, the use of biological triplicates has reduced the reproducibility between the cultures to 60%, so attempting the semiquantitation method is more difficult and the number of proteins differentially expressed is reduced compared to previous studies on *S. oneidensis* [147, 156, 158].

In light of this criteria, a number of proteins were identified as differentially expressed reproducibly across the cultures between each strain. Table 6.2 comprises the differentially expressed proteins identified as down-regulated in the Δ2426 cultures. A total of 22 proteins were found down-regulated in Δ2426 cultures at any given time point, with six of these proteins down-regulated under all of the time points. Out of the total 22 proteins down-regulated in Δ2426 cultures, 73% were annotated in the functional categories of hypothetical, transport and binding, and energy metabolism during the time-course explored. In addition, 23 proteins were identified as up-regulated in strain the Δ2426 in comparison to strain MR-1, with only one protein (HugA, a heme transport protein) up-regulated prior to the addition of Cr(VI) to the cultures (Table 6.3). The

177

**Table 6.2.** Up-Regulated Proteins from MR-1 Cultures Versus Δ2426 Cultures

| Locus | Time | Average Δ2426 | | | Average MR-1 | | | Description (Number) |
|---|---|---|---|---|---|---|---|---|
| | | Peptide Count | Spectra Count | % Coverage | Peptide Count | Spectra Count | % Coverage | |
| SO0970 | 0 hr | 25.3 | 55.0 | 46.8 | 39.0 | 105.3 | 61.4 | fumarate reductase flavoprotein subunit precursor (7) |
| SO1072 | 1 hr | 4.0 | 5.3 | 18.2 | 9.3 | 14.3 | 24.9 | chitin-binding protein, putative (17) |
| SO1180 | 0 hr | 5.0 | 5.0 | 18.4 | 10.7 | 11.7 | 34.1 | PhoH family protein (18) |
| SO1190 | 0 hr | 0.0 | 0.0 | 0.0 | 11.0 | 13.0 | 43.0 | conserved hypothetical protein (9) |
| SO1190 | 1 hr | 0.0 | 0.0 | 0.0 | 12.0 | 25.0 | 46.0 | conserved hypothetical protein (9) |
| SO1190 | 3 hr | 0.0 | 0.0 | 0.0 | 18.7 | 40.7 | 58.3 | conserved hypothetical protein (9) |
| SO1190 | 4 hr | 0.0 | 0.0 | 0.0 | 14.3 | 26.3 | 57.3 | conserved hypothetical protein (9) |
| SO1405 | 4 hr | 0.7 | 0.7 | 3.5 | 5.7 | 8.0 | 23.2 | transglutaminase family protein (18) |
| SO1779 | 4 hr | 6.0 | 7.7 | 13.2 | 12.0 | 18.3 | 19.3 | decaheme cytochrome c (omcA) (7) |
| SO1784 | 1 hr | 2.7 | 3.7 | 7.2 | 7.7 | 14.0 | 14.6 | ferrous iron transport protein B (feoB) (17) |
| SO2492 | 1 hr | 5.7 | 7.7 | 10.3 | 11.0 | 18.7 | 21.7 | oxidoreductase, acyl-CoA dehydrogenase family (18) |
| SO2796 | 1 hr | 7.3 | 9.3 | 13.4 | 16.0 | 24.0 | 24.2 | conserved hypothetical protein (9) |
| SO2912 | 0 hr | 13.3 | 18.7 | 21.3 | 21.7 | 38.0 | 36.9 | formate acetyltransferase (pflB) (7) |
| SO2916 | 0 hr | 22.3 | 38.0 | 33.3 | 31.0 | 76.0 | 57.8 | phosphate acetyltransferase (pta) (7) |
| SO3030 | 1 hr | 3.3 | 4.7 | 9.9 | 14.0 | 22.7 | 35.4 | siderophore biosynthesis protein (alcA) (17) |
| SO3030 | 3 hr | 2.7 | 2.7 | 7.7 | 8.0 | 11.3 | 20.1 | siderophore biosynthesis protein (alcA) (17) |
| SO3032 | 1 hr | 4.3 | 5.3 | 9.2 | 13.7 | 28.7 | 27.9 | siderophore biosynthesis protein, putative (17) |
| SO3032 | 3 hr | 5.3 | 9.0 | 12.1 | 9.0 | 25.3 | 18.2 | siderophore biosynthesis protein, putative (17) |
| SO3032 | 4 hr | 7.3 | 12.3 | 15.0 | 11.3 | 27.7 | 22.5 | siderophore biosynthesis protein, putative (17) |
| SO3033 | 1 hr | 4.3 | 6.3 | 10.6 | 12.7 | 29.3 | 26.1 | ferric alcaligin siderophore receptor (17) |
| SO3637 | 1 hr | 9.0 | 40.0 | 33.9 | 15.3 | 85.3 | 46.4 | survival protein surA (surA) (11) |
| SO3833 | 0 hr | 5.7 | 7.7 | 26.3 | 11.7 | 17.0 | 44.5 | peptide chain release factor 1 (prfA) (12) |
| SO3863 | 0 hr | 3.7 | 4.3 | 26.1 | 15.7 | 41.0 | 65.8 | molybdenum ABC transporter (modA) (17) |

**Table 6.2.** Continued

| Locus | Time | Average Δ2426 | | | Average MR-1 | | | Description (Number) |
|---|---|---|---|---|---|---|---|---|
| | | Peptide Count | Spectra Count | % Coverage | Peptide Count | Spectra Count | % Coverage | |
| SO3863 | 3 hr | 8.7 | 14.3 | 39.9 | 15.3 | 44.7 | 63.4 | molybdenum ABC transporter (modA) (17) |
| SO3863 | 4 hr | 7.0 | 8.3 | 42.2 | 16.7 | 51.7 | 69.3 | molybdenum ABC transporter (modA) (17) |
| SO3967 | 0 hr | 0.7 | 0.7 | 7.4 | 8.3 | 12.3 | 64.4 | molybdenum ABC transporter, putative (17) |
| SO3967 | 1 hr | 1.3 | 2.0 | 7.3 | 12.3 | 20.3 | 64.2 | molybdenum ABC transporter, putative (17) |
| SO3967 | 3 hr | 0.7 | 0.7 | 2.3 | 11.0 | 23.7 | 59.4 | molybdenum ABC transporter, putative (17) |
| SO3967 | 4 hr | 1.7 | 1.7 | 6.5 | 13.0 | 29.7 | 55.0 | molybdenum ABC transporter, putative (17) |
| SO4509 | 0 hr | 0.0 | 0.0 | 0.0 | 16.7 | 23.7 | 25.6 | formate dehydrogenase, alpha subunit (7) |
| SO4509 | 1 hr | 0.0 | 0.0 | 0.0 | 12.3 | 15.7 | 19.3 | formate dehydrogenase, alpha subunit (7) |
| SO4509 | 3 hr | 0.0 | 0.0 | 0.0 | 8.7 | 14.3 | 13.0 | formate dehydrogenase, alpha subunit (7) |
| SO4509 | 4 hr | 0.0 | 0.0 | 0.0 | 9.7 | 14.7 | 15.8 | formate dehydrogenase, alpha subunit (7) |
| SO4513 | 0 hr | 3.7 | 4.0 | 7.1 | 15.0 | 20.0 | 22.6 | formate dehydrogenase, alpha subunit (7) |
| SO4513 | 1 hr | 0.0 | 0.0 | 0.0 | 6.0 | 7.3 | 9.2 | formate dehydrogenase, alpha subunit (7) |
| SO4513 | 3 hr | 0.0 | 0.0 | 0.0 | 6.7 | 10.3 | 11.0 | formate dehydrogenase, alpha subunit (7) |
| SO4513 | 4 hr | 0.0 | 0.0 | 0.0 | 6.3 | 8.7 | 8.8 | formate dehydrogenase, alpha subunit (7) |
| SO4719 | 0 hr | 18.3 | 25.0 | 65.8 | 31.0 | 66.7 | 75.6 | conserved hypothetical protein (9) |
| SO4719 | 1 hr | 22.3 | 35.3 | 74.1 | 32.3 | 89.0 | 78.7 | conserved hypothetical protein (9) |
| SO4719 | 3 hr | 24.0 | 60.7 | 75.2 | 38.7 | 198.7 | 79.7 | conserved hypothetical protein (9) |
| SO4719 | 4 hr | 22.0 | 35.3 | 71.6 | 38.3 | 118.3 | 78.2 | conserved hypothetical protein (9) |
| SO4743 | 0 hr | 0.0 | 0.0 | 0.0 | 9.7 | 10.3 | 22.0 | TonB-dependent receptor, putative (17) |
| SO4743 | 1 hr | 0.0 | 0.0 | 0.0 | 16.0 | 22.7 | 35.7 | TonB-dependent receptor, putative (17) |
| SO4743 | 3 hr | 3.0 | 3.0 | 8.9 | 14.3 | 22.0 | 29.1 | TonB-dependent receptor, putative (17) |
| SO4743 | 4 hr | 2.0 | 2.0 | 4.3 | 9.7 | 17.3 | 21.7 | TonB-dependent receptor, putative (17) |

functional category distribution for up-regulated proteins was more diverse compared to the down-regulated proteins, with 11 different categories represented and no proteins reproducibly up-regulated across the time course. In addition, 14 of the up-regulated proteins were found at the 4 h time point, indicating that differential expression may be due to Cr response and not necessarily the role of SO2426 as a transcriptional regulator.

Proteins annotated under the functional role of hypothetical proteins were identified as being repressed in the Δ2426 mutant cultures. Two conserved hypothetical proteins (SO1190 and SO4719) were consistently down-regulated during the entire time course. SO1190 was not identified as being expressed in any of the Δ2426 cultures, while SO4719 was found with an average of ~21 ± 3 peptides corresponding to a steady expression level over the time course. In addition, SO4719 behaved similarly in the MR-1 cultures with an average peptide count of 70 ± 5 as well corresponding to a steady expression level during the time course. The corresponding mRNA levels corroborated these results demonstrating an elevated induction of these genes during the Cr(VI) transformation period for strain MR-1 [157]. The predicted protein sequence of SO4719 was aligned to the top blast hits from 16 other sequenced genomes of the genus *Shewanella* (Supplemental Figure S1). SO4719 demonstrated good alignment with a predicted extracellular solute-binding protein (family 1) from eight of the species/strains. This group of proteins are members of the LysR substrate binding protein family, which is part of a larger clan of periplasmic binding proteins. However, SO4719 is not annotated in the clan of periplasmic binding proteins, but rather the protein is located within the protein family of PfamB PB011652 according to the Pfam database [pfam.sanger.ac.uk, [254]]. This family was generated automatically from an alignment

**Table 6.3.** Down-Regulated Proteins in MR-1 Cultures Versus Δ2426 Cultures

| Locus | Time | Average Δ2426 | | | Average MR-1 | | | Description (number) |
|---|---|---|---|---|---|---|---|---|
| | | Peptide Count | Spectra Count | % Coverage | Peptide Count | Spectra Count | % Coverage | |
| SO0343 | 1 hr | 36.0 | 78.0 | 53.3 | 21.0 | 38.3 | 32.3 | aconitate hydratase 1 (acnA) (7) |
| SO0343 | 4 hr | 45.3 | 116.7 | 62.4 | 26.7 | 53.0 | 40.3 | aconitate hydratase 1 (acnA) (7) |
| SO0344 | 4 hr | 23.0 | 48.7 | 54.0 | 7.7 | 10.7 | 26.2 | methylcitrate synthase (prpC) (7) |
| SO0345 | 1 hr | 18.0 | 98.3 | 58.8 | 13.0 | 46.0 | 48.7 | methylisocitrate lyase (prpB) (7) |
| SO0346 | 4 hr | 9.3 | 14.3 | 37.7 | 1.3 | 1.7 | 8.5 | transcriptional regulator. GntR family (14) |
| SO0575 | 4 hr | 10.0 | 13.7 | 18.7 | 3.0 | 6.3 | 5.3 | RNA polymerase-associated protein (hepA) (16) |
| SO0798 | 3 hr | 6.0 | 6.3 | 12.1 | 0.7 | 0.7 | 1.3 | conserved hypothetical protein (9) |
| SO0840 | 4 hr | 45.0 | 72.0 | 40.6 | 19.0 | 34.0 | 19.8 | acetyl-CoA carboxylase multifunctional enzyme accADC (8) |
| SO0934 | 3 hr | 6.0 | 8.3 | 27.2 | 0.0 | 0.0 | 0.0 | conserved hypothetical protein (9) |
| SO0958 | 4 hr | 11.3 | 17.0 | 63.7 | 6.0 | 7.0 | 37.8 | alkyl hydroperoxide reductase, C subunit (ahpC) (4) |
| SO1484 | 4 hr | 8.3 | 11.3 | 24.9 | 1.3 | 2.3 | 4.9 | isocitrate lyase (aceA) (7) |
| SO1551 | 1 hr | 13.0 | 18.0 | 31.5 | 4.7 | 5.7 | 14.2 | GGDEF domain protein (18) |
| SO1551 | 4 hr | 16.3 | 33.7 | 31.8 | 5.0 | 7.3 | 14.4 | GGDEF domain protein (18) |
| SO1898 | 4 hr | 11.3 | 14.3 | 62.2 | 5.3 | 7.0 | 42.8 | transcriptional regulator, putative (14) |
| SO1930 | 4 hr | 39.0 | 69.7 | 46.3 | 20.3 | 34.3 | 26.9 | 2-oxoglutarate dehydrogenase, E1 component (sucA) (7) |
| SO2415 | 4 hr | 16.7 | 24.7 | 28.5 | 7.7 | 12.3 | 14.0 | ribonucleoside-diphosphate reductase, alpha subunit (nrdA) (8) |
| SO3089 | 4 hr | 5.7 | 8.3 | 23.9 | 0.0 | 0.0 | 0.0 | fatty oxidation complex, beta subunit (8) |
| SO3545 | 3 hr | 25.3 | 99.3 | 49.1 | 14.7 | 39.0 | 28.6 | OmpA family protein (3) |
| SO3585 | 4 hr | 8.3 | 14.0 | 39.1 | 0.7 | 0.7 | 1.8 | azoreductase, putative (4) |
| SO3669 | 0 hr | 10.7 | 19.7 | 21.9 | 5.3 | 7.3 | 13.4 | heme transport protein (hugA) (17) |
| SO3669 | 3 hr | 40.3 | 155.0 | 57.3 | 21.3 | 57.7 | 37.2 | heme transport protein (hugA) (17) |
| SO3669 | 4 hr | 42.7 | 158.7 | 60.1 | 24.7 | 74.0 | 39.4 | heme transport protein (hugA) (17) |
| SO3681 | 1 hr | 10.7 | 19.7 | 79.0 | 4.0 | 7.0 | 46.8 | universal stress protein family (4) |
| SO3914 | 3 hr | 20.7 | 35.3 | 28.3 | 8.0 | 14.0 | 10.6 | TonB-dependent receptor, putative (17) |

**Table 6.3.** Continued

| Locus | Time | Average Δ2426 | | | Average MR-1 | | | Description (number) |
|---|---|---|---|---|---|---|---|---|
| | | Peptide Count | Spectra Count | % Coverage | Peptide Count | Spectra Count | % Coverage | |
| SO3914 | 4 hr | 22.0 | 30.7 | 31.8 | 7.3 | 10.3 | 12.3 | TonB-dependent receptor, putative (17) |
| SO4523 | 1 hr | 27.3 | 58.7 | 49.7 | 9.0 | 16.0 | 21.5 | iron-regulated outer membrane virulence protein (irgA) (17) |
| SO4523 | 3 hr | 29.0 | 67.0 | 45.9 | 7.0 | 14.3 | 18.1 | iron-regulated outer membrane virulence protein (irgA) (17) |
| SO4523 | 4 hr | 34.3 | 84.3 | 58.7 | 11.7 | 19.0 | 27.7 | iron-regulated outer membrane virulence protein (irgA) (17) |
| SO4652 | 3 hr | 14.0 | 32.3 | 35.4 | 8.0 | 19.0 | 20.8 | sulfate ABC transporter, periplasmic sulfate-binding protein (sbp) (17) |
| SOA0048 | 1 hr | 24.7 | 31.3 | 38.5 | 13.0 | 16.0 | 25.5 | prolyl oligopeptidase family protein (11) |
| SOA0048 | 3 hr | 30.0 | 48.3 | 44.0 | 15.3 | 29.7 | 29.1 | prolyl oligopeptidase family protein (11) |

created by Prodom (prodom.prabi.fr) having the annotation of being a tungsten extracellular binding protein. The protein sequence also aligned well with a sulfate substrate binding protein from five of the species/strains. Therefore, SO4719 may be under transcriptional control of SO2426, but does not appear to demonstrate a Cr(VI) exposure-dependent regulation.

SO1190 is located in a putative operon with two other genes (*so1188* and *so1189*), which were identified as highly induced in the microarray work [157]. SO1190 was found with an average of 14 peptides over the time course and demonstrates the highest expression level at the 3 h time point during Cr(VI) transformation. SO1189 was identified solely in the MR-1 cultures during the 0, 1, and 3 h time points, with an average of 2 peptides, and SO1188 was not identified in this study. Our previous study testing acute chromate shock in *S. oneidensis* [147] identified both SO1188 and SO1189 expressed after 45 min of growth in the presence of 1 mM Cr(VI). Both proteins were identified with 2-4 peptides, which is the minimum for positive protein identification according to the criteria in the study. SO1188 was solely identified in the membrane fraction, where 2 peptides from SO1189 were identified in the crude fraction and the rest of the peptides in the membrane fraction. Based on these observations, SO1188-SO1190 may be under the transcriptional control of SO2426.

Due to SO1188 and SO1189 appearing to associate with the membrane fraction of *S. oneidensis* cell lysates previously [147]; this may be the reason the shotgun proteomics results presented here identify SO1189 with the minimum peptide count. The development of mass spectrometry proteome technology is not as mature with respect to sensitivity as microarray technology. In addition, the lysis method utilized here does not

183

obtain the level of membrane protein coverage found previously through fractionation [147, 156, 158]. Factors such as protein solubility and, in the case of this study, trypsin specificity, reduce the peptide count of membrane embedded proteins. These factors account for some of the discrepancy between microarray and proteomics technologies. For example, each protein contains 12 (SO1188) and 21 (SO1189) tryptic cleavage sites (lysines and arginines). Using biological replicates instead of technical replicates reduces the biological background and increases the confidence level that a protein found up-regulated reproducibly is actually responding to the condition presented to the culture. SO1189 is an example of a protein that was borderline up-regulated during the 3 h time point. Each LC-MS experiment comprised a unique culture of either strain MR-1 or Δ2426 (indicative of Culture 1, 2, and 3, respectively). Cultures 2 and 3 identified the protein with 3 and 5 tandem mass spectra, respectively, which Culture 3 passes the criteria of spectral count in the semiquantitation method. However, due to the absence of detection of SO1189 in Culture 1, this protein is not reproducibly up-regulated during the time point. If only Culture 3 had been sampled for proteome analysis, this protein would have been identified as differentially expressed for that particular time point.

The most prevalent category to demonstrate down-regulation of proteins was transport and binding. Two transport and binding proteins were repressed during the entire time course in the Δ2426 cultures. SO3967 encodes a putative molybdenum ABC transporter and was identified predominately in the MR-1 cultures (average of 18.5 peptides) with an increased induction at the 3 and 4 h time points. SO4743 encodes a putative TonB-dependent receptor and was predominantly found in the MR-1 cultures as well, with an average of 12.4 peptides during the time course. SO4743 demonstrated a

different trend with an increased peptide count identified at the 1 and 3 h time points. The 4 h time point demonstrated a similar abundance to that found prior to the addition of Cr(VI) in the media. The Cr(VI)-dependent regulation of this protein is implicated in the results presented here and may be one pathway in which Cr(VI) gains access into the cell. Previously, SO4743 was also found to be highly up-regulated following acute chromate shock [147] and appears to be induced irrespective of Cr(VI) concentration [156]. The repression of transport and binding proteins in strain Δ2426 during Cr(VI) transformation indicates that SO2426 may be a transcriptional regulator of these particular proteins and may operate in conjuction with Fur to regulate iron homeostasis.

A total of six proteins were found to be repressed in strain Δ2426 cultures belonging to the functional category energy metabolism. Included in the repressed proteins are two separate alpha subunits of a formate dehydrogenase complex (SO4509 and SO4513). There are a total of three formate dehydrogenase operons encoded in the *S. oneidensis* genome, *so0102-so0104*, *so4509-so4511, so4513-so4515. so0102-so0104* encode a nitrate inducible formate dehydrogenase with the alpha subunit, *so0102*, containing a selenocysteine residue. SO0102 was identified in both the wildtype MR-1 strain and the mutant Δ2426 strain with two or three unique peptides, respectively. The presence of SO0102 was identified solely prior to the introduction of Cr(VI) in the cultures. Interestingly, this nitrate-inducible operon appears to demonstrate a low-level constitutive expression in the absence of Cr(VI). One of the subunits of this operon possibly encodes a molybdenum center (SO0103) that coordinates with the selenocysteine containing alpha subunit (SO0102). Therefore, Cr(VI) may be interfering

185

with molybdenum uptake, which is causing up-regulation of the molybdendum ABC

transporter under Cr(VI) stress thereby allowing more molybdenum uptake into the cell.

*so4509-so4511* and *so4513-so4515* are organized in tandem and appear to be

constitutively expressed under wildtype conditions. Formate dehydrogenase is a multi-

protein enzyme complex, which uses the electron donor formate for respiration and

utilizes transition metals in the redox center [257]. Prior to the alpha subunit genes in

both operons, there are small genes annotated within the genome, *so4508* and *so4512*.

There is no evidence found in the proteome datasets to suggest these proteins are

expressed. They may represent regulatory sequences, which have characteristics

indicative of a protein-coding gene. SO4509, alpha subunit of formate dehydrogenase,

was solely identified in the MR-1 cultures over the entire time course. The alpha subunit

identified here is the only soluble member of the complex, which explains the ease of

identification of this protein [257]. Formate dehydrogenase in MR-1 has been shown to

be an important member of the electron transport chain to the terminal ferric reductase in

anaerobically grown MR-1 using Fe(III) citrate [102]. Cr(VI) may be inhibiting the

activity of formate dehydrogenase under the conditions tested here.

Other energy metabolism proteins identified included a fumarate reductase

flavoprotein subunit precursor (SO0970), PflB (formate acetyltransferase), and Pta

(phosphate acetyltransferase), which were all found repressed in the Δ2426 cultures prior

to chromate introduction. The decaheme cytochrome c, OmcA, was found to be

repressed in Δ2426 cultures after 4 h of exposure to Cr(VI). This protein is known to

play an important role in the reduction of Fe(III) and Mn(IV) in *S. oneidensis* MR-1 [107,

113, 258]. During the other time points sampled for proteome analysis, there is no change in the level of expression between strain MR-1 and Δ2426 cultures for OmcA.

As mentioned, 11 of the functional categories were represented as up-regulated in Δ2426 cultures over the time points sampled for proteome analysis. These categories included proteins involved in energy metabolism and iron transport. For the most part, the proteins repressed in Δ2426 cultures were members of two molybdenum ABC transporters. The proteins up-regulated in Δ2426 were involved in transport of iron (HugA and IrgA). IrgA was up-regulated during Cr(VI) exposure and transformation in Δ2426 cultures and was borderline up-regulated prior to Cr(VI) introduction. HugA was up-regulated in the control (0 h) and during active transformation, but borderline up-regulated during the 1 h time point. The up-regulation of these proteins during active Cr(VI) transformation indicates a possible intracellular iron deficiency more pronounced in the Δ2426 cultures versus wild-type cultures.

A putative operon involved in energy metabolism and comprising genes *so0344-so0346* was identified as up-regulated in strain Δ2426 (Table 6.3). *so0346* encodes a transcriptional regulator from the GntR family and is at the beginning of the operon. *so0344* (*prpC*) and *so0345* (*prpB*) are located directly downstream of *so0346*. *prpBC* encode methylisocitrate lyase and methylcitrate synthase, respectively, which are required for growth on propionate in *Salmonella enterica* [259]. SO0346 does share over 90% sequence similarity to other members of the genus *Shewanella*, indicating this is a conserved putative transcriptional regulator (data not shown). In addition, the operon structure also appears to be conserved with other *Shewanella* species as well. The GntR family of transcriptional regulators contain a N-terminal helix-turn-helix DNA-binding

domain with one of four subtypes of a C-terminal signaling domain [260, 261]. For SO0346, the C-terminal domain belongs to the FadR subtype, featuring an all alpha-helical structure [260], which are usually involved in regulating activity of proteins involved with amino acid metabolism pathways. This is in contrast to the structural organization of SO2426 and may indicate interaction between these two proteins. The GntR family of transcriptional regulators are known to dimerize with one another [260], the up-regulation of SO0346 may be a consequence of the protein dimerizing with itself in the absence of SO2426.

SO0346 was only detected in strains MR-1 and Δ2426 upon Cr(VI) introduction to the growth media (Supplemental Tables S1-S4). The level of expression for the other two members of the operon did not demonstrate any change in regulation prior to chromate introduction, indicating that SO0346 is not necessary for transcriptional activation of *so0344* and *so0345*. After 1 h in the presence of Cr(VI), SO0346 was identified in both strains, with only one of the triplicate cultures of strain MR-1 detecting this protein and all three cultures of strain Δ2426 having a similar detection level for this protein. After 3 h of Cr(VI) exposure, three peptides from SO0346 were found in two of the replicates from strain MR-1 cultures. SO0346 in strain Δ2426 cultures was borderline up-regulated at this time point with 5, 10, and 12 peptides, respectively identified in the replicate cultures. When approximately half of the Cr(VI) had been transformed in strain Δ2426, SO0346 was detected up-regulated and only one of the replicate cultures of strain MR-1 detected any peptides at this point corresponding to complete transformation of Cr(VI) (Table 6.3). The transcriptomic data indicates that the gene is up-regulated only after 24 h of exposure in strain Δ2426 [157]. Based on results

from previous studies [147, 156, 158, 248], SO0346 has not been found to be up-regulated in wild-type cultures exposed to various concentrations and intervals of chromate exposure. At this time, the exact mode of regulation of SO0346 with respect to the proposed function of SO2426 is not clear. Further investigation will be needed in order to decipher the relationship between SO2426 and SO0346 if indeed a relationship exists.

*Proteins differentially expressed during Cr(VI) transformation in strain MR-1*

Wild-type cultures of strain MR-1 prior to and following exposure to Cr(VI) were compared to identify proteins that are differentially expressed during acute shock and transformation of Cr(VI) (Tables 6.4 and 6.5). A total of 42 proteins were found to be up-regulated following addition of Cr(VI) with two primary functional categories represented (energy metabolism and transport and binding proteins). In contrast, 13 proteins were down-regulated following exposure to Cr(VI) in MR-1 cultures. There was a greater variety of functional categories represented in the down-regulated proteins; however energy metabolism dominated with five members down-regulated. Relatively few proteins identified as down-regulated indicated that the MR-1 cultures had not begun to be stressed by the transformation of Cr(VI) [158].

A number of the proteins found differentially expressed in MR-1 compared to Δ2426 were also found to demonstrate a Cr(VI) dependent response. For instance, AcnA, aconitate hydratase 1, was found to be up-regulated in response to Cr(VI) exposure in MR-1 cultures, but up-regulated in the Δ2426 cultures. AcnA and AcnB are well characterized in *E. coli* [262] with a role in the TCA cycle by catalyzing the isomerization of citrate to isocitrate. This important step in the TCA cycle is one reason

189

**Table 6.4.** Proteins Identified as Up-Regulated upon Cr(VI) Exposure in MR-1 Cultures

| Locus | Average 0hr MR-1 culture | | | Average Cr(VI) exposed MR-1 culture | | | Time point | Description (Number) |
| | Peptide Count | Spectra Count | % Coverage | Peptide Count | Spectra Count | % Coverage | | |
|---|---|---|---|---|---|---|---|---|
| SO0343 | 8.7 | 14.7 | 12.9 | 21.0 | 38.3 | 32.3 | 1 | aconitate hydratase 1 (acnA) (7) |
| SO0343 | 8.7 | 14.7 | 12.9 | 26.0 | 53.3 | 39.5 | 3 | aconitate hydratase 1 (acnA) (7) |
| SO0343 | 8.7 | 14.7 | 12.9 | 26.7 | 53.0 | 40.3 | 4 | aconitate hydratase 1 (acnA) (7) |
| SO0345 | 6.7 | 18.3 | 37.0 | 13.0 | 46.0 | 48.7 | 1 | methylisocitrate lyase (prpB) (7) |
| SO0345 | 6.7 | 18.3 | 37.0 | 16.0 | 45.0 | 53.1 | 3 | methylisocitrate lyase (prpB) (7) |
| SO0429 | 15.3 | 19.3 | 31.5 | 20.7 | 41.7 | 41.0 | 1 | peptidase, M13 family (3) |
| SO0429 | 15.3 | 19.3 | 31.5 | 24.3 | 55.3 | 45.2 | 4 | peptidase, M13 family (3) |
| SO0518 | 1.7 | 1.7 | 5.6 | 6.7 | 9.7 | 23.3 | 4 | outer membrane efflux family protein, putative (17) |
| SO0519 | 4.0 | 4.0 | 15.3 | 9.0 | 18.7 | 31.9 | 3 | cation efflux protein, putative (4) |
| SO0519 | 4.0 | 4.0 | 15.3 | 9.7 | 20.3 | 34.0 | 4 | cation efflux protein, putative (4) |
| SO0554 | 4.0 | 5.3 | 26.6 | 9.3 | 17.3 | 50.4 | 3 | hypothetical protein (9) |
| SO0554 | 4.0 | 5.3 | 26.6 | 10.3 | 16.0 | 45.8 | 4 | hypothetical protein (9) |
| SO1072 | 2.7 | 4.7 | 11.2 | 9.3 | 14.3 | 24.9 | 1 | chitin-binding protein, putative (17) |
| SO1072 | 2.7 | 4.7 | 11.2 | 8.7 | 10.3 | 27.8 | 3 | chitin-binding protein, putative (17) |
| SO1075 | 11.3 | 12.0 | 18.3 | 17.3 | 24.3 | 29.5 | 1 | conserved hypothetical protein (9) |
| SO1482 | 5.0 | 5.3 | 10.7 | 10.7 | 18.0 | 23.2 | 1 | TonB-dependent receptor, putative (17) |
| SO1490 | 23.0 | 46.7 | 59.2 | 29.7 | 107.3 | 65.2 | 4 | alcohol dehydrogenase II (adhB) (7) |
| SO1677 | 24.3 | 57.7 | 64.5 | 31.0 | 138.3 | 73.8 | 1 | acetyl-CoA acetyltransferase (atoB) (8) |
| SO1678 | 1.3 | 2.0 | 3.1 | 8.0 | 19.0 | 22.6 | 4 | methylmalonate-semialdehyde dehydrogenase (mmsA) (7) |
| SO1679 | 11.3 | 20.7 | 34.6 | 18.0 | 44.0 | 54.8 | 3 | acyl-CoA dehydrogenase family protein (8) |
| SO1679 | 11.3 | 20.7 | 34.6 | 20.0 | 49.3 | 52.6 | 4 | acyl-CoA dehydrogenase family protein (8) |
| SO1755 | 10.0 | 15.3 | 23.7 | 15.7 | 34.3 | 36.7 | 4 | phosphoglucomutase/phosphomannomutase family (7) |
| SO1816 | 6.0 | 7.7 | 28.6 | 12.7 | 30.0 | 45.9 | 3 | conserved hypothetical protein (9) |

**Table 6.4.** Continued

| Locus | Average 0hr MR-1 culture | | | Average Cr(VI) exposed MR-1 culture | | | Time point | Description (Number) |
|---|---|---|---|---|---|---|---|---|
| | Peptide Count | Spectra Count | % Coverage | Peptide Count | Spectra Count | % Coverage | | |
| SO1894 | 13.0 | 18.0 | 29.5 | 22.7 | 42.3 | 40.4 | 4 | acetyl-CoA carboxylase, biotin carboxylase, putative (8) |
| SO2903 | 33.3 | 214.0 | 86.6 | 45.0 | 482.7 | 91.9 | 1 | cysteine synthase A (cysK) (1) |
| SO2903 | 33.3 | 214.0 | 86.6 | 47.7 | 474.0 | 93.5 | 3 | cysteine synthase A (cysK) (1) |
| SO2903 | 33.3 | 214.0 | 86.6 | 46.3 | 526.7 | 91.2 | 4 | cysteine synthase A (cysK) (1) |
| SO3030 | 0.7 | 0.7 | 2,2 | 14.0 | 22.7 | 35.4 | 1 | siderophore biosynthesis protein (alcA) (17) |
| SO3030 | 0.7 | 0.7 | 2,2 | 8.0 | 11.3 | 20.1 | 3 | siderophore biosynthesis protein (alcA) (17) |
| SO3032 | 4.3 | 6.0 | 8.5 | 13.7 | 28.7 | 27.9 | 1 | siderophore biosynthesis protein, putative (17) |
| SO3032 | 4.3 | 6.0 | 8.5 | 11.3 | 27.7 | 22.5 | 4 | siderophore biosynthesis protein, putative (17) |
| SO3033 | 4.0 | 7.7 | 8.1 | 12.7 | 29.3 | 26.1 | 1 | ferric alcaligin siderophore receptor (17) |
| SO3033 | 4.0 | 7.7 | 8.1 | 13.3 | 33.3 | 24.6 | 3 | ferric alcaligin siderophore receptor (17) |
| SO3033 | 4.0 | 7.7 | 8.1 | 9.3 | 22.0 | 17.7 | 4 | ferric alcaligin siderophore receptor (17) |
| SO3145 | 8.3 | 22.7 | 43.7 | 13.3 | 46.3 | 61.2 | 4 | electron transfer flavoprotein, beta subunit (etfB) (7) |
| SO3237 | 6.7 | 13.7 | 38.2 | 11.7 | 43.0 | 48.3 | 3 | flagellin (4) |
| SO3238 | 5.3 | 12.0 | 35.0 | 10.7 | 37.3 | 53.7 | 4 | flagellin (4) |
| SO3430 | 16.3 | 23.3 | 47.4 | 35.3 | 133.7 | 77.1 | 1 | recA protein (recA) (6) |
| SO3430 | 16.3 | 23.3 | 47.4 | 37.3 | 129.3 | 75.2 | 3 | recA protein (recA) (6) |
| SO3430 | 16.3 | 23.3 | 47.4 | 36.7 | 122.0 | 85.3 | 4 | recA protein (recA) (6) |
| SO3462 | 0.0 | 0.0 | 0,0 | 7.7 | 10.7 | 17.1 | 1 | DNA repair protein RecN (recN) (6) |
| SO3462 | 0.0 | 0.0 | 0,0 | 11.0 | 14.0 | 28.6 | 3 | DNA repair protein RecN (recN) (6) |
| SO3462 | 0.0 | 0.0 | 0.0 | 8.0 | 12.0 | 19.4 | 4 | DNA repair protein RecN (recN) (6) |
| SO3599 | 14.0 | 28.3 | 56.5 | 24.0 | 88.7 | 71.9 | 1 | sulfate ABC transporter, periplasmic sulfate-binding (cysP) (17) |
| SO3599 | 14.0 | 28.3 | 56.5 | 22.7 | 90.7 | 67.1 | 4 | sulfate ABC transporter, periplasmic sulfate-binding (cysP) (17) |
| SO3667 | 4.0 | 4.3 | 40.6 | 20.0 | 87.7 | 85.8 | 1 | conserved hypothetical protein (9) |
| SO3667 | 4.0 | 4.3 | 40.6 | 16.0 | 69.3 | 74.6 | 3 | conserved hypothetical protein (9) |

**Table 6.4.** Continued

| Locus | Peptide Count | Spectra Count | % Coverage | Peptide Count | Spectra Count | % Coverage | Time point | Description (Number) |
|-------|---------------|---------------|------------|---------------|---------------|------------|------------|----------------------|
| | **Average 0hr MR-1 culture** | | | **Average Cr(VI) exposed MR-1 culture** | | | | |
| SO3667 | 4.0 | 4.3 | 40.6 | 17.3 | 64.7 | 76.6 | 4 | conserved hypothetical protein (9) |
| SO3669 | 5.3 | 7.3 | 13.4 | 27.0 | 78.3 | 46.9 | 1 | heme transport protein (hugA) (17) |
| SO3669 | 5.3 | 7.3 | 13.4 | 21.3 | 57.7 | 37.2 | 3 | heme transport protein (hugA) (17) |
| SO3669 | 5.3 | 7.3 | 13.4 | 24.7 | 74.0 | 39.4 | 4 | heme transport protein (hugA) (17) |
| SO3673 | 1.7 | 3.0 | 9.1 | 15.3 | 35.0 | 54.1 | 1 | hemin ABC transporter, periplasmic hemin-binding (hmuT) (17) |
| SO3673 | 1.7 | 3.0 | 9.1 | 15.7 | 43.0 | 52.6 | 3 | hemin ABC transporter, periplasmic hemin-binding (hmuT) (17) |
| SO3673 | 1.7 | 3.0 | 9.1 | 16.3 | 37.3 | 55.3 | 4 | hemin ABC transporter, periplasmic hemin-binding (hmuT) (17) |
| SO3726 | 1.7 | 2.3 | 6.3 | 8.0 | 13.7 | 31.6 | 1 | sulfate adenylyltransferase, subunit 1 (cysN) (5) |
| SO3727 | 3.7 | 4.7 | 18.8 | 13.0 | 19.0 | 43.4 | 1 | sulfate adenylyltransferase, subunit 2 (cysD) (5) |
| SO3727 | 3.7 | 4.7 | 18.8 | 12.0 | 22.7 | 45.6 | 3 | sulfate adenylyltransferase, subunit 2 (cysD) (5) |
| SO3737 | 10.0 | 11.7 | 22.5 | 20.7 | 39.7 | 35.7 | 1 | sulfite reductase (NADPH) hemoprotein beta-component (cysI) (5) |
| SO3737 | 10.0 | 11.7 | 22.5 | 23.3 | 47.3 | 42.3 | 3 | sulfite reductase (NADPH) hemoprotein beta-component (cysI) (5) |
| SO3914 | 5.3 | 9.0 | 11.1 | 12.3 | 17.0 | 19.2 | 1 | TonB-dependent receptor, putative (17) |
| SO4134 | 1.3 | 1.7 | 18.6 | 8.0 | 13.3 | 50.0 | 4 | conserved hypothetical protein (9) |
| SO4215 | 13.7 | 20.0 | 41.3 | 23.7 | 46.0 | 52.9 | 4 | cell division protein FtsZ (ftsZ) (4) |
| SO4476 | 0.7 | 0.7 | 3.7 | 7.7 | 14.3 | 27.2 | 4 | spheroplast protein y precursor, putative (18) |
| SO4480 | 0.0 | 0.0 | 0.0 | 7.3 | 10.3 | 14.2 | 3 | aldehyde dehydrogenase (aldA) (18) |
| SO4652 | 1.3 | 2.7 | 6.3 | 10.0 | 32.7 | 28.8 | 4 | sulfate ABC transporter, periplasmic sulfate-binding (sbp) (17) |
| SO4655 | 0.7 | 0.7 | 1.9 | 5.0 | 11.7 | 26.2 | 4 | sulfate ABC transporter, ATP-binding protein (cysA-2) (17) |
| SO4692 | 9.0 | 11.3 | 11.9 | 14.0 | 25.3 | 19.9 | 1 | AcrB/AcrD/AcrF family protein (4) |
| SO4743 | 9.7 | 10.3 | 23.1 | 16.0 | 22.7 | 35.7 | 1 | TonB-dependent receptor, putative (17) |
| SOA0100 | 9.7 | 15.0 | 21.4 | 18.7 | 35.0 | 40.9 | 4 | conserved hypothetical protein (9) |

**Table 6.5.**  Proteins Identified as Down-Regulated upon Cr(VI) Exposure in MR-1 Cultures

| | Average 0hr MR-1 culture | | | Average Cr(VI) exposed MR-1 culture | | | | |
|---|---|---|---|---|---|---|---|---|
| Locus | Peptide Count | Spectra Count | % Coverage | Peptide Count | Spectra Count | % Coverage | Time point | Description (Number) |
| SO0575 | 10.0 | 25.7 | 15.7 | 3.0 | 6.3 | 5.3 | 4 | RNA polymerase-associated protein (hepA) (16) |
| SO0848 | 27.3 | 41.7 | 38.0 | 8.0 | 13.3 | 15.2 | 1 | periplasmic nitrate reductase (napA) (7) |
| SO0848 | 27.3 | 41.7 | 38.0 | 4.3 | 5.0 | 9.8 | 3 | periplasmic nitrate reductase (napA) (7) |
| SO0848 | 27.3 | 41.7 | 38.0 | 5.0 | 7.7 | 9.7 | 4 | periplasmic nitrate reductase (napA) (7) |
| SO1519 | 9.0 | 10.3 | 21.1 | 3.3 | 3.7 | 9.1 | 4 | iron-sulfur cluster-binding protein (7) |
| SO1602 | 20.0 | 26.7 | 14.5 | 10.0 | 13.3 | 7.9 | 3 | multi-domain beta-ketoacyl synthase (18) |
| SO1602 | 20.0 | 26.7 | 14.5 | 11.3 | 16.0 | 8.3 | 4 | multi-domain beta-ketoacyl synthase (18) |
| SO1631 | 10.0 | 23.0 | 49.2 | 4.3 | 8.3 | 21.9 | 4 | uridylate kinase (pyrH) (13) |
| SO2218 | 14.0 | 21.0 | 31.2 | 8.3 | 19.7 | 24.5 | 4 | asparaginyl-tRNA synthetase (asnS) (12) |
| SO2248 | 8.0 | 12.0 | 17.4 | 3.7 | 5.7 | 12.1 | 3 | L-serine dehydratase 1 (sdaA) (7) |
| SO2427 | 26.7 | 54.0 | 39.0 | 20.3 | 35.0 | 30.3 | 4 | TonB-dependent receptor, putative (17) |
| SO2705 | 20.7 | 26.0 | 30.6 | 9.0 | 12.0 | 12.4 | 3 | DNA topoisomerase I (topA) (6) |
| SO3565 | 10.7 | 11.7 | 19.5 | 3.0 | 3.3 | 6.7 | 1 | 2,3-cyclic-nucleotide 2-phosphodiesterase (cpdB) (13) |
| SO3565 | 10.7 | 11.7 | 19.5 | 2.3 | 3.0 | 5.3 | 4 | 2,3-cyclic-nucleotide 2-phosphodiesterase (cpdB) (13) |
| SO3980 | 20.7 | 42.3 | 36.1 | 9.0 | 12.7 | 20.0 | 1 | cytochrome c552 nitrite reductase (7) |
| SO3980 | 20.7 | 42.3 | 36.1 | 5.0 | 7.3 | 12.4 | 3 | cytochrome c552 nitrite reductase (7) |
| SO3980 | 20.7 | 42.3 | 36.1 | 6.7 | 7.7 | 15.4 | 4 | cytochrome c552 nitrite reductase (7) |
| SO4513 | 15.0 | 20.0 | 20.6 | 6.0 | 7.3 | 9.2 | 1 | formate dehydrogenase, alpha subunit (7) |
| SO4513 | 15.0 | 20.0 | 20.6 | 6.7 | 10.3 | 11.0 | 3 | formate dehydrogenase, alpha subunit (7) |
| SO4513 | 15.0 | 20.0 | 20.6 | 6.3 | 8.7 | 8.8 | 4 | formate dehydrogenase, alpha subunit (7) |
| SO4659 | 6.0 | 8.7 | 21.2 | 0.0 | 0.0 | 0.0 | 3 | conserved hypothetical protein (9) |
| SO4659 | 6.0 | 8.7 | 21.2 | 0.0 | 0.0 | 0.0 | 4 | conserved hypothetical protein (9) |

why the isozyme AcnB has a compensatory pathway. AcnB has been found to be constitutively expressed, but easily inactivated due to demetallation [262]. AcnA is characterized as a compensatory counterpart to AcnB and is resistant to oxidative stress. This resistance is unusual for this type of [4Fe-4S] cluster enzyme. The presence of AcnA is generally as a minor component (i.e. no Cr(VI) added equals ~13 spectra identified); however, during Cr(VI) transformation the prevalence of AcnA increases up to 50 MS/MS spectra identified at the 4 h time point. This result is expected due to the transformation of Cr(VI) to Cr(III), which can proceed through the unstable redox active species Cr(V) [86, 126, 140, 146]. The fact that AcnA demonstrated increased expression in the Δ2426 cultures compared to MR-1 provides further evidence of the increased oxidative stress in strain Δ2426 cultures. In addition, the AcnB product was found not to (1) change expression level following Cr(VI) exposure and (2) was not down-regulated in the Δ2426 cultures. This is expected based on the known role of AcnB in *E. coli* as a house-keeping protein [262].

The alpha subunit of formate dehydrogenase (SO4513) demonstrated a contrasting expression pattern to AcnA. As described above, SO4513 was found to be highly up-regulated in MR-1 cultures and to be expressed in Δ2426 cultures only in the absence of Cr(VI). This protein was further found to be down-regulated following exposure to Cr(VI) in the MR-1 cultures. All three time points sampled following Cr(VI) introduction revealed approximately the same level of expression (~6 peptides on average identified) in contrast to prior to Cr(VI) introduction (15 peptides on average identified). This result indicates a perhaps complex control mechanism for expression of SO4513.

A number of proteins involved in sulfate uptake and metabolism were also identified as up-regulated upon exposure to Cr(VI) in the MR-1 cultures. CysK, CysP, CysN, CysD, CysI, Sbp, and CysA-2 were all found to be up-regulated post Cr(VI) introduction to MR-1 cultures. In fact, a majority of these proteins were identified as up-regulated during the initial shock time point (1 h post Cr(VI) introduction), indicating an initial up-take of Cr(VI) perhaps through the sulfate system thereby creating an intracellular sulphate deficiency, as has been implicated elsewhere [141]. CysK encodes the enzyme cysteine synthase A, which catalyzes the final synthesis step by condensing sulfide with *O*-acetyl-L-serine to yield cysteine [263]. In *Bacillus subtilis*, CysK was shown to negatively regulate the sulfur metabolism genes similar to the nitrogen metabolism regulatory system [263]. Therefore, the up-regulation of CysK under the conditions presented here implicates a possible deficiency of intracellular sulfur in response to the presence of Cr(VI) or oxidative stress. The possible regulatory role of CysK in *S. oneidensis* is not clear at this time. Two sulfate ABC transporters are encoded in the MR-1 genome (*so3599-so3602* and *so4652-4655*) with three members from these operons up-regulated following exposure to Cr(VI). CysP, the periplasmic sulfate-binding protein, is a member of the first operon and was found up-regulated during the 1 and 4 h time points, but demonstrated an equal expression level at the 3 h time point. Sbp is the periplasmic protein of the second operon and was found up-regulated following complete Cr(VI) transformation at the 4 h time point in addition to member CysA-2, the ATP-binding protein. Since Cr(VI) has been demonstrated as an inhibitor of sulfate uptake [141], there may be a deficiency within the cytoplasm of sulfide leading to the activation of genes necessary for sulfur acquisition and metabolism.

*Post-translational modifications of strain MR-1 during Cr(VI) transformation*

The 0, 1, 3, and 4 h MR-1 datasets were searched using InsPecT to determine MS/MS containing post-translational modifications (PTMs) overlooked during the initial Sequest search. Thompson *et al* [59] discusses the optimization of the scoring scheme used here, where the p-value threshold chosen for each dataset demonstrates a false discovery rate (FDR) of 2.0±0.2%. Therefore, each dataset will have a unique p-value cutoff due to the scoring distribution of that dataset. Table 6.6 contains the p-value cutoff chosen for each dataset with the corresponding FDRs observed. The p-value distribution was from 0.45 to 0.83 for the four time points of the MR-1 cultures sampled. The resulting FDR when considering PTM-containing peptides only was 12.0-19.0%. For example, the MR-1 culture harvested prior to Cr(VI) addition had a FDR of 2.1% with the p-value cutoff score of 0.45 and a total of 10,871 PTM peptides identified. All PTM peptide identifications can be accessed in Supplemental Table S5. Prior to Cr(VI) introduction into the cultures (0 h), many more PTM peptides were identified. This is in contrast to the non-modifications searches, where a comparable number of total proteins and peptides were identified under all of the time points (Table 6.1). The resulting optimization of the filtering threshold for identification of PTM peptides to a specific FDR leads to a more accurate comparison across the information in the datasets.

The most prevalent PTM in the 0 h dataset was oxidation of methionines, cysteines, tryosines, and tryptophan residues, followed by diacetylation of lysines and arginines. In contrast, there was approximately the same amount of monomethylations as diacetylations in the 4 h dataset and much fewer dimethylations identified. Relative to the total number of PTM peptides identified for the two respective datasets (Table 6.6),

196

**Table 6.6.** InsPecT p-value Thresholds and Corresponding FDRs for Strain MR-1 PTMs

| Culture Time point | p-value | Total Peptide Identifications | | | PTM Peptide Identifications | | |
|---|---|---|---|---|---|---|---|
| | | Reverse IDs | Total IDs | % FDR | Reverse IDs | Total IDs | % FDR |
| 0 hr | 0.45 | 733 | 68640 | 2.1 | 694 | 11565 | 12.0 |
| 1 hr | 0.70 | 670 | 74244 | 1.8 | 621 | 8166 | 15.2 |
| 3 hr | 0.83 | 759 | 71248 | 2.1 | 745 | 8736 | 17.1 |
| 4 hr | 0.75 | 712 | 65193 | 2.2 | 706 | 7415 | 19.0 |

there were more oxidized peptides found at the 4 h time point relative to the 0 h control. The oxidation identifications for both datasets primarily consisted of monooxidations. In addition, the other two Cr(VI) exposure time points, 1 and 3 h, exhibited the same overall distribution of PTMs characteristic of the 4 h time point. Even though phosphorylations are common signaling PTMs, this modification was not searched in the resulting datasets. In bacteria, most phosphorylations occur on histidine or aspartic acid residues consisting of temporally rapid events that are difficult to detect [52, 236, 264]. In addition, the method of fragmentation utilized in this study (collision induced dissociation) is not amenable for retaining the location of the phosphorylation site on the peptide resulting in a neutral loss of 98 Da as the primary fragment ion [59].

Due to the global differences in PTM characteristics for the time course, the PTM-containing peptides identified for each time point were organized according to the functional category role of the parent protein. Figure 6.4 depicts the category distribution of the peptides from each time point. The protein synthesis role category predominates across all time points as containing the greatest number of post-translationally modified peptides. This is expected due to the proteins encoding the ribosome comprising this role category, which have been shown previously to be highly modified [48]. Peptides with a role in amino acid biosynthesis demonstrated increased modifications indicative of increased control over this process (Figure 6.4). In addition, as exposure to Cr(VI) progresses, proteins involved in cellular processes become gradually more modified indicating a change in the level of regulation. However, a majority of the categories do not change over Cr(VI) transformation time, but rather the constituents that are modified vary.

198
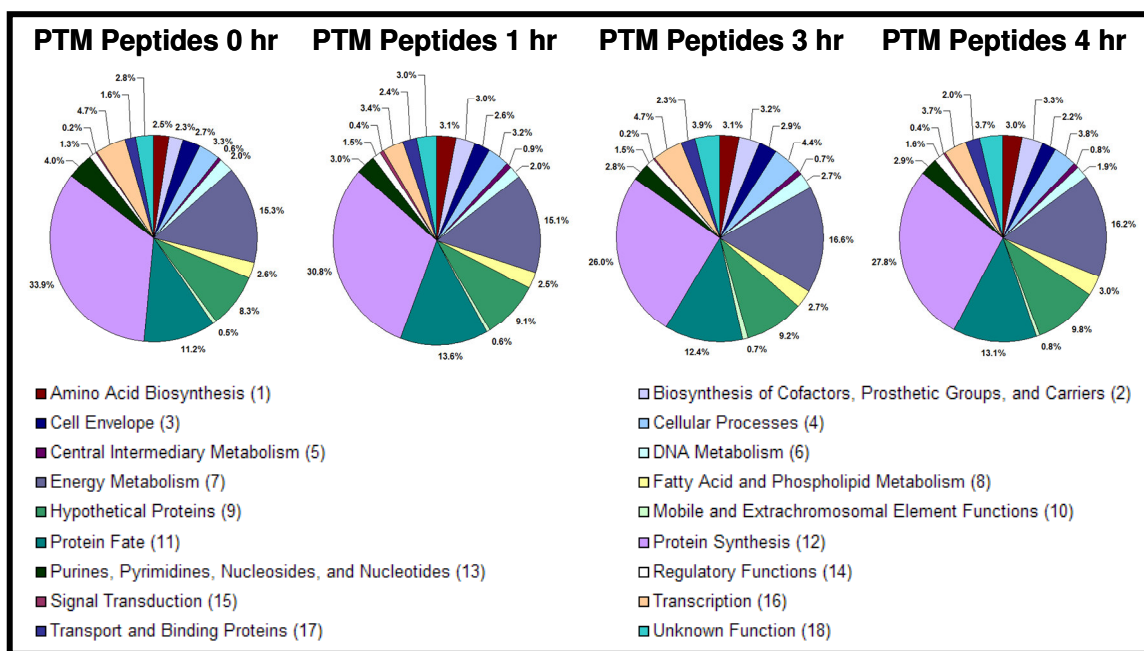
**Figure 6.4.** Post-translationally modified peptides identified by InsPecT for strain MR-1. Each slice represents a functional category and comprises the percent of peptides for the category identified out of the total PTM peptides for each time point (i.e. 0 hr, 1 hr, 3 hr, and 4 hr).

The functional category of signal transduction comprises proteins involved in gene expression regulation, namely the 53 two-component response regulators annotated in this group. SO2426, which has been knocked out in this study, is an orphan response regulator in this category demonstrating a number of PTMs over time. Over the time course presented here, signal transduction proteins identified as unmodified comprised 8-14 protein identifications. This is expected due to these members having a low copy number within a bacterial cell. By including differential chemical modifications for the InsPecT search, the identification rate of signal transduction proteins increased to at least 11-18 proteins identified over the time course.

A number of signal transduction proteins demonstrated Cr(VI)-dependent modifications during the time course evaluated here. Three proteins demonstrated differential modification states over the time course: SO2426, SO2742, and SO4472. SO2426 is a putative Cr(VI) transcriptional regulator shown here to be involved in controlling gene expression of a number of proteins. SO2742 is an orphan sensor histidine kinase/response regulator similar to SO2426, while NtrC (SO4472) is a nitrogen regulatory protein. Supplemental Figures S2-S4 are the annotated MS/MS of all the modified peptides documented for SO2742. There were a total of seven amino acids found to be chemically modified after 1 and 4 h of exposure to Cr(VI). An oxidized methionine was identified after Cr(VI) exposure for 1 h. After 4 h, two dioxidized methionines, a triply oxidized cysteine, a monomethylated lysine residue, and two arginines were dimethylated and diacetylated, respectively. Interestingly, one of these peptides included the monomethylated lysine, dioxidized methionine, and the dimethylated arginine would not have been identified with previous methods of PTM

search methods [51, 59], due to restrictions in both the computational requirements of algorithms and the flexibility in specifying PTMs in the algorithm parameters. However, the PTMs present on SO2742 are interesting and may point to a level of control for this protein after Cr(VI) transformation has occurred.

SO2426 is a response regulator that has been implicated to specifically respond to Cr(VI) exposure previously [147, 156]. Due to the importance of this protein in response to chromate exposure, understanding possible regulation by PTMs is necessary to decipher both regulation and function within the bacterium. Prior to Cr(VI) addition to the strain MR-1 cultures, the presence of SO2426 was not detected with PTMs and without any proteomic evidence from the non-modifications searches to suggest that the protein had even been translated at that time. Following addition of Cr(VI) to strain MR-1 cultures, SO2426 was identified in the proteome samples acquired, present with and without modified peptides. According to the identifications made by the algorithm InsPecT, a total of five putative spectra were identified as comprising a PTM peptide with two spectra found to be incorrect identifications after manual inspection. SO2426 was identified with 2-3 peptides for all three MR-1 cultures after exposure to Cr(VI) for 1 h. Following 3 h of exposure to Cr(VI), a total of six peptides passed the filter thresholds used for InsPecT. The 4 h time point yielded a total of 3 peptides for cultures 2 and 3 only. The identification of one modified peptide was found in culture 3 as well as the unmodified counterpart. The modification consisted of a methionine oxidation that may or may not be due to the sample preparation procedures used in this study. This same methionine residue was found to be oxidized after culture 2 was exposed to Cr(VI) for 4 h. The identification of a diacetylated arginine was found in culture 2 at the 3 h

timepoint. At this time, there is active transformation of Cr(VI) to presumably Cr(III) (Figure 6.2). The dicetylation was confidently identified by four fragment ions in the MS/MS and corresponds to R13 of SO2426 (Figure 6.5). This peptide also contains a triply oxidized cysteine and a doubly oxidized methionine. The acetylation may be regulatory in nature; however, the significance of this modification is unknown at this time. The regulation is possibly due to the relative proximity of the diacetylated arginine (R13) to an aspartic acid that is putatively phosphorylated aspartic acid (D57) [232].

**Conclusions**

Although this global proteomic study has provided a great deal of valuable information as to the response of *S. oneidensis* to chromate exposure during the transformation period, much more effort is needed in order to fully appreciate the physiological response to Cr(VI) exposure and stress. The deletion mutant strain Δ2426 demonstrated a deficiency in Cr(VI) transformation during the time course examined in this study. The deficiency may be due to genes that are involved in chromate transformation being under the control of SO2426. In regards to possible gene control of SO2426, a number of proteins were found to be heavily down-regulated in strain Δ2426 when compared to the wild-type strain MR-1 during parallel time points. Out of the differentially expressed proteins identified, the presence of a chromate reductase protein was not clear. However, as suggested by Viamajala *et al* [265], there may not be a dedicated chromate reductase encoded in the genome of *S. oneidensis*, but rather a number of reductase enzymes may recognize chromate as a substrate.
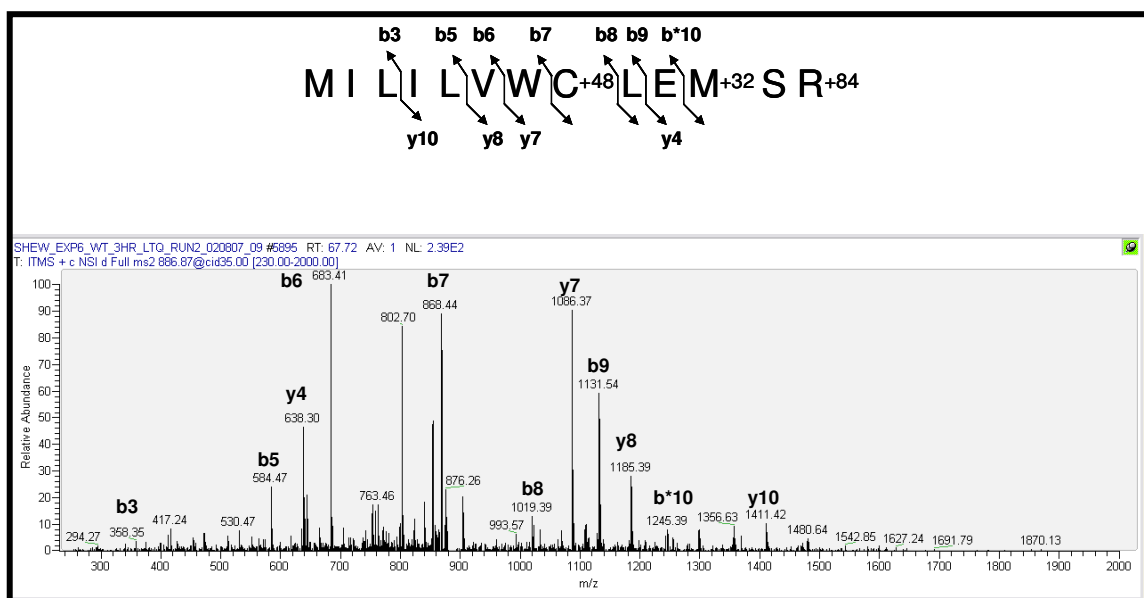
**Figure 6.5.** A peptide MS/MS from SO2426 with confirmed modifications.

The N-terminal peptide from SO2426 (MILILVWC$^{+48}$LEM$^{+32}$SR$^{+84}$) has two oxidation

events and a diacetylation event. Fragment ions y4, y7, y8, and y10 confirm the

diacetylated arginine residue, while fragment ions b8, b9, and b*10 confirm the triply

oxidized methionine. All three modifications are present in fragment ions y7, y8, and

y10.

## Chapter 7

## An Experimental Approach for Large-Scale Proteome Measurements from Small-Scale Amounts of Microbial Cultures and Communities

Melissa R. Thompson, Jennifer M. Froelich, Brian Erickson, Nathan C. VerBerkmoes, and Robert L. Hettich. Experimental Approach For Large-Scale Proteome Measurements From Small Amounts (low mg) Of Microbial Cultures. *In preparation for submission to Analytical Chemistry*, 2007. *Melissa Thompson performed sample preparation and LTQ measurements with the assistance of Jennifer M. Froelich, Brian Erickson, and Nathan C. VerBerkmoes. All data analysis was performed by Melissa R. Thompson. Supplemental material located at journal website.*

## Introduction

A foremost consideration in shotgun proteomics experiments has been the amount of biomass needed for deep proteome characterization by LC-MS. With many different methods of lysing microorganisms available to choose from, the most common methods utilize a manual disruption of the membrane to release the protein complement of the cell. Research in mass spectrometry-based proteomics has focused both method development and application efforts on microorganisms that are readily cultured in a laboratory. The limited range of organisms that can be cultured and studied with this technique is one of the many disadvantages to limiting the scope of investigation with these approaches. In addition, many tissue samples available from more complex organisms may be limiting as well [266].

In general, the amount of starting microbial cellular material has been a limiting factor in shotgun proteomics experiments. A relatively large quantity of biomass has

been the standard for an in-depth analysis of the protein complement in a given organism. Investigations published as recently as 2006 demonstrate the large amounts of cellular biomass (greater than 1 g) needed for an evaluation of the resulting proteome [63, 97, 267-269]. In 2003, Corbin *et al* [267] investigated the relationship between the transcriptome and the proteome of *Escherichia coli* MG1655. In order to obtain enough material for the proteome analysis, cultivation of ~3 x $10^9$ cells in a 100 mL culture was necessary. Furthermore, a recently published study by Ding *et al* [63] on various growth conditions of the dissimilatory metal-reducing bacterium *Geobacter sulfurreducens* also exploited large culturing conditions (a 1.5 L culture volume) to obtain a suitable quantity of starting cellular material. The significant quantity of starting material needed is also due to the considerable losses inherent to the sample processing steps, which involve lysis, fractionation, digestion, and cleanup. In addition to lysis, proteome experiments usually involve the fractionation of protein into a membrane-associated and a soluble fraction [63, 147, 156, 270]. Fractionation of cellular lysate requires greater amounts of starting material (at least 750 mg) due to the inherent losses caused by centrifugal separation and splitting the sample. However, fractionation was necessary previously, in order to routinely identify low-abundant proteins. The use of technical replicates also requires an increase in the amount of starting biomass. Developing a method where lysis and digestion occur in the same tube reduces the number of surfaces proteins/peptides come in contact with, thereby reducing sample losses, while maintaining a level of proteome coverage similar to the traditional lysis method.

The ability to efficiently rupture the cellular membrane of an organism, whether it is of macro- (mammalian tissue) or micro-size (bacteria) has been at the forefront of

205

biological research. In 1970, Coleman *et al* first described the enzyme lysozyme as an efficient agent to disrupt the cellular membrane of bacteria exposing the nucleic acid inside [271]. However, this technique is unsuitable for gram-positive bacteria, which have a different membrane composition [272]. Van Huynh *et al* described a method of lysing gram-positive bacteria using DMSO instead of a lytic enzyme [272]. Both of these methods are not suitable for shotgun proteomics experiments due to the use of excessive amounts of an enzyme for the lysozyme method [271] and employment of SDS detergent for the DMSO method [272]. The lysozyme protein would suppress the signal of endogenous proteins and SDS interferes with the activity of the commonly used protease trypsin [273] as well as the resulting LC-MS/MS analysis [274] in a shotgun proteomics experiment. Therefore, a new method of lysis had to be developed to alleviate these inherent problems associated with shotgun proteomics. Two methods of manual disruption, sonication and French press, of microbial cells have been used extensively [275, 276], and have been shown to be efficient in lysing bacterial cells [276]. The disadvantage of these lysis methods is their requirement for milliliter volumes of sample. As discussed below for microbial ecology studies, there will most likely not be that much starting material available. Designing a method of lysis that does not require a large volume for the suspended cells is critical in order to fully utilize the capabilities offered by shotgun proteomics experiments. To address this issue, Wang *et al* in 2005 developed a method of lysis that utilizes trifluroethylene (TFE) as a lysing agent followed by urea as the protein denaturant [266]. The disadvantage of this method is the increased dilution needed after the denaturing step to minimize trypsin inhibition in urea [277]. The method

206

described in this study uses Guanidine HCl as both a lysis agent and denaturant yielding in-depth proteome coverage.

To date, the number of microbial species that are cultivable has been estimated to be less than 1% [278-280]. In fact, the estimate for genome complexity in soil is equivalent to ~6000-10,000 *E. coli* genomes, but the complexity that is actually recovered through cultivation is equivalent to 40 *E. coli* genomes [281]. This creates a bias in the characterization and understanding of microbes with respect to both their evolutionary role and their metabolic role in the natural world [280]. Many ecologists and population geneticists understand the need for utilizing a proteomics approach for classifying metapopulation structures and the adaptive processes at work in the natural environment [282]. Micoorganisms play key metabolic roles in natural settings like degradation and transformation of metal and organic pollutants in the soil [71, 121, 129, 283] and are utilized in the treatment of sewage wastewater [284, 285]. A few studies on the proteome characterization of natural microbial communities have been published to date. Wilmes and Bond [284, 285] utilized two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) to characterize proteins present in a laboratory-scale activated sludge system. Powell *et al* [286] in 2005 used a similar yet distinct approach to characterize the dissolved organic matter (DOM) of marine environments utilizing SDS-PAGE LC-MS and MudPIT [33]. Recent work in our laboratory has demonstrated the comprehensive characterization of a natural microbial community found in acid mine drainage [65]. This microbial community thrives in the mine drainage, and was characterized by LC/LC-MS/MS. All of these studies discuss the critical need for better sample preparation methods, in particular for processing small samples for characterizing

natural microbial communities.  A comprehensive approach to the lysis of the microbes, chromatographic separation of the digested peptides, and analysis of shotgun proteomics data from microbial communities will greatly expand knowledge in microbial ecology.

## Materials and Methods

*Reagents, Sample Acquisition, and Sample Preparation*

Chemical reagents (i.e. Guanidine HCl) were acquired from Sigma Chemical Co. (St. Louis, MO) and were used as supplied without further purification.  Modified sequencing grade trypsin (Promega, Madison, WI) was used for all protein digestions.  Trifluoroethylene (TFE) was purchased from Fluka (Buchs, Switzerland).  HPLC-grade water and acetonitrile were obtained from Burdick & Jackson (Muskegon, MI), and 99% formic acid was purchased from EM Science (Darmstadt, Germany).  Wild-type *Shewanella oneidensis* MR-1 was cultivated under aerobic conditions as described in Brown *et al* [147] and was a gift from Dr. D. Thompson.  Wild-type *Rhodopseudomonas palustris* CGA0010 was cultured under photoheterotrophic conditions and was a gift from Dr. D. Pelletier of Oak Ridge National Laboratory.  The AMD biofilm was sampled from the Richmond Mine as described [65] and was a gift from Dr. J. Banfield of the University of California at Berkley.

Five separate lysis techniques were tested during the course of this study: sonication, Guanidine HCl, TFE, freeze-grinding, and bead-beating.  Figure 7.1 depicts a flow diagram of the experimental differences (sample size, preparation steps) between the lysis techniques of sonication and the two single tube (Guanidine HCL and TFE) methods described here.  Sonication was performed on *S. oneidensis* and *R. palustris* as a control lysis technique and is described in detail in [147] for *S. oneidensis*, with the only
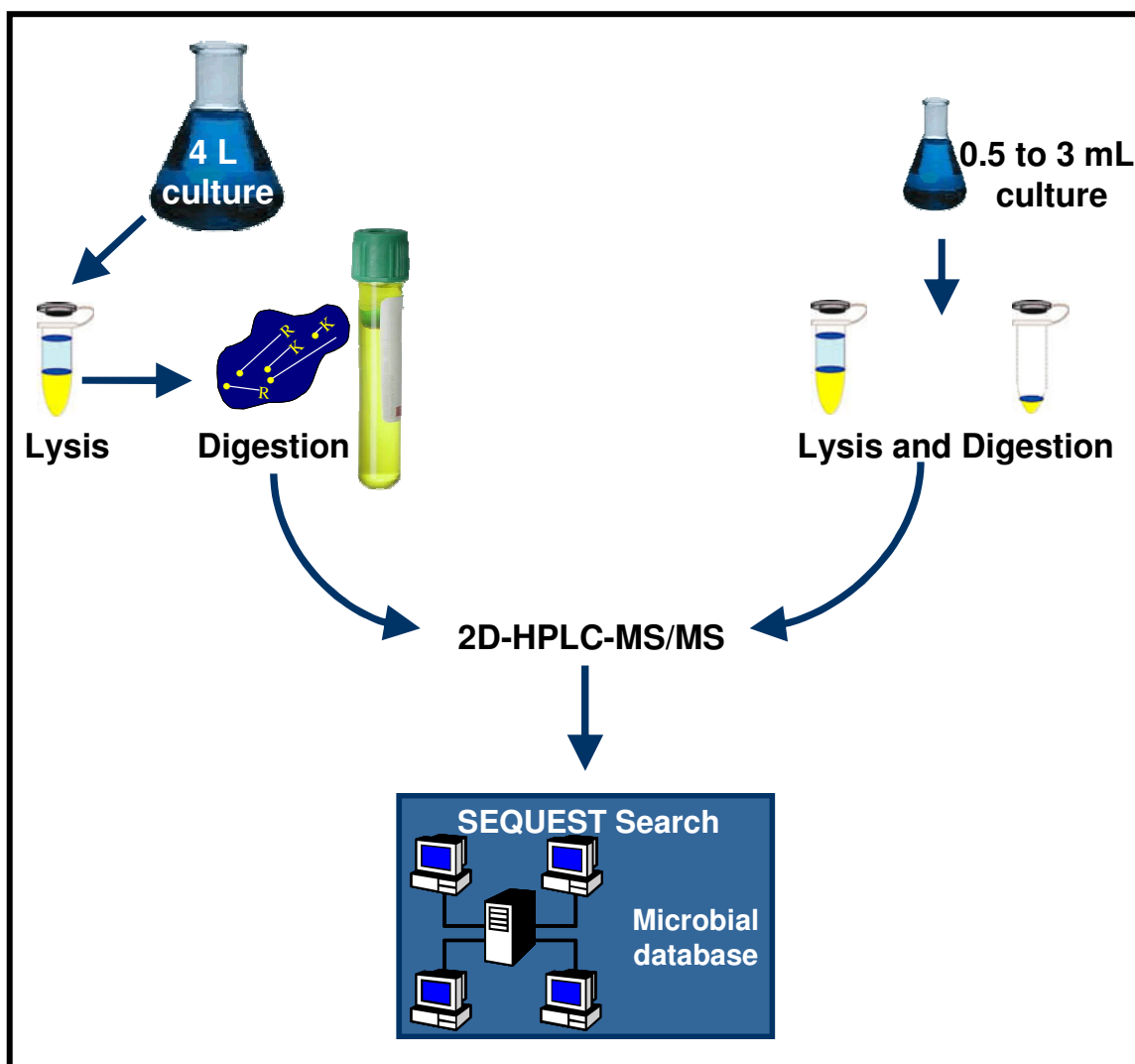
**Figure 7.1.** Schematic of the traditional sonication versus the single tube lysis method. In a traditional sonication lysis, this requires approximately 4 L of culture to obtain ~4 g of wet cell pellet. While, the single tube lysis method requires much less ~0.3 to 5.0 mL of culture for down to 1 mg of wet cell pellet.

modification for *R. palustris* being 10-30 s sonication pulses.  The Guanidine HCl lysis method was performed by lysing and simultaneously denaturing the resulting protein content with 6 M Guanidine/10 mM DTT dissolved in 50 mM Tris/ 10 mM $CaCl_2$ pH 7.6 (Tris Buffer) overnight at 37 °C with details described in Chapter 2.  Lysis using TFE was performed in the same manner as the Guanidine HCl method except TFE was added initially at a concentration of 50:50 TFE:Tris Buffer/10mM DTT.  For the AMD biofilm, two other methods were attempted in addition to guanidine-lysis and TFE.  The freeze-grinding method was performed by first flash-freezing the biofilm with liquid $N_2$ followed by pulverization of the biofilm for subsequent cellular lysis and digestion as described in Chapter 2.  A bead-beating method was also attempted for the AMD biofilm to determine if this method would destroy the biofilm structure and is also described in Chapter 2.  Details of the tryptic digestion following lysis can be found in Chapter 2.

*Protein Yield Quantification*

Wet cell pellets of 1, 5, and 10 mg from *R. palustris*; 1, 5, and 10 mg from *S. oneidensis* were lysed overnight in 6 M Guanidine HCl/ Tris Buffer.  Lysates were then centrifuged to pellet cellular debris and BCA analysis (Pierce, Rockford, IL) was performed according to the manufacturer's instructions.  A total of eight replicate cell lysates were analyzed for the 1 & 5 mg *R .palustris* and 1, 5, & 10 mg *S. oneidensis* samples, with the Biofilm samples and 10 mg *R. palustris* sample having two replicates. The average protein yield with standard deviation error bars was plotted in Excel (Microsoft, Redmond, WA).  The working absorbance range for the samples was 0.2 to 2.0.  Due to interfering chromophores present in the AMD biofilm, the BCA quantification was not performed.

*LC/LC-MS/MS Analysis*

All resulting lysis samples were analyzed via two-dimensional on-line liquid chromatography using identical Ultimate pumps (LC Packings; a Division of Dionex, San Francisco, CA) coupled to a LCQ ion trap (Thermo Electron, Waltham, MA) or a LTQ linear trapping quadrupole (Thermo Electron). The flow rate of the Ultimate pump at ~100 µL/min was split pre-column to achieve a final flow rate at the nanospray tip of ~300 nL/min. The samples were loaded onto a split-phase column (packed in-house with $C_{18}$ reverse phase and SCX chromatographic resin) as described in [147]. The split-phase column was placed behind a 15 cm $C_{18}$ analytical column (packed in-house [147]) and both were situated in front on a Thermo Electron nanospray source for the LCQ and a Proxeon nanospray source (Odense, Denmark) for the LTQ. The liquid chromatographic method used here for analyses with the LCQ and LTQ consisted of increasing step pulses of (0-500 mM) ammonium acetate salt, followed by a 2 hr 100% aqueous (95% $H_2O$, 5% ACN, 0.1% formic acid) to 50% organic (30% $H_2O$, 70% ACN, 0.1% formic acid) gradient. During the liquid chromatographic separation both the LCQ and LTQ were operated in a data-dependent mode [147, 156] and under the control of the Xcalibur software (Thermo Electron).

*Proteome Bioinformatics*

The following databases were used to search resulting MS/MS spectra with the algorithm Sequest [39] or DBDigger [44]. For the *S. oneidensis* dataset, the database (4,798 open reading frames) was downloaded from TIGR (www.tigr.org, Comprehensive Microbial Resource) and concatenated with a list of common contaminants (trypsin, keratin, etc.). The *R. palustris* dataset used the database annotated at Oak Ridge National

Laboratory and can be accessed at compbio.ornl.gov/rpal_proteome/databases. The AMD biofilm database used was from Tyson *et al* [67]. The following parameters were used for all searches: enzyme type, trypsin; Parent Mass Tolerance, 3.0; Fragment Ion Tolerance, 0.5; up to 4 missed cleavages allowed, and fully tryptic peptides only. Output files were then sorted with DTASelect [171] with the following filtering criteria for the Sequest searches: tryptic peptides only, delCN value of at least 0.08, and Xcorr values of at least 1.8 (+1), 2.5 (+2), 3.5 (+3). For the DBDigger searches the following criteria were applied using DTASelect [171]: tryptic peptides only, delCN of 0.08, and MASPIC [173] scores of at least 25 (+1), 30 (+2), and 45 (+3). The criteria used here have been tested in our laboratory [59, 65, 156, 184, 287] and give a false positive identification rate between 1-2% for bacterial isolates as well as simple microbial mixtures and communities. A comprehensive list of all proteins identified for a given sample size was created using the algorithm Contrast [171].

## **Results and Discussion**

*Lysis Efficiency*

In order to determine the efficiency of Guanidine HCl as a lysing agent with respect to bacterial cells, control lysing experiments (sonication) of *R. palustris* and *S. oneidensis* were performed (Figure 7.2), and protein was quantified using the BCA analysis. The results for a *R. palustris* cell pellet of 1 mg yielded 85.6 ± 18.9 µg/mL protein (equates to 94.16 µg of total protein) and the 5 mg pellet 286 ± 45.6 µg/mL protein (equates to 343.2 µg of total protein). The *S. oneidensis* 1 mg cell pellets yielded 276 ± 52.4 µg/mL protein and the 5 and 10 mg protein yields were initially above the working absorbance range, so a second set of replicates were diluted 2-fold and yielded

212
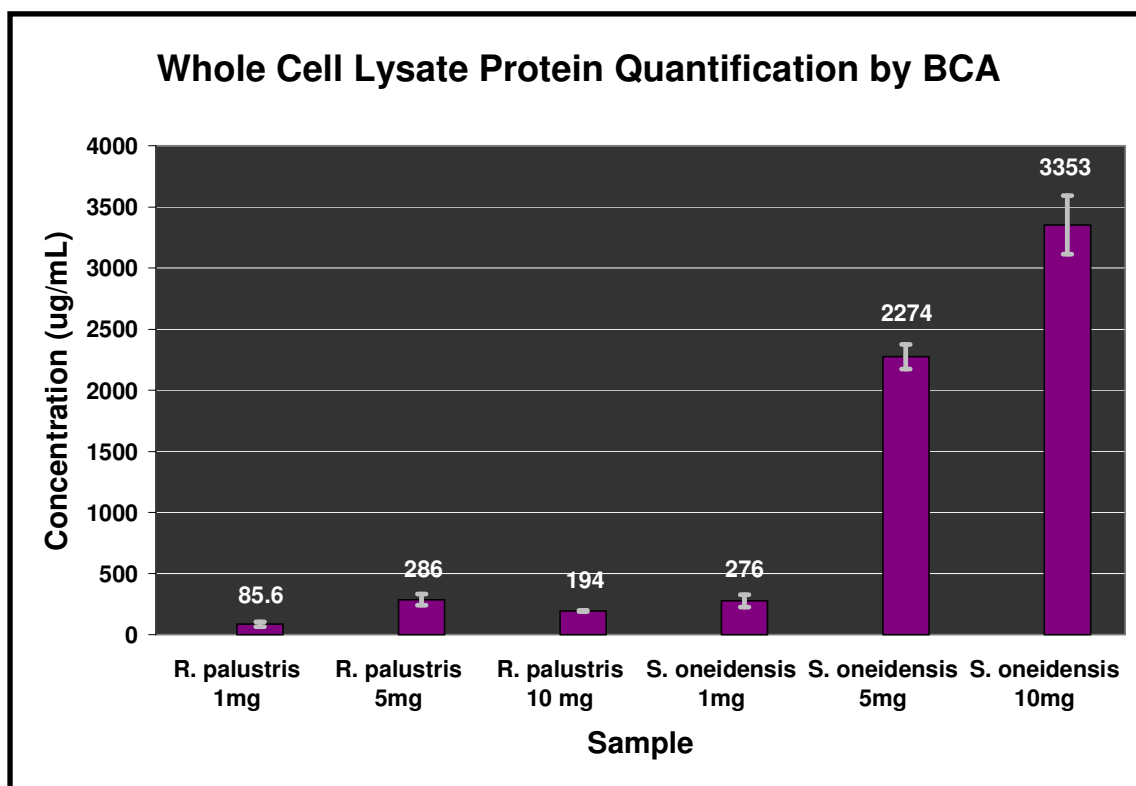
**Figure 7.2.** BCA protein quantification assay for different amounts of wet cell pellets. Protein quantification was performed with 1, 5 or 10 mg of *R. palustris* lysed by adding 6 M Guanidine or 1, 5, or 10 mg of *S. oneidensis* by the addition of 6 M Guanidine. Error bars represent triplicate measurements and values within the bar graph are the average protein measured for each cell pellet mass for the respective bacterial species.

2274 ± 102 and 3353 ± 241 µg/mL protein, respectively.  As apparent from above, this

method lyses *S. oneidensis* more efficiently than *R. palustris*, which is due to their

respective membrane compositions.  The membrane of *R. palustris* grown

photoheterophically contains many infoldings [288] that are not present in *S. oneidensis*;

which makes penetrating the membrane for lysis more difficult.  Therefore, a way to

improve the lysing efficiency of *R. palustris* is to add more denaturant (Guanidine HCl)

to the cell pellets.

*Bacterial Isolates Characterization*

       *S. oneidensis* wet cell pellets were lysed using sonication and Guanidnine HCl in

order to assess the proteome coverage identified with both methods of lysis.  From a total

of ~6 g of wet cell pellet harvested, ~4 g was lysed by sonication to yield ~82.5 mg of

crude protein.  The other 2 g of wet cell pellet was resuspended at a concentration of 1

g/mL, and then aliquoted into the smaller samples accordingly.  Using a LCQ ion trap

mass spectrometer, the number of proteins identified for each of the different cell pellet

sizes was similar to the 4 g sonication lysis sample, when the samples were desalted

offline by solid phase extraction (Table 7.1).  There is a marked decrease in identification

from samples desalted online, suggesting interference of the samples with the

chromatography system.  The Comprehensive Microbial Resource at TIGR

(www.tigr.org) organizes the *S. oneidensis* MR-1 predicted proteome into 18 functional

category assignments.  By examining the proteins identified in each category, the

distribution of the observed proteome is inferred.  According to this metric, lysing 1 and

5 mg of wet cell pellet provides a more comprehensive identification rate (more proteins

identified) versus sonication in eight of the 18 functional categories including the signal

**Table 7.1.** Protein Identifications Made Using the LCQ Ion Trap

| Organism | Biomass Size | Number of Proteins |
|---|---|---|
| *S. oneidensis* | 4 g | 754 |
| *S. oneidensis* | 50 mg | 700 |
| *S. oneidensis* | 25 mg | 782 |
| *S. oneidensis* | 10 mg | 698 |
| *S. oneidensis* | 10 mg | 502* |
| *S. oneidensis* | 5 mg | 791 |
| *S. oneidensis* | 5 mg | 567* |
| *S. oneidensis* | 1 mg | 820 |
| *S. oneidensis* | 1 mg | 515* |

* Online desalting of peptide mixture.

transduction category (Supplemental Table S1). Proteins involved in signal transduction are of low abundance normally and in the sonication lysis sample, no proteins were identified at the two-peptide level belonging to this category. However, from a cell pellet of 1 mg, a response regulator (SO4003) was identified with two peptides and the 5 mg cell pellet sample yielded the identification of a sensor histidine kinase (SO4173) with two peptides as well. These proteins are of low abundance, as evident by their level of detection here, but they are of great importance to cellular function. Identification of low abundance proteins is vital to understanding cellular responses to growth perturbations.

In order to determine if the resulting identifications were biased against membrane fraction proteins, the following samples were compared against identifications found in membrane fraction proteomes [147, 158]. For the control sonication sample, 60% of the membrane fraction proteins identified in previous studies was identified. The 1 and 5 mg samples did better, with 65-67% of the membrane fraction proteins being identified. However, using the LCQ ion trap without fractionating the samples into soluble and membrane fractions results in a loss of around 40% of the protein identifications. This method of lysis is not well suited for the LCQ ion trap instrumentation platform.

The LTQ linear trapping quadrupole has been described extensively as a more high-throughput instrument [31, 32, 156, 215]. Therefore, the *S. oneidensis* samples were also analyzed by this instrumentation platform (Table 7.2 and Supplemental Table S2). For the control sonication samples, a total of five sample analyses were conducted, yielding from 1229-1333 proteins identified under a single sample analysis and the total number of proteins identified was 1,831. Table 7.2 depicts the combined list from the top

216

**Table 7.2.** Functional Category Distribution for *S. oneidensis* MR-1 LTQ Dataset

| Number | Functional Category | 4 g | 5 mg | 1 mg |
|---|---|---|---|---|
| 1 | Amino Acid Biosynthesis | 58 | 52 | 40 |
| 2 | Biosynthesis of Cofactors, Prosthetic Groups, and Carriers | 86 | 82 | 63 |
| 3 | Cell Envelope | 94 | 94 | 73 |
| 4 | Cellular Processes | 148 | 143 | 122 |
| 5 | Central Intermediary Metabolism | 26 | 27 | 21 |
| 6 | DNA Metabolism | 64 | 67 | 47 |
| 7 | Energy Metabolism | 189 | 182 | 161 |
| 8 | Fatty Acid and Phospholipid Metabolism | 38 | 37 | 35 |
| 9 | Hypothetical Proteins | 375 | 373 | 250 |
| 10 | Mobile and Extrachromosomal Element Functions | 106 | 100 | 93 |
| 11 | Protein Fate | 119 | 115 | 92 |
| 12 | Protein Synthesis | 124 | 116 | 106 |
| 13 | Purines, Pyrimidines, Nucleosides, and Nucleotides | 55 | 55 | 52 |
| 14 | Regulatory Functions | 65 | 71 | 47 |
| 15 | Signal Transduction | 16 | 16 | 5 |
| 16 | Transcription | 41 | 41 | 35 |
| 17 | Transport and Binding Proteins | 70 | 79 | 43 |
| 18 | Unknown Function | 151 | 155 | 115 |
| | **Total** | **1825** | **1805** | **1400** |

three MS experiments. The 5 mg lysis was analyzed in triplicate and yielded protein

identifications from 1282-1472, with a total of 1,805 proteins identified from the three

analyses. The 1 mg lysis was analyzed once and detected 1308 proteins. Note that the

number of proteins identified by the LTQ is more than twice that of the LCQ, therefore it

is reasonable to assume there would be less bias with this instrumentation platform.

Indeed, when analyzing the functional category distribution between the 5 mg

lysis and the control sonication samples (Supplemental Table S2), nine of the eighteen

functional category assignments for the 5 mg sample are represented by equal or greater

number of proteins when compared to the control sonication sample. In fact, the category

of signal transduction contains 16 members for the LTQ dataset. Thirteen of the protein

members are shared between the two lysis techniques, indicating a similar performance

between the two analyses. Proteins localizing to membrane fractions in the previous

studies described above were also analyzed with the LTQ dataset. The results indicate

for the 5 mg lysis sample an overlap of 82% of membrane fraction proteins identified

from previous experiments were found. For the control sonication sample, between 45

and 47% of proteins were identified in the control sonication that localized to the

membrane fraction in previous experiments. The 1 mg lysis sample was intermediate,

with 66-68% of the membrane fraction proteins identified. Therefore, the smaller

biomass (1 and 5 mg) samples performed better with the LTQ than the LCQ with respect

to membrane fraction protein identification, but also performed better than the control

sonication sample for both instrumentation platforms.

The metabolically diverse bacterium *R. palustris* was also utilized to test the

ability of Guanidine HCl as both a lysis and reducing agent. As above, one aliquot was

218

sonicated while the other was used to test Guanidine HCl and TFE (results below). The

LTQ mass spectrometer was utilized in assessing the *R. palustris* sample with a total of

three control sonication samples and three 1 mg Guanidine HCl lysates to be compared.

A total of 1,897 proteins were identified in the 2 g sonicated *R. palustris* cell pellet, and

the 1 mg Guanidine HCl cell lysate yielded 1,797 proteins (Supplemental Table S3). As

above with *S. oneidensis*, there is not a significant difference in protein identifications

with sonication versus Guanidine HCl. However, the significant advantage of Guanidine

HCl appears to be the chemical's ability to efficiently lyse cells and simultaneously

denature the protein complement in a single tube experiment for tryptic digestion and

subsequent mass spectrometry analysis. The *R. palustris* predicted proteome is divided

into 16 functional category assignments, and these are used to compare the bias in

sampling between the two methods. First, the sonicated sample identified 39.4% of the

total predicted proteins in the genome, with unknowns and unclassified, lipid metabolism,

translation, amino acid metabolism, metabolism of cofactors and vitamins, and purine

and pyrimidine metabolism dominating with over 60% of the constituent members being

confidently identified. The 1 mg Guanidine HCl lysates performed similarly, with 37.3%

of the predicted proteins identified and the same functional categories dominated the

identifications.

Therefore, lysing 1-5 mg bacterial isolates with Guanidine HCl gives comparable

and less biased coverage of the resulting proteome, as compared to a 2-4 g control

sonication lysate. The results presented here suggest a new era in "shotgun" proteomics

experiments, where small < 10 mL of culture can be sampled and the subsequent

proteome analyzed. Sampling of a cell culture with a lower cell density leads to

understanding microbial processes and relationships in the environment. Plus, this leads to the ability to analyze environmental samples where the limiting factor will most likely be the amount of starting material present. In order to assess Guanidine HCl as a method for lysing environmental samples, the following acid mine drainage biofilm was tested.

*Comparison of Guanidine HCl versus Trifluoroethylene*

Recently, a similar small-scale approach has been published, highlighting the use of trifluoroethylene (TFE) as a lysing agent for mammalian tissue preparations [266]. TFE has been used in the past for stabilizing bacterial membrane proteins [289]; therefore its use as a lysing agent would be compatible with the system presented in this study. Instead of further denaturing the resulting proteins present in the cell lysates, TFE was added in a higher concentration initially and then diluted 6-fold to minimize interference with trypsin. In our study, *R. palustris* was lysed by sonication, Guanidine HCl, or TFE. A 1 mg wet cell pellet was used in the small sample lysing studies. These samples were analyzed in triplicate on the LTQ mass spectrometer, and the following results were obtained. For the Guanidine HCl lysate, 1334-1467 proteins were identified by mass spectrometry, with a total of 1797 found in the three replicate experiments. The TFE lysate identified between 1566-1635 proteins, with a total of 1992 found in three replicates.

These results were further analyzed in order to determine whether or not a bias in protein identification was present. The experimental reproducibility for these samples was determined; however, due to the nature of the lysing methods (both as lysing agent and denaturant), bias should not be present. Experimental reproducibility is defined as the percentage of proteins from the total list (i.e. 1797) that are identified in each

replicate experiment. Here, the replicate experiment is a technical replicate, since only one data-dependent LC/LC-MS/MS experiment can be performed for each lysate. For the Guanidine HCl lysate, the reproducibility was 58.2% and for the TFE lysate, 60.2%. This demonstrates that reproducible lysing, digestion, and mass spectrometric analysis of the resulting peptides occurs at a similar level between the two lysing methods. Furthermore, as depicted in Table 7.3, comparable identification across all 16 functional categories in the predicted proteome was achieved with these lysing techniques. Using either Guanidine HCl or TFE results in a similar proteome dataset for bacterial isolates.

*Proteome analysis of the AMD biofilm with the single-tube methods*

The Guanidine HCl, TFE, and freeze-grinding single-tube lysis methods as well as the bead-beating method were tested for the AMD biofilm. The community consists of five dominant members from both the bacterial and archeal domains found in a rigid matrix [65]. Table 7.4 illustrates the proteome measurement from various cell pellet sizes with all lysis methods using the algorithm DBDigger for searching. DBDigger was chosen for searching this dataset due to the increased speed of the algorithm for searching MS/MS data against large predicted protein databases [44]. A total of 527 proteins were identified from less than 20 mg of biofilm using the Guanidine HCl method and 391 from a comparable biofilm size with TFE. When a subsequent biofilm (C75) was tested with Guanidine HCl, freeze-grinding, and bead-beating with sample sizes from ~60-250 mg as well as a sonication control of 7.5 g, this resulted in even fewer protein identifications. The average identification rate for the Guanidine HCl method was 14 proteins for the two samples analyzed. The bead-beating was even worse with only up to 9 proteins identified. Freeze-grinding with liquid nitrogen seemed the most promising with an

221

**Table 7.3.** Functional Category Distribution for *R. palustris* LTQ Dataset

| Number | Functional Category | 2 g Sonication | 1 mg Guanidine | 1 mg TFE |
|---|---|---|---|---|
| 1 | Hypothetical | 30 | 44 | 45 |
| 2 | Unknowns and Unclassified | 269 | 284 | 278 |
| 3 | Replication and Repair | 53 | 30 | 40 |
| 4 | Energy Metabolism | 130 | 122 | 148 |
| 5 | Carbon and Carbohydrate Metabolism | 81 | 64 | 78 |
| 6 | Lipid Metabolism | 110 | 95 | 115 |
| 7 | Transcription | 76 | 65 | 70 |
| 8 | Translation | 140 | 133 | 131 |
| 9 | Cellular Processes | 250 | 236 | 279 |
| 10 | Amino Acid Metabolism | 125 | 110 | 122 |
| 11 | General Function Prediction | 206 | 172 | 184 |
| 12 | Metabolism of Cofactors and | 110 | 87 | 100 |
| 13 | Conserved Hypothetical | 77 | 62 | 71 |
| 14 | Transport | 121 | 176 | 193 |
| 15 | Signal Transduction | 76 | 77 | 95 |
| 16 | Purine and Pyrimidine Metabolism | 43 | 40 | 43 |
| | **Total** | **1897** | **1797** | **1992** |

**Table 7.4.** AMD Biofilm Proteome Measurements using DBDigger with the LTQ

| Biofilm Mass | Biofilm | Lysis Method | Protein IDs | Peptide IDs | Spectra |
|---|---|---|---|---|---|
| 29.1 mg | AB | TFE | 391 | 6767 | 10477 |
| 401.2 mg | AB | TFE | 158 | 1292 | 2197 |
| 16.5 mg | AB | Guanidine | 527 | 8681 | 21113 |
| 97.9 mg | AB | Guanidine | 551 | 7950 | 20886 |
| 386.2 mg | AB | Guanidine | 95 | 2514 | 9995 |
| 69.3 mg | C75 | Guanidine | column clogged no information available | | |
| 176.67 mg* | C75 | Guanidine | 15 | 26 | 87 |
| 248.54 mg* | C75 | Guanidine | 13 | 23 | 91 |
| 59.5 mg | C75 | Freeze-grinding | 21 | 34 | 50 |
| 178.43 mg* | C75 | Freeze-grinding | 29 | 55 | 145 |
| 249.88 mg* | C75 | Freeze-grinding | 29 | 55 | 162 |
| 59.7 mg | C75 | Bead-beating | 1 | 2 | 4 |
| 172.49 mg* | C75 | Bead-beating | 2 | 3 | 34 |
| 256.72 mg* | C75 | Bead-beating | 9 | 9 | 160 |
| 7.5 g Run 1 | C75 | Sonication | 2328 | 13197 | 25985 |
| 7.5 g Run 2 | C75 | Sonication | 2097 | 11372 | 22769 |
| 7.5 g Run 3 | C75 | Sonication | 2038 | 10945 | 19695 |

*Used SCX material on the back column only.

average of 26 proteins detected. Initially, the LC column was found to completely clog with the C75 biofilm. A modified method was used where only SCX was loaded on the back column. The goal was to obtain ~2000 proteins out of the approximately 14,000 predicted from the genome annotation [67], which had been identified in previous proteome measurements with larger starting sample [65, 290] as well as for the 7.5 g sonication lysis of the C75 biofilm (Table 7.4). A possible explanation as to why this microbial community did not yield a comprehensive proteome measurement similar to the bacterial isolates for the small scale method is due to the strong cellulose composition of the surrounding matrix (P. Wilmes, personal communication). The rigidity of the biofilm matrix creates a formidable challenge for the lysis procedure, which will require further investigation. One solution would be to enzymatically digest the matrix, thereby releasing the microbial cells. The disadvantage to this method is that the resulting proteome may be different after the incubation period is completed. In addition, by removing the scaffold during this incubation period, the microbial cells may go into shock due to the loss of the supporting structure. This would not allow for identification of low-abundant proteins of metabolic significance.

**Conclusions**

The need to sample complex environmental or biomedical samples with limited biomass has prompted the development of lysis methodologies that are unbiased and can be easily coupled with mass spectrometry-based proteomics as described here and elsewhere [266]. The method described here differs from other lysis techniques as being a true single-tube lysis method. From wet cell pellet to tryptically digested proteins, all steps are performed in the same eppendorf tube. In addition, there is no deleterious effect

on the tryptic digestion step with, as long as the Guanidine HCl is diluted to 1 M or less. The work presented here demonstrates both the ease and utility of performing a novel single-tube lysis method for the proteome-level characterization of bacterial isolates. An advantage to this lysis method is the unbiased means by which an equal coverage of transmembrane-domain and soluble cytosolic proteins can be achieved in the resulting LC-MS experiment. In addition, there is very little difference in the number of proteins identified utilizing 5 mg or less of wet cell pellet, in comparison to 2 g by a traditional sonication lysis technique. The single-tube lysis method must now be optimized further for the AMD biofilm, which illustrates the challenges of the transition from bacterial isolates to complex microbial communities within proteomics. Recently, a study by Wang *et al* [266] described the TFE single-tube lysis method for mouse brain samples in the range of 4.5-5.0 mg in size. Using the accurate mass and time (AMT) tag approach, the authors were able to confidently identify 491 proteins from three replicate lysates [266]. Our results show similar total proteins identified with one of the AMD microbial communities, indicating a possible stoichiometric limitation for single-tube lysis. Stoichiometric limitation as indicated here is specific for complex mixtures, where only the most abundant proteins will be identified from the sample leaving the low-abundant proteins undetectable with current methodologies. For example, far less than the mass weighed for the microbial community is actual biomass, but rather the bulk of the material is the supporting matrix.

# Chapter 8

## Conclusions

Proteomics serves as one of the key lynch-pins for systems biology studies, and has progressed dramatically from its humble beginnings with 2D-PAGE gels. Many labs around the world are employing MS-based proteomic measurements to study biological problems ranging from simple bacteria to complicated human systems. Even with these remarkable advancements and applications of proteomics, there are still many challenges for enhancing this approach to comprehensively characterize biological systems. For example, consider the simple bacteria *S. oneidensis*, which is the focus of this dissertation work. This bacterium has a genome consisting of ~5000 annotated open-reading frames that could be translated into proteins. Our proteome measurements have revealed the presence of ~1800 non-redundant proteins under any single growth condition. The question then arises as to how many proteins are actually translated under a single growth condtion, and whether these proteins can be comprehensively detected by the MS technology employed here. Clearly, the dynamic range of protein expression is large, possibly reaching $10^5$ in bacterial systems, reflecting the range from highly abundant house-keeping proteins (such as ribosomal proteins) to low abundance (< 10 copies per cell) transcription factors. The complexity and dynamic range of the proteome samples provide an enormous analytical challenge to the shotgun MS technique. However, the advancements of on-line multidimensional LC-MS/MS based approaches have almost kept up with these challenges. Returning to our example with *S. oneidensis*, other

226

evidence indicates that this organism may only express about 40% of its proteome under a single growth condition, thereby maintaining a reserve of untapped proteins for other environmental scenarios. This would imply that of the ~2000-2500 proteins expressed under a single condition, our MS-based technology can detect roughly 75-90% of the proteins. This clearly demonstrates that this approach is not merely skimming the surface of the most abundant (and possibly uninteresting) house-keeping proteins, but rather is digging quite deep into the proteome. This provides strong hope for reconstructing detailed pathways and networks of the protein machinery that is operational for a bacteria's life processes.

One interesting aspect of these shotgun proteomics experiments is the detection of a large number of unknown proteins resulting from expression of hypothetic genes. This direct measurement of definitive proteins strongly validates the validity of the genome annotation and aids in helping resolve the realities of gene expression. Without these non-targeted MS-based proteomics experiments, the realities and possible functional roles of these unknown proteins would be completely absent.

Within a shotgun proteomics experiment, many proteins are found to be expressed constitutively; albeit at varying levels under various growth or stress conditions. Therefore, quantitative proteomics allows for the determination of either relative or absolute abundance levels for proteins identified from a whole-cell lysate experiment. These differentially expressed proteins are typically used to compare two different growth conditions for the organism. Quantification allows for the identification of proteins which may be induced or repressed in response to a specific stimulus. For example, a central objective of this dissertation was to explore wild-type *S. oneidensis*

227

versus chromate-exposed cultures, to identify differentially expressed proteins that would provide insight into how this organism deals with metal stress. The measured protein abundances using shotgun proteomics is challenging due to the amino acid distribution within the peptides, the confidence in the calculated abundance value for a given protein, and the labeling method used for a peptide/protein. For this bacterium, most proteins will not vary significantly in abundance, since they are necessary for critical life functions and may not be influenced by the presence of chromate. However, the more interesting proteins will be sensitive to this metal shock, which will impact their abundances. The confidence in abundance measurements is challenging for proteins having a 2-3 fold change in expression level in response to growth changes.

For bacterial systems, there are three standard methods of protein quantification: iTRAQ [1], isotopic labeling, and semiquantitation (or unlabeled quantification). iTRAQ labels the N-terminal amine and the ε-amino group of lysines with one of four tags of the same mass, which works well for bacteria. The resulting MS/MS will contain sequencing information identifying the peptide as well as quantification information from the reporter group giving relative abundances. However, at the time of the studies described here, iTRAQ was not available for general use. Isotopic labeling requires providing an isotopically "heavy" nutrient during growth in order for subsequent quantification of the intact peptide. This method is not amenable for growth conditions chosen for *S. oneidensis* here due to the requirement in using a minimal medium for isotopic labeling of proteins. Therefore, semiquantitation was chosen as the method of quantification here. Its benefits include: no additional sample preparation or special media growth considerations, no requirement for a dedicated mass spectrometry method, and a clear

identification of those proteins that have dramatic changes in expression level. As discussed below, the FDR for these proteins is below 1%, which allows for a confident conclusion that the protein is actually differentially expressed. Ultimately, additional improvements must be made in order to extract meaningful information from smaller changes in protein abundance.

The most important consideration of any aspect of a shotgun proteomic dataset is how the resulting FDR will affect the conclusions drawn. The FDR for the work presented here was set to ~2.0% for peptides, which reflects a calculated protein FDR of ~11.5% at the two-peptide level. However, only ~14% of the total protein identifications in a given proteomics dataset arise from a two-peptide identification. A majority (~77%) of proteins are identified from five or more peptides, which have a calculated FDR of 0.9% for the protein identification. These conservative filter levels increase the confidence of the resulting dataset for proteins identified and the resulting quantification performed. However, more attention must be drawn to the issue of false identifications in shotgun proteomics datasets, because there are a large number of proteins that are subsequently identified and in many cases quantified.

Understanding the role of PTMs in cell cycle control and signaling pathways is critical for providing a complete representation of the Cr(VI) stress response in *S. oneidensis*. In order to predict the presence and location of a PTM, there are a number of principles that must first be understood. In light of the previous drawbacks for investigating PTMs in shotgun proteomics datasets (detailed in Chapter 5), one of the aims in the research performed in the dissertation was the assessment of three separate algorithms which provide the option of searching for PTMs. The FDR for PTM-

containing peptides was found to be ~10% when searching the dataset with InsPecT, which is in contrast to the 50% found utilizing DBDigger.  In addition, InsPecT was found to search the shotgun proteomics dataset in approximately half the time required by DBDigger using a desktop computer.  Another metric to gauge the success of a PTM search is identifying an increased number of proteins.  However, there is not an increase in the total number of proteins identified by adding the option of including PTMs.  There is a slight decrease in the total number of peptides and spectra that are identified by specifying PTMs corresponding to an elevated false negative rate.  The increased thresholds necessary to compensate for the increased FDR in the PTM searches cause an elevation in the false negative rate (a decrease in peptide and spectral count).  A better metric may be to determine if information on the relative modification level of a protein is changed based on the growth condition.  However, much more work is needed in order to better understand how to search and interpret the resulting datasets.

All of the methods development performed in this dissertation played a role in understanding chromium exposure in *S. oneidensis* at the molecular level.  A total of four large-scale proteome studies were performed to assess at the molecular level the response of *S. oneidensis* to chromium exposure during various times and dosages of exposure.  Both technical and biological replicate measurements were made.  The first study performed involved acute chromate exposure in *S. oneidensis* to a sub-lethal concentration for a brief period of time (up to 90 min).  This study was followed by the molecular level characterization of the response of *S. oneidensis* to chromium exposure after complete transformation of Cr(VI) (at the 24 h time point).  Based on proteins found differentially expressed following acute and chronic exposure at fixed Cr(VI)

concentrations, a dosage-response study was undertaken. The final study involved

characterizing both wild type strain MR-1 and a gene knockout mutant strain (Δ2426)

during transformation of Cr(VI) to presumably Cr(III). *so2426* was chosen for deletion

due to the protein and transcript identified as up-regulated following acute shock

exposure under various Cr(VI) dosages.

The first study performed involved understanding acute chromate exposure in *S.*

*oneidensis* to a sub-lethal concentration. Prior to this study, very little was known about

how *S. oneidensis* responded to chromate exposure in culture. Previous work focused on

extremely low concentrations of chromate, which neither caused a stress response nor

were an accurate representation of Cr(VI) concentration in contaminated sites [2-8]. The

first study examined the response of *S. oneidensis* MR-1 to 1 mM of Cr(VI) for 45 or 90

min. A concentration of 1 mM Cr(VI) was chosen based on the growth response curves

of *S. oneidensis* as a function of chromate concentration [9]. No Cr(VI) was reduced in

the acute study; rather the goal of the study was to identify the initial transcriptome and

proteome response of *S. oneidensis.* Semi-quantification was performed on the resulting

proteome dataset, taking into account reproducible differences in sequence coverage,

peptide count, and spectral count. The sulfur metabolism pathway predicted from the

genome sequence of *S. oneidensis* is depicted as a KEGG map in Figure 8.1, where the

protein components highlighted in blue within the pathway were found up-regulated and

the green pathway components were not differentially expressed. However, semi-

quantification is instrument-dependent, where the increased dynamic range of the LTQ

requires a more stringent level of filtering for quantification of proteins detected with this

instrumentation. The benefit of the increased dynamic range of the LTQ allowed for the
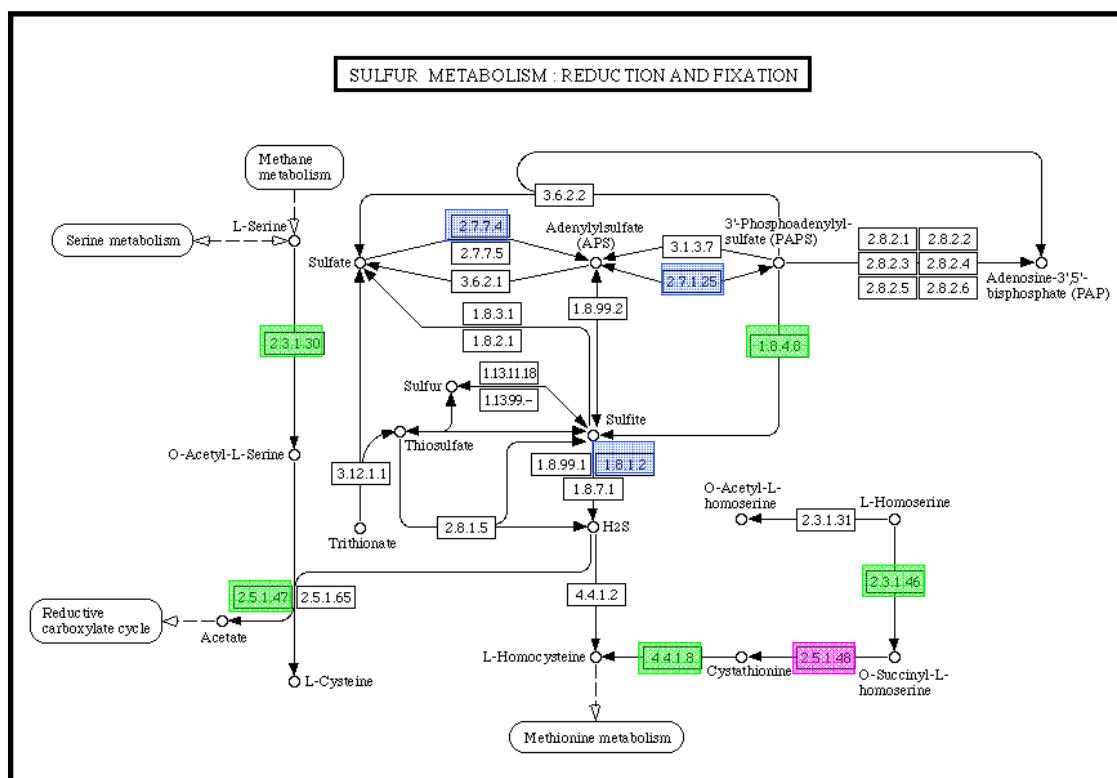
**Figure 8.1.** The sulfur metabolism pathway of *S. oneidensis* MR-1.

Depicted is the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway of sulfur

metabolism (www.genome.jp/kegg/pathway.html).  Proteins shaded in blue were found

up-regulated during Cr(VI) shock, shaded in green were not differentially expressed, and

shaded in pink was not identified in the proteome dataset.

detection and quantification of a putative transcriptional regulator (SO2426) that

appeared to respond overwhelmingly to Cr(VI) in the transcriptome data. Previously, this

level of detection was not possible with the older instrumentation (LCQ). Now, detection

and quantification of low abundance proteins, such as transcriptional regulators, is

possible. The more sensitive detection level leads to a greater corroboration with the

transcriptome data, which in general has a much more sensitive level of detection

If *S. oneidensis* is to be utilized for bioremediation purposes, then understanding the

effects of long-term exposure to environmental factors such as chromate is necessary

prior to implementation of *S. oneidensis* as an agent of bioremediation in contaminated

sites. Therefore, a proteome study was undertaken where *S. oneidensis* was exposed to

0.3 mM Cr(VI) for 24 h correlating to a chronic exposure. This concentration was

chosen based on the growth results of the acute exposure study [9]. The concentration of

1 mM demonstrated a biphasic growth pattern, where there was a long lag period of ~40

h before growth resumed. Whereas a 0.3 mM Cr(VI) concentration would allow the

bacterial cells to recover and resume growth after a lag period between 20-30 h, thus

allowing enough biomass to be produced for the proteome experiments. The proteome

results from the chronic exposure study were very surprising. The assumption prior to

data analysis was that a similar proteome profile to the acute exposure study would be

observed. However, when data analysis was complete, a unique proteome profile was

obtained. The results indicated a high level of stress response in *S. oneidensis* in addition

to the presumable activation of a lytic phage based on up-regulated phage related

proteins. Out of the 76 proteins identified up-regulated after chronic Cr(VI) exposure in

*S. oneidensis*, 36 proteins up-regulated corresponded to regions of the genome known to

harbor phage insertion sites. The chronic exposure work was the first integrated transcriptomic and proteomic study to comprehensively evaluate chronic Cr(V) exposure in *S. oneidensis.*

*S. oneidensis* was exposed to Cr(VI) concentrations of 0, 0.3, 0.5, and 1 mM for 30 min. These doses were also chosen based on growth response curves found in the acute exposure study [9]. In addition, Cr(VI) contaminated sites are known to comprise many different concentrations of Cr(VI), with a gradient ranging from low to highly contaminated regions within a particular site. Understanding whether or not *S. oneidensis* will respond to and reduce Cr(VI) at various concentrations is critical before use of the bacterium as an agent for bioremediation. The use of *S. oneidensis* at sites where the Cr(VI) contamination is minimal would not be useful if the bacterium does not respond to lower concentrations. The time period of exposure response for the dosage-response study and the acute exposure study were comparable, creating the assumption that the proteome profiles for the two studies should be comparable (at least for the shared dosage). This indeed was the case when the profiles were compared. In addition, the overall response for each dosage demonstrated a number of similarities as well. The overall profile indicated that a number of transport and binding proteins, as well as a putative transcriptional regulator (SO2426), were highly up-regulated for all dosages tested. However, a number of proteins were found to demonstrate a dosage-dependent response to Cr(VI) as well. These proteins were found to increase in their relative abundance with increasing dosage and included examples such as transporter proteins and the putative transcriptional regulator (SO2426). The results of the dosage-response

234

study indicated a dosage dependent activation for SO2426, which was chosen for evaluation further as a positive regulator of Cr(VI) gene activation.

Due to results from previous studies involving acute Cr(VI) exposure and the dosage-dependent response of *S. oneidensis*, a proteome study of Cr(VI) transformation was performed. This study sampled time periods during Cr(VI) transformation of both wild-type strain MR-1 and a deletion mutant derived from strain MR-1. The SO2426 putative transcriptional regulator was knocked out in order to test how *S. oneidensis* would respond to Cr(VI) exposure during the transformation period of 4 h in length. SO2426 was chosen for deletion based on the previous studies, which indicated a clear activation of this gene in response to Cr(VI) exposure. SO2426 previously was found to either not be expressed prior to Cr(VI) introduction or displayed a minimal level of expression (~2 peptides identified). Therefore, we knew that SO2426 must play a key role in regulation of protein activation in response to Cr(VI) if this protein is indeed a transcriptional regulator as the genome annotation provided states. Another aspect to this study was the use of biological triplicate cultures for proteome analysis. Very few proteome studies published to date have utilized biological replicates for measurements or for quantification of protein abundance. There is more difficulty in determining protein abundance from biological replicates versus technical replicates. However, the differentially expressed proteins found in a biological replicate are more significant. One caveat of using biological replicates is the overall decrease in reproducibility of proteins identified in comparison to using technical replicates. However, the Cr(VI) transformation study confirmed the hypothesis of the importance of SO2426 in response to Cr(VI) exposure. The Δ2426 cultures demonstrated an impaired ability to transform

Cr(VI) and a number of proteins were identified that indicate importance in Cr(VI) transport across the *S. oneidensis* outer and inner membranes. Conclusions drawn from these proteome experiments indicate (1) that Cr(VI) may be a competitive inhibitor of sulfate, (2) Cr(III) may be more toxic to *S. oneidensis* than Cr(VI), (3) proteome results demonstrated a dosage-dependent stress response to Cr(VI), and (4) the characterization of a putative metal response regulator (SO2426), which appears to be a global regulator of not only Cr(VI), but a number of other transition metals.

The evidence for competitive inhibition of sulfate by chromate arises from the induction of proteins annotated as involved in sulfate uptake and metabolism. In the acute shock study, three of the nine sulfur metabolism pathway components were identified up-regulated (Figure 8.1). The functions of the up-regulated components within in the pathway are enzymes that catalyze production of sulfate from other sulfur-containing compounds, which indicates an intracellular sulfate deficiency. In addition, there are two sulfate ABC transporters annotated in the *S. oneidensis* genome (*so3599-so3602* and *so4652-so4655*) and members of both transporters are found up-regulated during exposure to Cr(VI). Finally, four members of the cysteine biosynthesis pathway were found up-regulated following exposure to 1 mM Cr(VI). If sulfate deficiency were not an issue, *S. oneidensis* would not need to increase the production of these proteins. There may also be similarities between the hydration radius of chromate and sulfate, which leads to the chromate interacting with the sulfate transporters. Competitive inhibition of sulfate could be caused by more Cr(VI) present in the media, which outcompetes sulfate for binding at the ABC transporters. By increasing the sulfate concentration, one may be able to rescue *S. oneidensis* from chromate shock.

Prior to the studies presented here, the consensus in the field of microbiology was that the Cr(VI) species led to the toxicity phenotype observed when microbes are exposed to chromium [6, 10-18]. After *S. oneidensis* was grown in the presence of Cr for 24, all Cr(VI) was presumably transformed to Cr(III). Cr(VI) and Cr(III) are the two stable forms of chromium commonly found in the environment [12]. The conclusion drawn from the chronic study is that Cr(III) may in fact be more toxic than Cr(VI). In this study, the proteome profile observed was unique in comparison to the other Cr(VI) exposure proteome studies reported. The activation of two lytic prophages appeared to be the signature response of *S. oneidensis* after the transformation of Cr(VI) to presumably Cr(III). The presence of Cr(III) was not tested; as well as there was not a noticeable precipitate reported in the media. The proteome dataset allowed for the confirmation of the lytic phage cycle by providing confirmation of translation of these genes by seizing the host cell's translation apparatus. The results of this chronic exposure study were similar to the transcriptome results reported for *S. oneidensis* exposure to both ionizing radiation and artificial UV radiation [19, 20]. The actual response of *S. oneidensis* towards long-term exposure to Cr could be deciphered by inactivation of the integrated phages or the use of a strain that does not have an integrated phage. Also, one could test the mutagenesis rate of intracellular Cr(III) by a whole genome microarray as shown for *E. coli* [21]. This could detect the mutation rate of the *S. oneidensis* chromosome, thereby determining if critical house-keeping genes have nucleotide sequence changes following long-term exposure to Cr(III)-containing products. In addition, O'Brian *et al* [22] demonstrated that Cr interacts with the phosphate backbone of double-stranded DNA molecules. This interaction will cause the

DNA replication machinery to stall, with DNA replication arrest in cells that are actively reducing Cr(VI). The stalled replication machinery may be one of the reasons as to the observed phenotype of elongated, filamentous bacterial cells at the 24 h sampling point.

In addition to the physiological responses just described, we found that upon Cr(VI) exposure there was evidence of stress response from the expressed proteome datasets. Characteristic proteins found activated in response to Cr(VI) intracellular stress include RecN, TopB, SO1648 (cold shock domain family protein), IbpA, and ClpB. There was a dosage-dependent response found for stress response proteins (see Chapter 4 for details). The dosage-dependence of stress response-related proteins indicates a possible tolerance level to Cr(VI), which upon increased concentrations of Cr(VI) there is a trigger that indicates increased intracellular stress. However, at this time there is not an indication as to the identity of the protein(s) responsible for activating the stress response demonstrated in the acute shock or dosage response studies.

Finally, the characterization of a putative DNA-binding response regulator, SO2426, as a metal response regulator was the culmination of the proteome studies on Cr exposure. Prior to the acute Cr(VI) shock study, there was no evidence that SO2426 was translated into protein or was induced at the transcriptome and proteome level under Cr(VI) conditions. Further work determined that SO2426 putatively controlled gene expression at the transcript level of genes involved with sulfate metabolism and iron homeostasis after creation of a knockout of *so2426* [23]. The same proteins identified as down-regulated in the mutant cultures both prior to and following Cr(VI) addition were found up-regulated during acute Cr(VI) shock; indicating that the gene expression is controlled by an actively expressed and translated *so2426*. This work demonstrates the

238

utility and value of performing global non-targeted proteome studies to identify putative roles for proteins expressed under given growth conditions. In addition, the expectation is that information gleaned from the Cr(VI) proteome studies will provide fodder for future proteome studies in environmental microbiology to better understand the complex role of Cr(VI) biogeochemical cycling in the environment.

Measurements using mass spectrometry based proteomics have enhanced the scientific knowledge on microbial and eukaryotic systems. Research in the areas of microbial ecology and human disease would not have progressed as quickly without global proteome studies. First, comprehensive proteome studies provide novel non-targeted biochemical information on protein presence and abundance. In addition, the area of shotgun proteomics is helping to validate genome annotation. The large number of hypothetical genes annotated led many individuals to question the validity of genome annotation. Therefore, with the advent of shotgun proteomics, the identification and in some cases tentative functional assignments can be given to hypothetical proteins.

The research to date is very helpful, but also can be overwhelming. Repositories for MS/MS data must be created to organize, standardize, and increase the availability of information obtained from the large-scale proteome studies. HUPO, the human proteome project, has begun the process of providing guidelines for proteome datasets [24]. The organizers were visionary, because they realized both the value of the proteome data and the need for community wide availability of the information. In the future, other areas must following the HUPO example, otherwise the field of proteomics will suffer the same problems as faced when nucleotide sequencing of the human genome was

239

completed [25-28].  Some of these issues include organizing the large amount of sequencing information and dissemination of the genome sequence.

Since the induction of routine proteome measurements, the scientific questions posed have become more complex.  Previously, measurements on pure microbial cultures and human cell lines were considered to be significant studies.  However, as shotgun proteomics experiments have become more routine, there is an increasing interest in investigating microbial environmental communities [29-31] and complex mammalian tissues [32-35].  The current platform is limited to relatively abundant members within a particular community; therefore there is an increased concern of the overall dynamic range capabilities of the experimental platform.  In order to compensate for the current dynamic range limitations, improvements in chromatographic separation and the mass spectrometry instruments will be needed.  The mass spectrometer duty cycle is more efficient now than five years ago; however, many peptides/proteins are still not sampled during the experiments.  Low abundant peptides/proteins sampling will improve by adding a third dimension of separation.  Without dramatic advancements to the instrumentation, the goal of detecting proteins found at reduced expression levels is not possible.  In the end, an enrichment of low-abundant community members will be necessary to understand their contribution.

# List of References

1.      Ideker, T., et al., *Integrated genomic and proteomic analyses of a systematically perturbed metabolic network.* Science, 2001. **292**: p. 929-934.

2.      O'Farrell, P.H., *High resolution two-dimensional electrophoresis of proteins.* J Biol Chem, 1975. **250**(10): p. 4007-21.

3.      Lovell-Badge, R.H. and M.J. Evans, *Changes in protein synthesis during differentiation of embryonal carcinoma cells, and a comparison with embryo cells.* J Embryol Exp Morphol, 1980. **59**: p. 187-206.

4.      Krakauer, T., et al., *Analysis of the heterogeneity of the mouse H-2K, D, and L gene products.* J Immunol, 1980. **124**(5): p. 2149-56.

5.      Niall, H.D. and P. Edman, *Two structurally distinct classes of kappa-chains in human immunoglobulins.* Nature, 1967. **216**(5112): p. 262-3.

6.      Edman, P. and G. Begg, *A protein sequenator.* European Journal of Biochemistry, 1967. **1**(1): p. 80-91.

7.      Fales, H.M., G.W. Milne, and M.L. Vestal, *Chemical ionization mass spectrometry of complex molecules.* Journal of the American Chemical Society, 1969. **91**(13): p. 3682-5.

8.      de Hoffman, E. and V. Stroobant, *1.  Ion Sources*, in *Mass Spectrometry: Principles and Applications*. 2002, John Wiley & Sons, LTD: New York. p. 11-61.

9.      Hillenkamp, F., et al., *Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers.* Analytical Chemistry, 1991. **63**(24): p. 1193A-1203A.

10.     Berndt, P., U. Hobohm, and H. Langen, *Reliable automatic protein identification from matrix-assisted laser desorption/ionization mass spectrometric peptide fingerprints.* Electrophoresis, 1999. **20**(18): p. 3521-6.

11.     Laiko, V.V., M.A. Baldwin, and A.L. Burlingame, *Atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry.* Analytical Chemistry, 2000. **72**(4): p. 652-7.

12.     Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.

13.     Siuzdak, G., *The emergence of mass spectrometry in biochemical research.* PNAS, 1994. **91**(24): p. 11290-7.

14.     Canas, B., et al., *Mass spectrometry technologies for proteomics.* Briefings in Functional Genomics and Proteomics, 2006. **4**(4): p. 295-320.

15.     Peng, J. and S.P. Gygi, *Proteomics:  the move to mixtures.* Journal of Mass Spectrometry, 2001. **36**: p. 1083-1091.

16.     Tyers, M. and M. Mann, *From genomics to proteomics.* Nature, 2003. **422**: p. 193-197.

17.     Rompp, A., et al., *Examples of Fourier transform ion cyclotron resonance mass spectrometry developments: from ion physics to remote access biochemical mass spectrometry.* Eur J Mass Spectrom (Chichester, Eng), 2005. **11**(5): p. 443-56.

18.     Guan, S., A.G. Marshall, and M.C. Wahl, *MS/MS with high detection efficiency and mass resolving power for product ions in Fourier transform ion cyclotron resonance mass spectrometry.* Anal Chem, 1994. **66**(8): p. 1363-7.

19.	Wolters, D.A., M.P. Washburn, and J.R. Yates III, *An automated multidimensional protein identification technology for shotgun proteomics.* Analytical Chemistry, 2001. **73**: p. 5683-5690.

20.	Liu, H., R.G. Sadygov, and J.R. Yates, 3rd, *A model for random sampling and estimation of relative protein abundance in shotgun proteomics.* Analytical Chemistry, 2004. **76**(14): p. 4193-201.

21.	Marshall, A., *Fourier transform ion cyclotron resonance mass spectrometry.* Acc. Chem. Res., 1985. **18**: p. 316-322.

22.	Syka, J.E.P., et al., *Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications.* Journal of Proteome Research, 2004. **3**: p. 621-626.

23.	Hu, Q., et al., *The Orbitrap:  a new mass spectrometer.* Journal of Mass Spectrometry, 2005. **40**: p. 430-443.

24.	Sundqvist, B., et al., *Californium-252 plasma desorption time of flight mass spectroscopy of proteins.* Biomed Mass Spectrom, 1984. **11**(5): p. 242-57.

25.	VerBerkmoes, N.C., et al., *Integrating "Top-Down" and "Bottom-Up" mass spectrometric approaches for proteomic analysis of Shewanella oneidensis.* Journal of Proteome Research, 2002. **1**: p. 239-252.

26.	Pandey, A. and M. Mann, *Proteomics to study genes and genomes.* Nature, 2000. **405**: p. 837-846.

27.     Chernushevich, I.V., A.V. Loboda, and B.A. Thomson, *An introduction to quadrupole-time-of-flight mass spectrometry.* J Mass Spectrom, 2001. **36**(8): p. 849-65.

28.     Hopfgartner, G., et al., *Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules.* J Mass Spectrom, 2004. **39**(8): p. 845-55.

29.     Stafford Jr., G.C., et al., *Recent improvements in and analytical applications of advanced ion trap technology.* International Journal of Mass Spectrometry and Ion Processes, 1984. **60**: p. 85-98.

30.     Schwartz, J.C., M. Senko, and J.E.P. Syka, *A two-dimensional quadrupole ion trap mass spectrometer.* Journal of the American Society for Mass Spectrometry, 2002. **13**: p. 659-669.

31.     Blackler, A.R., et al., *Quantitative comparison of proteomic data quality between a 2D and 3D quadrupole ion trap.* Analytical Chemistry, 2006. **78**(4): p. 1337-44.

32.     Riter, L.S., et al., *Comparison of the Paul ion trap to the linear ion trap for use in global proteomics.* Proteomics, 2006. **6**: p. 1735-1740.

33.     Washburn, M.P., D.A. Wolters, and J.R. Yates III, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nature Biotechnology, 2001. **19**: p. 242-247.

34.     Cargile, B.J., et al., *Gel based isoelectric focusing of peptides and the utility of isoelectric point in protein identification.* Journal of Proteome Research, 2004. **3**(1): p. 112-9.

245

35.     Essader, A.S., et al., *A comparison of immobilized pH gradient isoelectric focusing and strong-cation-exchange chromatography as a first dimension in shotgun proteomics.* Proteomics, 2005. **5**(1): p. 24-34.

36.     Courchesne, P.L., et al., *Optimization of capillary chromatography ion trap-mass spectrometry for identification of gel-separated proteins.* Electrophoresis, 1998. **19**(6): p. 956-67.

37.     Figeys, D., et al., *Data-dependent modulation of solid-phase extraction capillary electrophoresis for the analysis of complex peptide and phosphopeptide mixtures by tandem mass spectrometry: application to endothelial nitric oxide synthase.* Anal Chem, 1999. **71**(13): p. 2279-87.

38.     Lipton, M.S., et al., *AMT tag approach to proteomic characterization of Deinococcus radiodurans and Shewanella oneidensis.* Methods of Biochemical Analysis, 2006. **49**: p. 113-34.

39.     Eng, J.K., A.L. McCormack, and J.R. Yates III, *An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database.* Journal of the American Society for Mass Spectrometry, 1994. **5**: p. 976-989.

40.     Tanner, S., et al., *InsPecT: identification of posttranslationally modified peptides from tandem mass spectra.* Analytical Chemistry, 2005. **77**(14): p. 4626-39.

41.     Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-67.

42. Tabb, D.L., C.G. Fernando, and M.C. Chambers, *MyriMatch: Highly accurate tandem mass spectral peptide identification by multivariate hyergeometric analysis.* Journal of Proteome Research, 2007: p. [8].

43. Geer, L.Y., et al., *Open mass spectrometry search algorithm.* Journal of Proteome Research, 2004. **3**(5): p. 958-64.

44. Tabb, D.L., et al., *DBDigger: reorganized proteomic database identification that improves flexibility and speed.* Analytical Chemistry, 2005. **77**(8): p. 2464-74.

45. VerBerkmoes, N.C., et al., *Mass spectrometric approaches for characterizing bacterial proteomes.* Expert Rev Proteomics, 2004. **1**(4): p. 433-47.

46. Mirzaei, H. and F. Regnier, *Identification of yeast oxidized proteins: chromatographic top-down approach for identification of carbonylated, fragmented, and cross-linked proteins in yeast.* Journal of Chromatography A, 2007. **1141**: p. 22-31.

47. Gianazza, E., J. Crawford, and I. Miller, *Detecting oxidative post-translational modifications in proteins.* Amino Acids, 2006.

48. Strader, M.B., et al., *Characterization of the 70S ribosome from Rhodopseudomonas palustris using an integrated "top-down" and "bottom-up" mass spectrometric approach.* Journal of Proteome Research, 2004. **3**: p. 965-978.

49. Forbes, A.J., et al., *Targeted analysis and discovery of posttranslational modifications in proteins from methanogenic archaea by top-down MS.* PNAS, 2004. **101**: p. 2678-2683.

50. Mann, M. and O.N. Jensen, *Proteomic analysis of post-translational modifications.* Nature Biotechnology, 2003. **21**: p. 255-261.

51.    MacCoss, M.J., et al., *Shotgun identification of protein modifications from protein complexes and lens tissue.* PNAS, 2002. **99**(12): p. 7900-5.

52.    Stock, J.B., A.M. Stock, and J.M. Mottonen, *Signal transduction in bacteria.* Nature, 1990. **344**(6265): p. 395-400.

53.    Eisenbach, M., *Control of bacterial chemotaxis.* Molecular Microbiology, 1996. **20**(5): p. 903-10.

54.    Baker, M.D., P.M. Wolanin, and J.B. Stock, *Systems biology of bacterial chemotaxis.* Current Opinion in Microbiology, 2006. **9**(2): p. 187-92.

55.    Olsen, J.V., et al., *Global, in vivo, and site-specific phosphorylation dynamics in signaling networks.* Cell, 2006. **127**(3): p. 635-48.

56.    Cantin, G.T. and J.R. Yates, 3rd, *Strategies for shotgun identification of post-translational modifications by mass spectrometry.* Journal of Chromatography A, 2004. **1053**(1-2): p. 7-14.

57.    Blagoev, B., et al., *Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics.* Nature Biotechnology, 2004. **22**(9): p. 1139-45.

58.    Garcia, B.A., J. Shabanowitz, and D.F. Hunt, *Analysis of protein phosphorylation by mass spectrometry.* Methods, 2005. **35**(3): p. 256-64.

59.    Thompson, M.R., D.K. Thompson, and R.L. Hettich, *Systematic Assessment of the Benefits and Caveats in Mining Microbial Post-Translational Modifications from Shotgun Proteomic Data; Response of Shewanella oneidensis to Chromate Exposure.* Journal of Proteome Research, 2007. **Accepted.**

60. Gupta, N., et al., *Whole proteome analysis of post-translational modifications: Applications of mass-spectrometry for proteogenomic annotation.* Genome Research, 2007.

61. Myers, C.R. and K.H. Nealson, *Bacterial manganese reduction and growth with manganese oxide as the sole electron acceptor.* Science, 1988. **240**: p. 1319-1320.

62. Nealson, K.H. and D. Saffarini, *Iron and manganese in anaerobic respiration: environmental significance, physiology, and regulation.* Annu Rev Microbiol, 1994. **48**: p. 311-43.

63. Ding, Y.H., et al., *The proteome of dissimilatory metal-reducing microorganism Geobacter sulfurreducens under various growth conditions.* Biochimica et Biophysica Acta, 2006. **1764**(7): p. 1198-206.

64. Nealson, K.H. and C.R. Myers, *Iron reduction by bacteria:  A potential role in the genesis of banded iron formations.* American Journal of Science, 1990. **290-A**: p. 35-45.

65. Ram, R.J., et al., *Community proteomics of a natural microbial biofilm.* Science, 2005. **308**: p. 1915-1920.

66. Arnold, R.G., et al., *Regulation of Dissimilatory Fe(III) Reduction Activity in Shewanella putrefaciens.* Applied and Environmental Microbiology, 1990. **56**(9): p. 2811-2817.

67. Tyson, G.W., et al., *Community structure and metabolism through reconstruction of microbial genomes from the environment.* Nature, 2004. **428**: p. 37-43.

68.     Larimer, F., et al., *Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodopseudomonas palustris.* Nature Biotechnology, 2004. **22**: p. 55-61.

69.     Heidelberg, J.F., et al., *Genome sequence of the dissimilatory metal ion-reducing bacterium Shewanella oneidensis.* Nature Biotechnology, 2002. **20**: p. 1118-1123.

70.     Frazier, M.E., et al., *Realizing the potential of the genome revolution:  the genomes to life program.* Science, 2003. **300**: p. 290-293.

71.     Seshadri, R., et al., *Genome sequence of the PCE-dechlorinating bacterium Dehalococcoides ethenogenes.* Science, 2005. **307**(5706): p. 105-8.

72.     Parkhill, J., et al., *Genome sequence of Yersinia pestis, the causative agent of plague.* Nature, 2001. **413**(6855): p. 523-7.

73.     Rondon, M.R., et al., *Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms.* Applied and Environmental Microbiology, 2000. **66**(6): p. 2541-7.

74.     Venter, J.C., et al., *Environmental genome shotgun sequencing of the Sargasso Sea.* Science, 2004. **304**(5667): p. 66-74.

75.     Mongodin, E.F., J.B. Emerson, and K.E. Nelson, *Microbial metagenomics.* Genome Biol, 2005. **6**(10): p. 347.

76.     Martin, H.G., et al., *Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities.* Nature Biotechnology, 2006.

77.     Tringe, S.G., et al., *Comparative metagenomics of microbial communities.* Science, 2005. **308**(5721): p. 554-7.

78.    Gill, S.R., et al., *Metagenomic analysis of the human distal gut microbiome.* Science, 2006. **312**(5778): p. 1355-9.

79.    Myers, C.R. and K.H. Nealson, *Respiration-linked proton translocation coupled to anaerobic reduction of manganese(IV) and iron(III) in Shewanella putrefaciens MR-1.* Journal of Bacteriology, 1990. **172**(11): p. 6232-8.

80.    Venkateswaren, K., et al., *Polyphasic taxonomy of the genus Shewanella and description of Shewanella oneidensis sp. nov.* International Journal of Systematic Bacteriology, 1999. **49**: p. 705-724.

81.    Myers, C.R. and J.M. Myers, *Role of menaquinone in the reduction of fumarate, nitrate, iron (III) and manganese (IV) by Shewanella putrefaciens MR-1.* FEMS Microbiology Letters, 1993. **114**: p. 215-222.

82.    Petrovskis, E.A., T.M. Vogel, and P. Adriaens, *Effects of electron acceptors and donors on transformation of tetrachloromethane by Shewanella putrefaciens MR-1.* FEMS Microbiology Letters, 1994. **121**(3): p. 357-63.

83.    Saffarini, D.A., et al., *Anaerobic respiration of Shewanella putrefaciens requires both chromosomal and plasmid-borne genes.* FEMS Microbiology Letters, 1994. **119**: p. 271-278.

84.    Scott, J.H. and K.H. Nealson, *A biochemical study of the intermediary carbon metabolism of Shewanella putrefaciens.* Journal of Bacteriology, 1994. **176**(11): p. 3408-11.

85.    Moser, D.P. and K.H. Nealson, *Growth of the facultative anaerobe Shewanella putrefaciens by elemental sulfur reduction.* Applied and Environmental Microbiology, 1996. **62**(6): p. 2100-5.

86.    Myers, C.R., et al., *Chromium(VI) reductase activity is associated with the cytoplasmic membrane of anaerobically grown Shewanella putrefaciens MR-1.* Journal of Applied Microbiology, 2000. **88**: p. 98-106.

87.    Krause, B. and K.H. Nealson, *Physiology and enzymology involved in denitrification by Shewanella putrefaciens.* Applied and Environmental Microbiology, 1997. **63**(7): p. 2613-8.

88.    Nealson, K.H., D.P. Moser, and D.A. Saffarini, *Anaerobic electron acceptor chemotaxis in Shewanella putrefaciens.* Applied and Environmental Microbiology, 1995. **61**(4): p. 1551-4.

89.    Myers, C.R. and K.H. Nealson. *Iron mineralization by bacteria:  metabolic coupling of iron reduction to cell metabolism in Alteromonas putrefaciens strain MR-1*. in *Conference on Iron Biominerals*. 1990. University of New Hampshire: New York:  Plenum Press.

90.    Nealson, K.H. and C.R. Myers, *Microbial reduction of manganese and iron: new approaches to carbon cycling.* Applied and Environmental Microbiology, 1992. **58**(2): p. 439-43.

91.    Serres, M.H. and M. Riley, *Structural domains, protein modules, and sequence similarities enrich our understanding of the Shewanella oneidensis MR-1 proteome.* Omics, 2004. **8**(4): p. 306-21.

92.    Bencharit, S. and M.J. Ward, *Chemotactic responses to metals and anaerobic electron acceptors in Shewanella oneidensis MR-1.* Journal of Bacteriology, 2005. **187**(14): p. 5049-53.

93.     Saffarini, D.A. and K.H. Nealson, *Sequence and genetic characterization of etrA, an fnr analog that regulates anaerobic respiration in Shewanella putrefaciens MR-1.* Journal of Bacteriology, 1993. **175**(24): p. 7938-44.

94.     Maier, T.M. and C.R. Myers, *Isolation and characterization of a Shewanella putrefaciens MR-1 electron transport regulator etrA mutant: reassessment of the role of EtrA.* Journal of Bacteriology, 2001. **183**(16): p. 4918-26.

95.     Beliaev, A.S., et al., *Microarray transcription profiling of a Shewanella oneidensis etrA mutant.* Journal of Bacteriology, 2002. **184**(16): p. 4612-6.

96.     Thompson, D.K., et al., *Transcriptional and proteomic analysis of a ferric uptake regulator (Fur) mutant of Shewanella oneidensis:  possible involvement of fur in energy metabolism, transcriptional regulation, oxidative stress.* Applied and Environmental Microbiology, 2002. **68**: p. 881-892.

97.     Wan, X.F., et al., *Transcriptomic and proteomic characterization of the Fur modulon in the metal-reducing bacterium Shewanella oneidensis.* Journal of Bacteriology, 2004. **186**(24): p. 8385-400.

98.     Myers, C.R. and K.H. Nealson, *Microbial reduction of manganese oxides: Interactions with iron and sulfur.* Geochimica et Cosmochimica Acta, 1988. **52**: p. 2727-2732.

99.     Lovley, D.R., E.J. Phillips, and D.J. Lonergan, *Hydrogen and Formate Oxidation Coupled to Dissimilatory Reduction of Iron or Manganese by Alteromonas putrefaciens.* Applied and Environmental Microbiology, 1989. **55**(3): p. 700-706.

100.    Myers, C.R. and J.M. Myers, *Localization of cytochromes to the outer membrane of anaerobically grown Shewanella putrefaciens MR-1*. Journal of Bacteriology, 1992. **174**(11): p. 3429-38.

101.    Myers, C.R. and J.M. Myers, *Fumarate reductase is a soluble enzyme in anaerobically grown Shewanella putrefaciens MR-1*. FEMS Microbiology Letters, 1992. **98**: p. 13-20.

102.    Myers, C.R. and J.M. Myers, *Ferric reductase is associated with the membranes of anaerobically grown Shewanella putrefaciens MR-1*. FEMS Microbiology Letters, 1993. **108**: p. 15-22.

103.    Myers, C.R. and J.M. Myers, *Ferric iron reduction-linked growth yields of Shewanella putrefaciens MR-1*. The Journal of Applied Bacteriology, 1994. **76**(3): p. 253-8.

104.    Myers, C.R. and J.M. Myers, *Cloning and sequence of cymA, a gene encoding a tetraheme cytochrome c required for reduction of iron(III), fumarate, and nitrate by Shewanella putrefaciens MR-1*. Journal of Bacteriology, 1997. **179**(4): p. 1143-52.

105.    Myers, C.R. and J.M. Myers, *Outer membrane cytochromes of Shewanella putrefaciens MR-1: spectral analysis, and purification of the 83-kDa c-type cytochrome*. Biochimica et Biophysica Acta, 1997. **1326**(2): p. 307-18.

106.    Beliaev, A.S. and D.A. Saffarini, *Shewanella putrefaciens mtrB encodes an outer membrane protein required for Fe(III) and Mn(IV) reduction*. Journal of Bacteriology, 1998. **180**(23): p. 6292-6297.

254

107. Myers, J.M. and C.R. Myers, *Isolation and sequence of omcA, a gene encoding a decaheme outer membrane cytochrome c of Shewanella putrefaciens MR-1, and detection of omcA homologs in other strains of S. putrefaciens.* Biochimica et Biophysica Acta, 1998. **1373**(1): p. 237-51.

108. Leys, D., et al., *Structure and mechanism of the flavocytochrome c fumarate reductase of Shewanella putrefaciens MR-1.* Nature Structural Biology, 1999. **6**(12): p. 1113-7.

109. Myers, J.M. and C.R. Myers, *Role of the tetraheme cytochrome CymA in anaerobic electron transport in cells of Shewanella putrefaciens MR-1 with normal levels of menaquinone.* Journal of Bacteriology, 2000. **182**(1): p. 67-75.

110. Beliaev, A.S., et al., *MtrC, an outer membrane decahaem c cytochrome required for metal reduction in Shewanella putrefaciens MR-1.* Molecular Microbiology, 2001. **39**(3): p. 722-30.

111. Tsapin, A.I., et al., *Purification and properties of a low-redox-potential tetraheme cytochrome c3 from Shewanella putrefaciens.* Journal of Bacteriology, 1996. **178**(21): p. 6386-8.

112. Myers, C.R. and J.M. Myers, *Isolation and characterization of a transposon mutant of Shewanella putrefaciens MR-1 deficient in fumarate reductase.* Letters in Applied Microbiology, 1997. **25**(3): p. 162-8.

113. Myers, J. and C. Myers, *Role for outer membrane cytochromes OmcA and OmcB of Shewanella putrefaciens MR-1 in reduction of manganese dioxide.* Applied and Environmental Microbiology, 2001. **67**: p. 260-269.

114. Tsapin, A.I., et al., *Identification of a small tetraheme cytochrome c and a flavocytochrome c as two of the principal soluble cytochromes c in Shewanella oneidensis strain MR1.* Applied and Environmental Microbiology, 2001. **67**(7): p. 3236-44.

115. Saffarini, D.A., S.L. Blumerman, and K.J. Mansoorabadi, *Role of menaquinones in Fe(III) reduction by membrane fractions of Shewanella putrefaciens.* Journal of Bacteriology, 2002. **184**(3): p. 846-8.

116. Maier, T.M. and C.R. Myers, *The outer membrane protein Omp35 affects the reduction of Fe(III), nitrate, and fumarate by Shewanella oneidensis MR-1.* BMC Microbiology, 2004. **4**: p. 23.

117. Meyer, T.E., et al., *Identification of 42 possible cytochrome C genes in the Shewanella oneidensis genome and characterization of six soluble cytochromes.* Omics, 2004. **8**(1): p. 57-77.

118. Bouhenni, R., A. Gehrke, and D. Saffarini, *Identification of genes involved in cytochrome c biogenesis in Shewanella oneidensis, using a modified mariner transposon.* Applied and Environmental Microbiology, 2005. **71**: p. 4935-4937.

119. Yang, F., et al., *Characterization of purified c-type heme-containign peptides and identification of c-type heme-attachemnt sites in Shewanella oneidensis cytochromes using mass spectrometry.* Journal of Proteome Research, 2005. **4**: p. 846-854.

120. Lower, S.K., M.F. Hochella Jr., and T.J. Beveridge, *Bacterial recognition of mineral surfaces:  nanoscale interactions between Shewanella and $\alpha$-FeOOH.* Science, 2001. **292**: p. 1360-1363.

121. James, B.R., *The challenge of remediating chromium-contaminated soil.* Environmental Science & Technology, 1996. **30**: p. A248-A251.

122. Langard, S., *Metals in the Environment.* 1980, Academy Press Inc.: New York. p. 111-132.

123. Riley, R.G., J.M. Zachara, and F.J. Wobber. Report DOE/ER-0547T, Chemical Contaminants on DOE Lands and Selection of Contaminant Mixtures for Subsurface Science Research. 1992, Washington, D.C.: United States Department of Energy.

124. Palmer, C.D. and P.R. Wittbrodt, *Processes affecting the remediation of chromium-contaminated sites.* Environmental Health Perspectives, 1991. **92**: p. 25-40.

125. Shupack, S.I., *The chemistry of chromium and some resulting analytical problems.* Environ Health Perspect, 1991. **92**: p. 7-11.

126. Cervantes, C., et al., *Interactions of chromium with microorganisms and plants.* FEMS Microbiology Reviews, 2001. **25**: p. 335-347.

127. Dubiel, M., et al., *Microbial iron respiration can protect steel from corrosion.* Applied and Environmental Microbiology, 2002. **68**: p. 1440-1445.

128. Park, C.H., et al., *Remediation and Beneficial Reuse of Contaminated Sediments*, in *Batelle Press*. 2002: Columbus, OH. p. 103-111.

129. Brakstad, O.G. and A.G. Lodeng, *Microbial diversity during biodegradation of crude oil in seawater from the North Sea.* Microbial Ecology, 2005. **49**(1): p. 94-103.

130. de Lorenzo, V., *Blueprint of an oil-eating bacterium.* Nature Biotechnology, 2006. **24**(8): p. 952-3.

131. Nyman, J.L., et al., *Biogeochemical elimination of chromium(VI) from contaminated water.* Biorem. J., 2002. **6**: p. 39-55.

132. Bencheikh-Latmani, R., et al., *Global transcriptional profiling of Shewanella oneidensis MR-1 during Cr(VI) and U(VI) reduction.* Applied and Environmental Microbiology, 2005. **71**: p. 7453-7460.

133. Viamajala, S., et al., *Chromate/nitrite interactions in Shewanella oneidensis MR-1: Evidence for multiple hexavalent chromium [Cr(VI)] reduction mechanisms dependent on physiological growth conditions.* Biotechnology and Bioengineering, 2002. **78**: p. 770-778.

134. Daulton, T.L., et al., *Electron energy loss spectroscopy techniques for the study of microbial chromium(VI) reduction.* Journal of Microbiological Methods, 2002. **50**: p. 39-54.

135. Middleton, S.S., et al., *Cometabolism of Cr(VI) by Shewanella oneidensis MR-1 produces cell-associated reduced chromium and inhibits growth.* Biotechnology and Bioengineering, 2003. **83**(6): p. 627-37.

136. Wielinga, B., et al., *Iron promoted reduction of chromate by dissimilatory iron-reducing bacteria.* Environmental Science & Technology, 2001. **35**: p. 522-527.

137. Caccavo, F., N.B. Ramsing, and J.W. Costerton, *Morphological and metabolic responses to starvation by the dissimilatory metal-reducing bacterium Shewanella alga BrY.* Applied and Environmental Microbiology, 1996. **62**: p. 4678-4682.

138. Park, C.H., et al., *Purification to homogeneity and characterization of a novel Pseudomonas putida chromate reductase.* Applied and Environmental Microbiology, 2000. **66**: p. 1788-1795.

139. Ishibashi, Y., C. Cervantes, and S. Silver, *Chromium reduction in Pseudomonas putida.* Applied and Environmental Microbiology, 1990. **56**: p. 2268-2270.

140. Suzuki, T., et al., *NAD(P)H-dependent chromium (VI) reductase of Pseudomonas ambigua G-1: a Cr(V) intermediate is formed during the reduction of Cr(VI) to Cr(III).* Journal of Bacteriology, 1992. **174**(16): p. 5340-5.

141. Riedel, G.F., *The relationship between chromium (VI) uptake, sulfate uptake, and chromium toxicity in the estuarine diatom Thalassiosira pseudonana.* Aquatic Toxicology, 1985. **7**: p. 191-204.

142. Katz, S.A. and H. Salem, *The toxicology of chromium with respect to its chemical speciation: a review.* Journal of Applied Toxicology, 1993. **13**: p. 217-224.

143. Singh, J., et al., *Chromium-induced genotoxicity and apoptosis: relationship to chromium carcinogenesis.* Oncol. Rep., 1998. **5**: p. 1307-1318.

144. Hossain, M.A., M. Alam, and D.R. Yonge, *Estimating the dual-enzyme kinetic parameters for Cr (VI) reduction by Shewanella oneidensis MR-1 from soil column experiments.* Water Research, 2005. **39**(14): p. 3342-8.

145. Hossain, M.A., et al., *Finite element modeling of Cr(VI) reduction by Shewanella oneidensis MR-1 employing the dual-enzyme kinetic model.* Computers & Geosciences, 2005. **31**: p. 1286-1292.

146. Viamajala, S., et al., *Toxic effects of Chromium(VI) on anaerobic and aerobic growth of Shewanella ondeidensis MR-1.* Biotechnol. Prog., 2004. **20**: p. 87-95.

147. Brown, S.D., et al., *Molecular Dynamics of the Shewanella oneidensis Response to Chromate Stress.* Molecular & Cellular Proteomics, 2006. **5**: p. 1054-1071.

148. Summers, A.O. and G.A. Jacoby, *Plasmid-determined resistance to boron and chromium compounds in Pseudomonas aeruginosa.* Antimicrob. Agents Chemother., 1978. **13**: p. 637-640.

149. Ohtake, H., C. Cervantes, and S. Silver, *Decreased chromate uptake in Pseudomonas fluorescens carrying a chromate resistance plasmid.* Journal of Bacteriology, 1987. **169**: p. 3853-3556.

150. Cervantes, C. and H. Ohtake, *Plasmid-determined resistance to chromate in Pseudomonas aeruginosa.* FEMS Microbiology Letters, 1988. **56**: p. 173-176.

151. Cervantes, C., et al., *Cloning, nucleotide sequence, and expression of the chromate resistance determinant of Pseudomonas aeruginosa plasmid pUM505. .* Journal of Bacteriology, 1990. **172**: p. 287-291.

152. Bopp, L.H., A.M. Chakrabarty, and H.L. Ehrlich, *Chromate resistance plasmid in Pseudomonas fluorescens.* Journal of Bacteriology, 1983. **155**: p. 1105-1109.

153. Bopp, L.H. and H.L. Erlich, *Chromate resistance and reduction in Pseudomonas fluorescens strain LB300.* Arch. Microbiol., 1988. **150**: p. 426-431.

154. Nies, A., D.H. Nies, and S. Silver, *Cloning and expression of plasmid genes encoding resistances to chromate and cobalt in Alcaligenes eutrophus.* Journal of Bacteriology, 1989. **171**: p. 5065-5070.

155. Nies, A., D.H. Nies, and S. Silver, *Nucleotide sequence and expression of a plasmid-encoded chromate resistance determinant from Alcaligenes eutrophus.* Journal of Biological Chemistry, 1990. **265**: p. 5648-5653.

260

156. Thompson, M.R., et al., *Dosage-Dependent Proteome Response of Shewanella oneidensis MR-1 to Acute Chromate Challenge.* Journal of Proteome Research, 2007. **6**(5): p. 1745-1757.

157. Chourey, K., et al., *The orphan response regulator SO2426 of Shewanella oneidensis MR-1 is involved in the cellular response to metal stress.* 2007. **In preparation for submission to Journal of Bacteriology.**

158. Chourey, K., et al., *Global Molecular and Morphological Effects of 24-h Chromium Exposure on Shewanella oneidensis MR-1.* Applied and Environmental Microbiology, 2006. **72**: p. 6331-6344.

159. Thompson, M.R., et al., *Proteomic Comparison of Shewanella oneidensis MR-1 Wild-Type and a Response Regulator Deletion Strain under Conditions of Chromate Transformation.* 2007. **In preparation for submission to Applied and Environmental Microbiology.**

160. Thompson, M.R., et al., *Experimental Approach For Large-Scale Proteome Measurements From Small Amounts (low mg) Of Microbial Cultures.* 2007. **In preparation for submission to Analytical Chemistry.**

161. Rice, R.H., G.E. Means, and W.D. Brown, *Stabilization of bovine trypsin by reductive methylation.* Biochimica et Biophysica Acta, 1977. **492**(2): p. 316-21.

162. Smith, P.K., et al., *Measurement of protein using bicinchoninic acid.* Analytical Biochemistry, 1985. **150**(1): p. 76-85.

163. Chao, L. and I.E. Liener, *Autoactivation of Porcine Trypsinogen in the Presence and Absence of Calcium.* Biochim Biophys Acta, 1965. **96**: p. 508-16.

164.    Wollum, A.G., *Cultural methods for soil microorganisms*, in *Methods of Soil Analysis. Part 2-Chemical and Microbiological Properties*, A.L. Page, R.H. Miller, and D.R. Keency, Editors. 1982, Soil Science Society of America: Madison, WI. p. 781-802.

165.    Qian, W.J., et al., *Comparative proteome analyses of human plasma following in vivo lipopolysaccharide administration using multidimensional separations coupled with tandem mass spectrometry.* Proteomics, 2005. **5**(2): p. 572-84.

166.    McDonald, W.H., et al., *Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins fromcomplex mixtures:  single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT.* International Journal of Mass Spectrometry, 2002. **219**: p. 245-251.

167.    de Hoffman, E. and V. Stroobant, *2.  Mass Analyzers*, in *Mass Spectrometry: Principles and Applications*. 2002, John Wiley & Sons, LTD.: New York. p. 63-132.

168.    Paul, W. and H. Steinwedel, *Patent No.  2.939,952*. 1960: United States Patent.

169.    de Hoffman, E. and V. Stroobant, *3.  Tandem Mass Spectrometry (MS/MS)*, in *Mass Spectrometry:  Principles and Applications*. 2002, John Wiley & Sons, LTD: New York. p. 133-181.

170.    Schwartz, J.C. and I. Jardine, *Quadrupole ion trap mass spectrometry.* Methods in Enzymology, 1996. **270**: p. 552-86.

171.    Tabb, D.L., W.H. McDonald, and J.R. Yates III, *DTASelect and Contrast:  Tools for assembling and comparing protein identifications from shotgun proteomics.* Journal of Proteome Research, 2002. **1**: p. 21-26.

172.    Tabb, D.L., et al., *Determination of peptide and protein ion charge states by Fourier transformation of isotope-resolved mass spectra.* Journal of the American Society for Mass Spectrometry, 2006. **17**(7): p. 903-15.

173.    Narasimhan, C., et al., *MASPIC: intensity-based tandem mass spectrometry scoring scheme that improves peptide identification at high confidence.* Anal Chem, 2005. **77**(23): p. 7581-93.

174.    Keller, A., et al., *Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search.* Analytical Chemistry, 2002. **74**(20): p. 5383-92.

175.    Washburn, M.P., et al., *Analysis of quantitative proteomic data generated via multidimensional protein identification technology.* Analytical Chemistry, 2002. **74**: p. 1650-1657.

176.    Wu, C.C., et al., *Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis.* Analytical Chemistry, 2004. **76**(17): p. 4951-9.

177.    Gygi, S.P., et al., *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.* Nature Biotechnology, 1999. **17**(10): p. 994-9.

178.    Wu, W.W., et al., *Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D Gel- or LC-MALDI TOF/TOF.* Journal of Proteome Research, 2006. **5**: p. 651-658.

179.    Shadforth, I.P., et al., *i-Tracker:  For quantitative proteomics using iTRAQ (TM).* BMC Genomics, 2005. **6**: p. 145-150.

180.  Zieske, L.R., *A perspective on the use of iTRAQ (TM) reagent technology for protein complex and profiling studies.* Journal of Experimental Botany, 2006. **57**: p. 1501-1508.

181.  Kirkpatrick, D.S., S.A. Gerber, and S.P. Gygi, *The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications.* Methods, 2005. **35**(3): p. 265-73.

182.  Zybailov, B.L., L. Florens, and M.P. Washburn, *Quantitative shotgun proteomics using a protease with broad specificity and normalized spectral abundance factors.* Molecular Biosystems, 2007. **3**(5): p. 354-60.

183.  Zhang, B., et al., *Detecting differential and correlated protein expression in label-free shotgun proteomics.* Journal of Proteome Research, 2006. **5**: p. 2909-2918.

184.  VerBerkmoes, N.C., et al., *Determination and comparison of the baseline proteomes of the versatile microbe Rhodopseudomonas palustris under its major metabolic states.* Journal of Proteome Research, 2006. **5**: p. 287-298.

185.  Peng, J., et al., *Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome.* Journal of Proteome Research, 2003. **2**(1): p. 43-50.

186.  Giometti, C.S., et al., *Analysis of the Shewanella oneidensis proteome by two-dimensional gel electrophoresis under nondenaturing conditions.* Proteomics, 2003. **3**: p. 777-785.

187.  Kolker, E., et al., *Global profiling of Shewanella oneidensis MR-1: Expression of hypothetical genes and improved functional annotations.* PNAS, 2005. **102**: p. 2099-2104.

188. Wang, C.C., and Newton, A. (1969) *Iron transport in Escherichia coli: relationship between chromium sensitivity and high iron requirement in mutants of Escherichia coli.* Journal of Bacteriology, 1969. **98**: p. 1135-1141.

189. Andrews, S.C., A.K. Robinson, and F. Rodrıguez-Quinones, *Bacterial iron homeostasis.* FEMS Microbiology Reviews, 2003. **27**: p. 215-237.

190. de Lorenzo, V., et al. *Iron transport in bacteria.* in *American Society for Microbiology.* 2004. Washington, D.C.

191. Andrews, S.C., et al., *Overproduction, purification and characterization of the bacterioferritin of Escherichia coli and a C-terminally extended variant.* European Journal of Biochemistry, 1993. **213**: p. 329-338.

192. Qiu, X., et al., *Comparative analysis of differentially expressed genes in Shewanella oneidensis MR-1 following exposure to UVC, UVB, and UVA radiation.* Journal of Bacteriology, 2005. **187**: p. 3556-3564.

193. Palma, M., et al., *Transcriptome analysis of the response of Pseudomonas aeruginosa to hydrogen peroxide.* Journal of Bacteriology, 2004. **186**: p. 248-252.

194. Salunkhe, P., et al., *Genome-wide transcriptional profiling of the steady-state response of Pseudomonas aeruginosa to hydrogen peroxide.* Journal of Bacteriology, 2005. **187**: p. 2565-2572.

195. Zheng, M., et al., *OxyR and SoxRS regulation of fur.* Journal of Bacteriology, 1999. **181**: p. 4639-4643.

196. Benov, L., Kredich, N. M., and Fridovich, I. (1996) *The mechanism of the auxotrophy for sulfur-containing amino acids imposed upon Escherichia coli by superoxide.* Journal of Biological Chemistry, 1996. **271**: p. 21037-21040.

197. Newton, G.L., et al., *Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes.* Journal of Bacteriology, 1996. **178**: p. 1990-1995.

198. Mitaku, S., T. Hirokawa, and T. Tsuji, *Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces.* Bioinformatics, 2002. **18**: p. 608-616.

199. Ferguson, G.P., et al., *Methylglyoxal production in bacteria: suicide or survival?* Arch. Microbiol., 1998. **170**: p. 209-219.

200. McGinnis, S. and T.L. Madden, *BLAST: at the core of a powerful and diverse set of sequence analysis tools.* Nucleic Acids Research, 2004. **32**: p. W20-W25.

201. Ackerley, D.F., et al., *Chromate-reducing properties of soluble flavoproteins from Pseudomonas putida and Escherichia coli.* Applied and Environmental Microbiology, 2004. **70**: p. 873-882.

202. Gonzalez, C.F., et al., *ChrR, a soluble quinone reductase of Pseudomonas putida that defends against $H_2O_2$.* Journal of Biological Chemistry, 2005. **280**: p. 22590-22595.

203. Qiu, X., et al., *Transcriptome analysis applied to survival of Shewanella oneidensis MR-1 exposed to ionizing radiation.* Journal of Bacteriology, 2006. **188**: p. 1199-1204.

204. Aswad, D.W., *Stoichiometric methylation of porcine adrenocorticotropin by protein carboxyl methyltransferase requires deamidation of asparagine 25. Evidence for methylation at the alpha-carboxyl group of atypical L-isoaspartyl residues.* Journal of Biological Chemistry, 1984. **259**: p. 10714-10721.

205. Murray, E.D., Jr. and S. Clarke, *Synthetic peptide substrates for the erythrocyte protein carboxyl methyltransferase. Detection of a new site of methylation at isomerized L-aspartyl residues.* Journal of Biological Chemistry, 1984. **259**: p. 10722-10732.

206. Visick, J.E., H. Cai, and S. Clarke. 1998. , *The L-isoaspartyl protein repair methyltransferase enhances survival of aging Escherichia coli subjected to secondary environmental stresses.* Journal of Bacteriology, 1998. **180**: p. 2623-2629.

207. Elias, D.A., et al., *Confirmation of the expression of a large set of conserved hypothetical proteins in Shewanella oneidensis MR-1.* Journal of Microbiological Methods, 2006. **66**(2): p. 223-33.

208. Jonscher, K.R. and J.R. Yates III, *The quadrupole ion trap mass spectrometer--a small solution to a big challenge.* Analytical Biochemistry, 1997. **244**: p. 1-15.

209. Kolker, E., et al., *Initial proteome analysis of model microorganism Haemophilus influenzae strain Rd KW20.* Journal of Bacteriology, 2003. **185**: p. 4593-4602.

210. Florens, L., et al., *A proteomic view of the Plasmodium falciparum life cycle.* Nature, 2002. **419**: p. 520-526.

211. Romine, M.F., et al., *Validation of Shewanella oneidensis MR-1 small proteins by AMT tag-based proteome analysis.* OMICS, 2004. **8**: p. 239-254.

212. Elias, D.A., et al., *Global detection and characterization of hypothetical proteins in Shewanella oneidensis MR-1 using LC-MS based proteomics.* Proteomics, 2005. **5**: p. 3120-3130.

213.    Gadd, G.M., *Bioremedial potential of microbial mechanisms of metal mobilization and immobilization.* Curr Opin Biotechnol, 2000. **11**: p. 271-9.

214.    Lovley, D.R. and E.J. Phillips, *Reduction of Chromate by Desulfovibrio vulgaris and Its c(3) Cytochrome.* Applied and Environmental Microbiology, 1994. **60**(2): p. 726-728.

215.    Xie, H. and T.J. Griffin, *Trade-off between high sensitivity and increased potential for false positive peptide sequence matches using a two-dimensional linear ion trap for tandem mass spectrometry-based proteomics.* Journal of Proteome Research, 2006. **5**(4): p. 1003-9.

216.    Luo, Q., et al., *Preparation of 20-μm-i.d. silica-based monolithic columns and their performance for proteomics analyses.* Analytical Chemistry, 2005. **77**: p. 5028-5035.

217.    Alvarez, A.H., R. Moreno-Sánchez, and C. Cervantes, *Chromate efflux by means of the ChrA chromate resistance protein from Pseudomonas aeruginosa.* Journal of Bacteriology, 1999. **181**: p. 7398-7400.

218.    Muskotal, A., et al., *Interaction of FliS flagellar chaperone with flagellin.* FEBS Letters, 2006. **580**(16): p. 3916-20.

219.    Thomas, J.G. and F. Baneyx, *Roles of the Escherichia coli small heat shock proteins IbpA and IbpB in thermal stress management:  comparison with ClpA, ClpB, and HtpG in vivo.* Journal of Bacteriology, 1998. **180**: p. 5165-5172.

220.    Tsung, K., S. Inouye, and M. Inouye, *Factors affecting the efficiency of protein synthesis in Escherichia coli. Production of a polypeptide of more than 6000 amino acid residues.* J Biol Chem, 1989. **264**(8): p. 4428-33.

221.    Jiang, X., et al., *Escherichia coli prlC gene encodes a trypsin-like proteinase regulating the cell cycle.* J Biochem (Tokyo), 1998. **124**(5): p. 980-5.

222.    Weber, A., S.A. Kogl, and K. Jung, *Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in Escherichia coli.* Journal of Bacteriology, 2006. **188**(20): p. 7165-75.

223.    Han, K.K. and A. Martinage, *Post-translational chemical modification(s) of proteins.* The International Journal of Biochemistry, 1992. **24**(1): p. 19-28.

224.    Sourjik, V., *Receptor clustering and signal processing in E. coli chemotaxis.* Trends in Microbiology, 2004. **12**(12): p. 569-76.

225.    Humbard, M.A., S.M. Stevens, Jr., and J.A. Maupin-Furlow, *Posttranslational modification of the 20S proteasomal proteins of the archaeon Haloferax volcanii.* Journal of Bacteriology, 2006. **188**(21): p. 7521-30.

226.    Guan, J. and M.R. Chance, *Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry.* Trends in Biochemical Sciences, 2005. **30**: p. 583-592.

227.    Jaakkola, P., et al., *Targeting of HIF-$\alpha$ to the von Hippel-Lindau Ubiquitylation Complex by $O_2$-Regulated Prolyl Hydroxylation.* Science, 2001. **292**(5516): p. 468-472.

228.    Mulkidjanian, A.Y., *Ubiquinol oxidation in the cytochrome bc1 complex: reaction mechanism and prevention of short-circuiting.* Biochim Biophys Acta, 2005. **1709**(1): p. 5-34.

229.  Roth, M.J., et al., *Precise and parallel characterization of coding polymorphisms, alternative splicing and modifications in human proteins by mass spectrometry.* Molecular & Cellular Proteomics, 2005. **4**(1002-1008).

230.  Garcia, B.A., J. Shabanowitz, and D.F. Hunt, *Characterization of histones and their post-translational modifications by mass spectrometry.* Current Opinion in Chemical Biology, 2007. **11**(1): p. 66-73.

231.  Nielsen, M.L., M.M. Savitski, and R.A. Zubarev, *Extent of modifications in human proteome samples and their effect on dynamic range of analysis in shotgun proteomics.* Molecular & Cellular Proteomics, 2006. **5**(12): p. 2384-91.

232.  Lewis, R.J., et al., *Phosphorylated aspartate in the structure of a response regulator protein.* Journal of Molecular Biology, 1999. **294**: p. 9-15.

233.  Kleinnijenhuis, A.J., et al., *Analysis of Histidine Phosphorylation Using Tandem MS and Ion-Electron Reactions.* Analytical Chemistry, 2007.

234.  Vasilescu, J., et al., *Proteomic analysis of ubiquitinated proteins from human MCF-7 breast cancer cells by immunoaffinity purification and mass spectrometry.* Journal of Proteome Research, 2005. **4**(6): p. 2192-200.

235.  Havilio, M. and A. Wool, *Large-scale unrestricted identification of post-translational modifications using tandem mass spectrometry.* Analytical Chemistry, 2007: p. (7).

236.  Annan, R.S. and S.A. Carr, *The essential role of mass spectrometry in characterizing protein structure: mapping posttranslational modifications.* Journal of Protein Chemistry, 1997. **16**(5): p. 391-402.

237. Florens, L., et al., *Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors.* Methods, 2006. **40**: p. 303-311.

238. Allen, J.F., *Redox control of transcription:  sensors, response regulators, activators and repressors.* FEBS Letters, 1993. **332**: p. 203-207.

239. Kato, A. and E.A. Groisman, *Connecting two-component regulatory systems by a protein that protects a response regulator from dephosphorylation by its cognate sensor.* Genes Dev, 2004. **18**(18): p. 2302-13.

240. Stock, J.B., M.N. Levit, and P.M. Wolanin, *Information processing in bacterial chemotaxis.* Science's STKE, 2002.

241. Djordjevic, S. and A.M. Stock, *Structural analysis of bacterial chemotaxis proteins: components of a dynamic signaling system.* J Struct Biol, 1998. **124**(2-3): p. 189-200.

242. Groh, J.L., et al., *A method adapting microarray technology for signature-tagged mutagenesis of Desulfovibrio desulfuricans G20 and Shewanella oneidensis MR-1 in anaerobic sediment survival experiment.* Applied and Environmental Microbiology, 2005. **71**: p. 7064-7074.

243. Liu, Y., et al., *Transcriptome analysis of Shewanella oneidensis MR-1 in response to elevated salt conditions.* Journal of Bacteriology, 2005. **187**: p. 2501-2507.

244. Shioi, J., C.V. Dang, and B.L. Taylor, *Oxygen as attractant and repellent in bacterial chemotaxis.* Journal of Bacteriology, 1987. **169**(7): p. 3118-23.

245. Jeanmougin, F., et al., *Multiple sequence alignment with Clustal X.* Trends in Biochemical Sciences, 1998. **23**: p. 403-405.

246.     Kleene, S.J., M.L. Toews, and J. Adler, *Isolation of glutamic acid methyl ester from an Escherichia coli membrane protein involved in chemotaxis.* The Journal of Biological Chemistry, 1977. **252**(10): p. 3214-8.

247.     Gralnick, J.A., C.T. Brown, and D.K. Newman, *Anaerobic regulation by an atypical Arc system in Shewanella oneidensis.* Molecular Microbiology, 2005. **56**(5): p. 1347-57.

248.     Brown, S.D., et al., *Cellular response of Shewanella oneidenis to strontium stress.* Applied and Environmental Microbiology, 2006. **72**(1): p. 890-900.

249.     Mizoguchi, H., H. Mori, and T. Fujio, *Escherichia coli minimum genome factory.* Biotechnology and Applied Biochemistry, 2007. **46**: p. 157-167.

250.     Lipton, M.S., et al., *Global analysis of the Deinococcus radiodurans proteome by using accurate mass tags.* PNAS, 2002. **99**: p. 11049-11054.

251.     Karp, N.A., et al., *Impact of replicate types on proteomic expression analysis.* Journal of Proteome Research, 2005. **4**(5): p. 1867-71.

252.     Schmid, A.K., et al., *Global whole-cell FTICR mass spectrometric proteomics analysis of the heat shock response in the radioresistant bacterium Deinococcus radiodurans.* Journal of Proteome Research, 2005. **4**: p. 709-718.

253.     Zybailov, B., et al., *Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae.* Journal of Proteome Research, 2006. **5**(9): p. 2339-47.

254.     Finn, R.D., et al., *Pfam: clans, web tools and services.* Nucleic Acids Res, 2006. **34**(Database issue): p. D247-51.

255. Bigas, A., et al., *Non-viability of Haemophilus parasuis fur-defective mutants.* Vet Microbiol, 2006. **118**(1-2): p. 107-16.

256. Idone, V., et al., *Effect of an orphan response regulator on Streptococcus mutans sucrose-dependent adherence and cariogenesis.* Infect Immun, 2003. **71**(8): p. 4351-60.

257. Jormakka, M., B. Byrne, and S. Iwata, *Formate dehydrogenase--a versatile enzyme in changing environments.* Curr Opin Struct Biol, 2003. **13**(4): p. 418-23.

258. Myers, C.R. and J.M. Myers, *MtrB is required for proper incorporation of the cytochromes OmcA and OmcB into the outer membrane of Shewanella putrefaciens MR-1.* Applied and Environmental Microbiology, 2002. **68**(11): p. 5585-94.

259. Palacios, S. and J.C. Escalante-Semerena, *prpR, ntrA, and ihf functions are required for expression of the prpBCDE operon, encoding enzymes that catabolize propionate in Salmonella enterica serovar typhimurium LT2.* J Bacteriol, 2000. **182**(4): p. 905-10.

260. Rigali, S., et al., *Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies.* Journal of Biological Chemistry, 2002. **277**(15): p. 12507-15.

261. Gorelik, M., et al., *Structural characterization of GntR/HutC family signaling domain.* Protein Science, 2006. **15**(6): p. 1506-11.

262. Varghese, S., Y. Tang, and J.A. Imlay, *Contrasting sensitivities of Escherichia coli aconitases A and B to oxidation and iron depletion.* J Bacteriol, 2003. **185**(1): p. 221-30.

263. Albanesi, D., et al., *Bacillus subtilis cysteine synthetase is a global regulator of the expression of genes involved in sulfur assimilation.* J Bacteriol, 2005. **187**(22): p. 7631-8.

264. Hess, J.F., R.B. Bourret, and M.I. Simon, *Phosphorylation assays for proteins of the two-component regulatory system controlling chemotaxis in Escherichia coli.* Methods in Enzymology, 1991. **200**: p. 188-204.

265. Viamajala, S., et al., *Chromate reduction in Shewanella oneidensis MR-1 is an inducible process associated with anaerobic growth.* Biotechnology Progress, 2002. **18**(2): p. 290-5.

266. Wang, H., et al., *Development and evaluation of a micro- and nanoscale proteomic sample preparation method.* Journal of Proteome Research, 2005. **4**: p. 2397-2403.

267. Corbin, R.W., et al., *Toward a protein profile of Escherichia coli: comparison to its transcription profile.* PNAS, 2003. **100**: p. 9232-9237.

268. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.* Molecular & Cellular Proteomics, 2004. **3**(12): p. 1154-69.

269. Zhu, W., et al., *Shotgun proteomics of Methanococcus jannaschii and insights into methanogenesis.* Journal of Proteome Research, 2004. **3**: p. 538-548.

270. Callister, S.J., et al., *Application of the Accurate Mass and Time Tag Approach to the Proteome Analysis of Sub-cellular Fractions Obtained from Rhodobacter sphaeroides 2.4.1. Aerobic and Photosynthetic Cell Cultures.* Journal of Proteome Research, 2006. **5**(8): p. 1940-1947.

271.    Coleman, S.E., I. van de Rijn, and A.S. Bleiweis, *Lysis of Grouped and Ungrouped Streptococci by Lysozyme.* Infect Immun, 1970. **2**(5): p. 563-569.

272.    Van Huynh, N., et al., *A procedure for the preparation of bacterial DNA that employs dimethyl sulfoxide to induce the lysis of cells.* Analytical Biochemistry, 1989. **176**: p. 464-467.

273.    Yu, Y.Q., et al., *Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins.* Analytical Chemistry, 2003. **75**(21): p. 6023-8.

274.    Bosserhoff, A., J. Wallach, and R.W. Frank, *Micropreparative separation of peptides derived from sodium dodecyl sulphate-solubilized proteins.* Journal of Chromatography, 1989. **473**(1): p. 71-7.

275.    De Leij, L. and B. Witholt, *Structural heterogeneity of the cytoplasmic and outer membranes of Escherichia coli.* Biochimica et Biophysica Acta, 1977. **471**(1): p. 92-104.

276.    Guerlava, P., V. Izac, and J.L. Tholozan, *Comparison of different methods of cell lysis and protein measurements in Clostridium perfringens: application to the cell volume determination.* Curr Microbiol, 1998. **36**(3): p. 131-5.

277.    Wilkinson, J.M., *Fragmentation of Polypeptides by Enzymic Methods*. Practical Protein Chemistry: A Handbook. , ed. A. Darbre. 1986, New York: John Wiley and Sons.

278.    Woese, C.R., *Whither microbiology? Phylogenetic trees.* Curr Biol, 1996. **6**(9): p. 1060-3.

279. Schloss, P.D. and J. Handelsman, *Biotechnological prospects from metagenomics.* Curr Opin Biotechnol, 2003. **14**(3): p. 303-10.

280. Handelsman, J., et al., *Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products.* Chemistry & Biology, 1998. **5**(10): p. R245-9.

281. Torsvik, V. and L. Ovreas, *Microbial diversity and function in soil: from genes to ecosystems.* Current Opinion in Microbiology, 2002. **5**(3): p. 240-5.

282. Biron, D.G., et al., *Population proteomics: an emerging discipline to study metapopulation ecology.* Proteomics, 2006. **6**(6): p. 1712-5.

283. Paul, D., et al., *Accessing microbial diversity for bioremediation and environmental restoration.* Trends Biotechnol, 2005. **23**(3): p. 135-42.

284. Wilmes, P. and P.L. Bond, *The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms.* Environmental Microbiology, 2004. **6**(9): p. 911-20.

285. Wilmes, P. and P.L. Bond, *Towards exposure of elusive metabolic mixed-culture processes: the application of metaproteomic analyses to activated sludge.* Water Science and Technology, 2006. **54**(1): p. 217-26.

286. Powell, M.J., et al., *Marine proteomics: generation of sequence tags for dissolved proteins in seawater using tandem mass spectrometry.* Marine Chemistry, 2005. **95**: p. 183-198.

287. Savidor, A., et al., *Expressed peptide tags: an additional layer of data for genome annotation.* Journal of Proteome Research, 2006. **5**(11): p. 3048-58.

288. Varga, A.R. and L.A. Staehelin, *Spatial differentiation in photosynthetic and non-photosynthetic membranes of Rhodopseudomonas palustris.* Journal of Bacteriology, 1983. **154**(3): p. 1414-30.

289. Deshusses, J.M., et al., *Exploitation of specific properties of trifluoroethanol for extraction and separation of membrane proteins.* Proteomics, 2003. **3**(8): p. 1418-24.

290. Lo, I., et al., *Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria.* Nature, 2007. **446**(7135): p. 537-41.

291. Gan, C.S., et al., *Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ).* Journal of Proteome Research, 2007. **6**(2): p. 821-7.

292. Larsen, M.R., et al., *Analysis of posttranslational modifications of proteins by tandem mass spectrometry.* Biotechniques, 2006. **40**(6): p. 790-8.

293. Viamajala, S., B.M. Peyton, and J.N. Petersen, *Modeling chromate reduction in Shewanella oneidensis MR-1:  Development of a novel dual-enzyme kinetic model.* Biotechnology and Bioengineering, 2003. **83**: p. 790-797.

294. Tang, Y.J., et al., *Evaluation of the effects of various culture conditions on Cr(VI) reduction by Shewanella oneidensis MR-1 in a novel high-throughput mini-bioreactor.* Biotechnology and Bioengineering, 2006. **95**(1): p. 176-84.

295. Hermjakob, H., et al., *The HUPO PSI's molecular interaction format--a community standard for the representation of protein interaction data.* Nature Biotechnology, 2004. **22**(2): p. 177-83.

296.    Cuticchia, A.J., *Future vision of the GDB human genome database.* Hum Mutat, 2000. **15**(1): p. 62-7.

297.    Karp, P.D., S. Paley, and J. Zhu, *Database verification studies of SWISS-PROT and GenBank.* Bioinformatics, 2001. **17**(6): p. 526-32; discussion 533-4.

298.    Ashurst, J.L. and J.E. Collins, *Gene annotation: prediction and testing.* Annu Rev Genomics Hum Genet, 2003. **4**: p. 69-88.

299.    Skupski, M.P., et al., *The Genome Sequence DataBase: towards an integrated functional genomics resource.* Nucleic Acids Research, 1999. **27**(1): p. 35-8.

300.    Domon, B. and R. Aebersold, *Challenges and Opportunities in Proteomics Data Analysis.* Molecular & Cellular Proteomics, 2006. **5**(10): p. 1921-1926.

301.    Norbeck, A.D., et al., *Proteomic approaches to bacterial differentiation.* Journal of Microbiological Methods, 2006. **67**(3): p. 473-86.

302.    Lacerda, C.M., L.H. Choe, and K.F. Reardon, *Metaproteomic analysis of a bacterial community response to cadmium exposure.* Journal of Proteome Research, 2007. **6**(3): p. 1145-52.

# Appendix

**Appendix A1: All Peptides Identified in Cr-Shocked *S. oneidensis* MR-1 using an LTQ Mass Spectrometer**

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO0002 | | | | | 2 | | | | proton/peptide symporter family protein |
| SO0003 | 2 | 4 | | | | 3 | 2 | 3 | tRNA modification GTPase TrmE (trmE) |
| SO0004 | 26 | 18 | 19 | 18 | 24 | 15 | 20 | 16 | inner membrane protein, 60 kDa |
| SO0005 | | 2 | | | | 3 | | 2 | conserved hypothetical protein TIGR00278 |
| SO0006 | 2 | 3 | 3 | 3 | 2 | 3 | 2 | 2 | ribonuclease P protein component (rnpA) |
| SO0007 | | 2 | | | | | | | ribosomal protein L34 (rpmH) |
| SO0008 | 10 | 12 | 9 | 13 | 13 | 14 | 13 | 14 | chromosomal replication initiator protein DnaA (dnaA) |
| SO0009 | 6 | 8 | 10 | 12 | 8 | 7 | 9 | 11 | DNA polymerase III, beta subunit (dnaN) |
| SO0010 | | 2 | | 2 | | 2 | | | DNA replication and repair protein RecF (recF) |
| SO0011 | 44 | 49 | 48 | 55 | 56 | 44 | 56 | 66 | DNA gyrase, B subunit (gyrB) |
| SO0012 | | | | 2 | | | 2 | | glutathione S-transferase family protein |
| SO0014 | 32 | 42 | 37 | 40 | 34 | 35 | 34 | 44 | glycyl-tRNA synthetase, beta subunit (glyS) |
| SO0015 | 3 | | | | 3 | | 2 | 4 | glycyl-tRNA synthetase, alpha subunit (glyQ) |
| SO0016 | | | | | | | | 2 | DNA-3-methyladenine glycosidase I (tag) |
| SO0017 | 7 | 8 | 2 | 2 | 4 | 8 | | 4 | conserved hypothetical protein |
| SO0018 | 2 | 3 | 2 | 2 | | | 3 | 2 | conserved hypothetical protein |
| SO0019 | | 5 | | 3 | | | 2 | 3 | conserved hypothetical protein |
| SO0020 | 23 | 21 | 19 | 20 | 21 | 24 | 13 | 22 | fatty oxidation complex, beta subunit (fadA) |
| SO0021 | 43 | 53 | 45 | 48 | 43 | 40 | 47 | 52 | fatty oxidation complex, alpha subunit (fadB) |
| SO0022 | | 3 | 2 | 2 | | | 4 | 4 | prolidase (pepQ) |
| SO0025 | 2 | 3 | | | 2 | 2 | | 3 | conserved hypothetical protein |
| SO0026 | | | | 2 | | | 3 | | transcriptional regulator, ArsR family |
| SO0027 | 3 | | | | 3 | 5 | | | protoporphyrinogen oxidase, putative |
| SO0029 | 9 | 16 | 5 | 9 | 11 | 17 | 6 | 9 | potassium uptake protein TrkA (trkA) |
| SO0030 | 10 | 8 | 4 | 7 | 4 | 2 | 4 | 8 | sun protein (sun) |
| SO0031 | 7 | 10 | 5 | 7 | 7 | | 4 | 7 | methionyl-tRNA formyltransferase (fmt) |
| SO0032 | 3 | 6 | 6 | 7 | 8 | 2 | 5 | 6 | polypeptide deformylase (def-1) |
| SO0033 | 3 | 4 | 5 | 3 | 2 | | 5 | 3 | LysM domain protein |
| SO0037 | | 2 | 2 | | | | | 2 | Sua5/YciO/YrdC/YwlC family protein |
| SO0038 | 2 | 3 | 2 | 2 | 2 | | | 3 | coproporphyrinogen III oxidase, aerobic (hemF) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO0039 | 3 | 3 | 3 | 2 | 4 | | 3 | 4 | conserved hypothetical protein |
| SO0042 | 3 | 3 | | | | 2 | | | carbonic anhydrase, family 3 |
| SO0045 | 2 | | | 4 | 2 | | | | Rrf2 family protein |
| SO0048 | 15 | 15 | 8 | 14 | 17 | 7 | 11 | 12 | peptidase, M23/M37 family |
| SO0049 | 16 | 9 | 12 | 12 | 9 | 9 | 14 | 14 | phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent (gpmA) |
| SO0050 | 10 | 11 | 9 | 9 | 10 | 9 | 9 | 10 | rhodanese domain protein |
| SO0052 | 9 | 16 | 20 | 16 | 17 | 14 | 18 | 17 | protein-export protein SecB (secB) |
| SO0053 | 5 | 7 | 8 | 9 | 4 | 3 | 8 | 7 | glycerol-3-phosphate dehydrogenase (NAD(P)+) (gpsA) |
| SO0054 | 2 | | | | | | | | conserved hypothetical protein TIGR00275 |
| SO0055 | | | | 2 | | | 3 | 3 | conserved domain protein |
| SO0060 | | 2 | | | | | | | sensor histidine kinase |
| SO0061 | 3 | | | 2 | | 2 | | | lipoprotein, NLP/P60 family |
| SO0062 | 6 | 5 | 3 | 4 | 4 | 3 | 4 | 3 | hypothetical protein |
| SO0063 | | | 3 | 2 | 3 | | 2 | 2 | conserved hypothetical protein |
| SO0065 | 9 | 6 | 7 | 3 | 3 | 4 | 2 | | molybdenum cofactor biosynthesis protein Mog (mog) |
| SO0066 | 23 | 28 | 22 | 27 | 28 | 27 | 24 | 31 | conserved hypothetical protein |
| SO0067 | | | | | | | | 2 | penicillin-binding protein 1C, putative |
| SO0069 | 3 | 3 | 4 | 4 | 5 | 2 | 5 | 6 | transport protein, putative |
| SO0070 | 5 | 8 | 6 | 9 | 6 | 8 | 9 | 5 | ATP-binding transport protein NatA (natA) |
| SO0071 | 9 | 9 | 10 | 12 | 10 | 6 | 11 | 12 | hydrolase, alpha/beta hydrolase fold family |
| SO0072 | | 4 | | | | 3 | | | transcriptional regulator, GntR family |
| SO0073 | 7 | 9 | 8 | 9 | 5 | 5 | 5 | 8 | ABC transporter, ATP-binding protein |
| SO0075 | 12 | 19 | 11 | 11 | 14 | 18 | 9 | 9 | AMP-binding family protein |
| SO0076 | 38 | 51 | 44 | 50 | 42 | 26 | 27 | 38 | hypothetical protein |
| SO0080 | | | | 2 | | | | | conserved hypothetical protein |
| SO0084 | 3 | 8 | 10 | 10 | 7 | 7 | 13 | 10 | HAD-superfamily hydrolase, subfamily IA, variant 1 family protein |
| SO0095 | 12 | 11 | 7 | 7 | 6 | 8 | 5 | 4 | imidazolonepropionase (hutI) |
| SO0096 | 5 | 2 | 3 | | 2 | | | | histidine utilization repressor (hutC) |
| SO0097 | 23 | 25 | 25 | 25 | 24 | 21 | 20 | 20 | urocanate hydratase (hutU) |
| SO0098 | 15 | 22 | 13 | 9 | 20 | 10 | 13 | 12 | histidine ammonia-lyase (hutH) |
| SO0102 | 7 | 12 | 6 | 10 | 6 | 7 | 5 | 5 | formate dehydrogenase, nitrate-inducible, iron-sulfur subunit (fdnH) |
| SO0104 | 3 | 4 | 2 | 5 | 4 | | 3 | 5 | fdhE protein (fdhE) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
|---|---|---|---|---|---|---|---|---|---|
| SO0105 | 7 | 6 | 2 | 4 | 10 | 4 | 3 | 4 | L-seryl-tRNA selenium transferase (selA) |
| SO0106 | 6 | 5 | 4 | 5 | 9 | 5 | 6 | 7 | selenocysteine-specific translation elongation factor (selB) |
| SO0108 | | | 2 | 2 | | | | | conserved hypothetical protein |
| SO0109 | | | | 2 | 2 | 2 | 2 | 2 | conserved hypothetical protein |
| SO0110 | 6 | 6 | 2 | 4 | 8 | 6 | 7 | 4 | conserved hypothetical protein |
| SO0112 | 4 | 3 | 3 | 4 | | 6 | 5 | 5 | conserved hypothetical protein |
| SO0113 | | | | | | | | 2 | conserved hypothetical protein |
| SO0114 | | 2 | 2 | 2 | | | 2 | 2 | conserved hypothetical protein |
| SO0118 | 3 | | 2 | | 3 | 3 | 2 | 2 | hypothetical protein |
| SO0119 | | | 2 | | | | | | conserved hypothetical protein |
| SO0120 | 2 | 2 | 3 | | | 2 | 3 | 4 | hypothetical protein |
| SO0121 | 12 | 10 | 9 | 13 | 10 | 9 | 11 | 12 | conserved hypothetical protein |
| SO0123 | | | 3 | 2 | 2 | | 2 | | acyltransferase family protein |
| SO0130 | 13 | 11 | 3 | 7 | 4 | 5 | | 5 | protease, putative |
| SO0131 | | 2 | | | | | | | conserved hypothetical protein |
| SO0132 | 2 | | | 2 | | 2 | | | flagellar protein FliL, putative |
| SO0135 | 7 | 7 | 6 | 6 | 8 | 5 | 7 | 6 | hemolysin protein, putative |
| SO0137 | | 2 | | 3 | 2 | | | | molybdopterin biosynthesis MoeB protein (moeB) |
| SO0138 | 4 | 3 | 5 | 3 | 5 | | | | molybdopterin biosynthesis MoeA protein (moeA) |
| SO0139 | 5 | 9 | 10 | 10 | 5 | 5 | 9 | 10 | ferritin (ftn) |
| SO0141 | 5 | 4 | 3 | 4 | 6 | 6 | 3 | 2 | sensory box protein |
| SO0142 | 2 | 3 | 2 | | | | | 3 | 3,4-dihydroxy-2-butanone 4-phosphate synthase (ribB) |
| SO0144 | 12 | 19 | 13 | 17 | 15 | 13 | 9 | 12 | protease II (ptrB) |
| SO0148 | | 2 | 2 | | | | | 3 | hypothetical protein |
| SO0152 | 17 | 15 | 13 | 17 | 12 | 10 | 8 | 13 | conserved hypothetical protein |
| SO0162 | 22 | 29 | 24 | 25 | 28 | 15 | 24 | 30 | phosphoenolpyruvate carboxykinase (ATP) (pckA) |
| SO0163 | 3 | | | | | 2 | | 2 | chaperonin HslO (hslO) |
| SO0164 | | | | | | | 2 | | heat shock protein 15 (hslR) |
| SO0165 | 10 | 17 | 10 | 10 | 10 | 15 | 10 | 15 | general secretion pathway protein C (gspC) |
| SO0166 | 25 | 31 | 33 | 38 | 24 | 31 | 35 | 43 | general secretion pathway protein D (gspD) |
| SO0167 | 17 | 21 | 14 | 17 | 21 | 18 | 14 | 15 | general secretion pathway protein E (gspE) |
| SO0168 | 6 | 11 | 6 | 10 | 9 | 7 | 6 | 9 | general secretion pathway protein F (gspF) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO0169 | 6 | 10 | 8 | 9 | 6 | 9 | 7 | 7 | general secretion pathway protein G (gspG) |
| SO0171 | 2 | 2 | | | | | | | general secretion pathway protein I (gspI) |
| SO0172 | | 2 | 3 | 2 | 3 | 6 | 3 | 3 | general secretion pathway protein J (gspJ) |
| SO0173 | 2 | | 2 | | 2 | | 2 | | general secretion pathway protein K (gspK) |
| SO0174 | 12 | 14 | 13 | 8 | 8 | 7 | 9 | 12 | general secretion pathway protein L (gspL) |
| SO0175 | 4 | 5 | 2 | 3 | 6 | 3 | 5 | 5 | general secretion pathway protein M (gspM) |
| SO0176 | 7 | 12 | 10 | 8 | 7 | 6 | 10 | 9 | general secretion pathway protein N (gspN) |
| SO0190 | | 2 | | 2 | | | 2 | 2 | MutT/nudix family protein |
| SO0191 | | 2 | | 2 | | | 5 | 6 | cysQ protein (cysQ-1) |
| SO0194 | | 2 | 2 | | 2 | 3 | | | acyltransferase family protein |
| SO0196 | 6 | 10 | 6 | 9 | 9 | 4 | 4 | 9 | selenide, water dikinase (selD) |
| SO0197 | 5 | 10 | 4 | 6 | 11 | 13 | 6 | 4 | fatty acid desaturase, family 1 |
| SO0198 | 5 | 6 | 6 | 5 | 5 | 5 | 5 | 2 | transcriptional regulator, TetR family |
| SO0203 | | | 2 | | | | | | hypothetical protein |
| SO0206 | 4 | 5 | 6 | 4 | 2 | | 2 | 6 | tRNA (uracil-5-)-methyltransferase (trmA) |
| SO0208 | 5 | 9 | 8 | 5 | 6 | 6 | 7 | 9 | RNA-binding protein |
| SO0215 | | | 2 | | 2 | | 2 | | pantothenate kinase (panK) |
| SO0217 | 63 | 72 | 65 | 68 | 62 | 65 | 61 | 66 | translation elongation factor Tu (tufB) |
| SO0218 | | | | 2 | | | | | preprotein translocase, SecE subunit (secE) |
| SO0219 | 22 | 20 | 14 | 18 | 16 | 18 | 10 | 18 | transcription antitermination protein NusG (nusG) |
| SO0220 | 15 | 16 | 18 | 17 | 16 | 22 | 14 | 16 | ribosomal protein L11 (rplK) |
| SO0221 | 47 | 53 | 42 | 47 | 44 | 57 | 43 | 46 | ribosomal protein L1 (rplA) |
| SO0222 | 38 | 41 | 34 | 40 | 39 | 48 | 35 | 34 | ribosomal protein L10 (rplJ) |
| SO0223 | 17 | 18 | 21 | 18 | 19 | 18 | 21 | 24 | ribosomal protein L7/L12 (rplL) |
| SO0224 | 166 | 194 | 153 | 166 | 170 | 172 | 155 | 175 | DNA-directed RNA polymerase, beta subunit (rpoB) |
| SO0225 | 121 | 153 | 128 | 134 | 143 | 138 | 142 | 144 | DNA-directed RNA polymerase, beta subunit (rpoC) |
| SO0226 | 13 | 18 | 15 | 17 | 14 | 12 | 17 | 17 | ribosomal protein S12 (rpsL) |
| SO0227 | 32 | 37 | 33 | 38 | 38 | 42 | 30 | 33 | ribosomal protein S7 (rpsG) |
| SO0228 | 44 | 62 | 50 | 46 | 42 | 34 | 29 | 42 | translation elongation factor G (fusA-1) |
| SO0229 | 60 | 70 | 63 | 66 | 60 | 63 | 59 | 64 | translation elongation factor Tu (tufA) |
| SO0230 | 16 | 17 | 17 | 18 | 17 | 18 | 15 | 14 | ribosomal protein S10 (rpsJ) |
| SO0231 | 28 | 34 | 29 | 31 | 30 | 34 | 18 | 29 | ribosomal protein L3 (rplC) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO0232 | 25 | 24 | 24 | 28 | 24 | 27 | 22 | 24 | ribosomal protein L4 (rplD) |
| SO0233 | 14 | 12 | 13 | 13 | 12 | 22 | 11 | 10 | ribosomal protein L23 (rplW) |
| SO0234 | 35 | 39 | 41 | 40 | 43 | 38 | 39 | 42 | ribosomal protein L2 (rplB) |
| SO0235 | 11 | 14 | 11 | 13 | 10 | 11 | 12 | 12 | ribosomal protein S19 (rpsS) |
| SO0236 | 23 | 25 | 22 | 24 | 21 | 23 | 22 | 24 | ribosomal protein L22 (rplV) |
| SO0237 | 42 | 47 | 39 | 42 | 38 | 50 | 40 | 42 | ribosomal protein S3 (rpsC) |
| SO0238 | 19 | 21 | 21 | 20 | 22 | 23 | 20 | 20 | ribosomal protein L16 (rplP) |
| SO0239 | 11 | 11 | 12 | 15 | 11 | 15 | 11 | 10 | ribosomal protein L29 (rpmC) |
| SO0240 | 16 | 16 | 15 | 15 | 15 | 16 | 14 | 18 | ribosomal protein S17 (rpsQ) |
| SO0241 | 19 | 21 | 20 | 23 | 17 | 18 | 19 | 19 | ribosomal protein L14 (rplN) |
| SO0242 | 21 | 25 | 25 | 25 | 21 | 29 | 26 | 24 | ribosomal protein L24 (rplX) |
| SO0243 | 37 | 40 | 38 | 40 | 33 | 42 | 37 | 40 | ribosomal protein L5 (rplE) |
| SO0244 | 13 | 12 | 10 | 15 | 13 | 10 | 13 | 10 | ribosomal protein S14 (rpsN) |
| SO0245 | 19 | 16 | 15 | 18 | 17 | 26 | 11 | 15 | ribosomal protein S8 (rpsH) |
| SO0246 | 28 | 33 | 29 | 30 | 28 | 34 | 26 | 28 | ribosomal protein L6 (rplF) |
| SO0247 | 26 | 27 | 23 | 23 | 20 | 31 | 20 | 22 | ribosomal protein L18 (rplR) |
| SO0248 | 28 | 29 | 29 | 24 | 21 | 30 | 22 | 25 | ribosomal protein S5 (rpsE) |
| SO0249 | 7 | 11 | 9 | 9 | 7 | 9 | 5 | 7 | ribosomal protein L30 (rpmD) |
| SO0250 | 24 | 26 | 23 | 26 | 23 | 26 | 21 | 21 | ribosomal protein L15 (rplO) |
| SO0251 | 7 | 11 | 6 | 7 | 5 | 12 | 4 | 5 | preprotein translocase, SecY subunit (secY) |
| SO0253 | 26 | 22 | 24 | 25 | 22 | 26 | 24 | 21 | ribosomal protein S13 (rpsM) |
| SO0254 | 22 | 25 | 23 | 22 | 22 | 22 | 21 | 20 | ribosomal protein S11 (rpsK) |
| SO0255 | 31 | 40 | 32 | 35 | 34 | 54 | 28 | 39 | ribosomal protein S4 (rpsD) |
| SO0256 | 42 | 42 | 42 | 46 | 39 | 36 | 41 | 43 | DNA-directed RNA polymerase, alpha subunit (rpoA) |
| SO0257 | 17 | 22 | 19 | 21 | 17 | 24 | 19 | 18 | ribosomal protein L17 (rplQ) |
| SO0259 | 2 | | 2 | 2 | | 4 | 3 | 4 | cytochrome c biogenesis protein CcmE (ccmE) |
| SO0261 | | | | | 3 | | | | heme exporter protein CcmC (ccmC) |
| SO0263 | 6 | 7 | 9 | 10 | 9 | 7 | 8 | 8 | heme exporter protein CcmA (ccmA) |
| SO0264 | 7 | 6 | 4 | 3 | 4 | 7 | 3 | 3 | cytochrome c (scyA) |
| SO0265 | 30 | 27 | 22 | 28 | 36 | 28 | 24 | 31 | conserved hypothetical protein |
| SO0266 | 8 | 11 | 11 | 7 | 11 | 11 | 9 | 13 | cytochrome c-type biogenesis protein CcmF (ccmF-1) |
| SO0267 | 13 | 21 | 18 | 9 | 17 | 18 | 16 | 15 | thiol:disulfide interchange protein DsbE (dsbE) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO0268 | 9 | 9 | 6 | 4 | 6 | 10 | 6 | 9 | cytochrome c-type biogenesis protein CcmH (ccmH) |
| SO0272 | 2 | 6 | 8 | 9 | 9 | 4 | 10 | 14 | competence/damage-inducible protein CinA (cinA) |
| SO0273 | 4 | 8 | 9 | 4 | 8 | 4 | 6 | 7 | conserved hypothetical protein |
| SO0274 | 17 | 14 | 13 | 14 | 19 | 15 | 16 | 19 | phosphoenolpyruvate carboxylase (ppc) |
| SO0278 | | 5 | 3 | 5 | 5 | | 7 | 6 | argininosuccinate synthase (argG) |
| SO0280 | 29 | 31 | 34 | 31 | 28 | 27 | 26 | 25 | penicillin-binding protein 1A (mrcA) |
| SO0281 | | 5 | 2 | 4 | 4 | 3 | | | type IV pilus biogenesis protein PilM |
| SO0282 | 5 | 7 | 3 | 5 | 5 | 6 | 3 | 2 | type IV pilus biogenesis protein PilN |
| SO0283 | 2 | 4 | | | 5 | | 3 | 4 | type IV pilus biogenesis protein PilO |
| SO0284 | 3 | 3 | 2 | 4 | | 5 | 3 | 2 | type IV pilus biogenesis protein PilP |
| SO0285 | 2 | 4 | 4 | 7 | 4 | 9 | 5 | 4 | type IV pilus biogenesis protein PilQ |
| SO0286 | 6 | 8 | 7 | 6 | 6 | 8 | 6 | 10 | shikimate kinase (aroK) |
| SO0287 | | 2 | 2 | | | | | | 3-dehydroquinate synthase (aroB) |
| SO0288 | 2 | 5 | 4 | 6 | 2 | 2 | 3 | 4 | damX domain protein |
| SO0289 | 2 | 3 | | 2 | | | 3 | 3 | DNA adenine methylase (dam) |
| SO0292 | 7 | 7 | 5 | 6 | 7 | 5 | 6 | 7 | ribulose-phosphate 3-epimerase (rpe) |
| SO0293 | 2 | | | | | | | | phosphoglycolate phosphatase (gph) |
| SO0294 | 8 | 9 | 13 | 13 | 9 | 8 | 4 | 10 | tryptophanyl-tRNA synthetase (trpS) |
| SO0295 | | | | | | 2 | | | transcriptional regulator, LysR family |
| SO0297 | 6 | 6 | 7 | 8 | 6 | 7 | 10 | 8 | lipoprotein, putative |
| SO0298 | 6 | 5 | 4 | 10 | 7 | 4 | 4 | 6 | phosphoheptose isomerase (gmhA) |
| SO0300 | 19 | 25 | 15 | 19 | 21 | 16 | 24 | 26 | lipoprotein, putative |
| SO0301 | 5 | 8 | 3 | 6 | 6 | 2 | 3 | 6 | conserved hypothetical protein TIGR00096 |
| SO0302 | | | 2 | | | | 2 | 2 | hypothetical protein |
| SO0311 | 4 | 16 | 3 | 8 | 7 | 6 | 4 | 7 | conserved hypothetical protein |
| SO0314 | 11 | 17 | 5 | 6 | 37 | 16 | 3 | 11 | ornithine decarboxylase, inducible (speF) |
| SO0316 | | | 2 | | | | | | conserved hypothetical protein TIGR00481 |
| SO0322 | | 2 | | | | | | | hypothetical protein |
| SO0323 | | | | | | | | 2 | hypothetical protein |
| SO0325 | | 2 | | 2 | 3 | 2 | | | dsrE-related protein |
| SO0326 | 5 | 4 | 5 | 8 | 7 | 7 | 5 | 4 | hypothetical protein |
| SO0330 | | | | | | | | 2 | conserved hypothetical protein |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO0333 | 8 | 11 | 16 | 17 | 11 | 8 | 13 | 16 | thiol:disulfide interchange protein DsbA (dsbA) |
| SO0335 | 8 | 8 | 8 | 7 | 6 | 8 | 6 | 7 | conserved hypothetical protein |
| SO0340 | 6 | 9 | 8 | 12 | 11 | 3 | 10 | 13 | branched-chain amino acid aminotransferase (ilvE) |
| SO0341 | 4 | 6 | | | 3 | 5 | | 4 | sensory box protein |
| SO0342 | | | 3 | 4 | 2 | 5 | 2 | 6 | conserved hypothetical protein |
| SO0343 | 13 | 9 | 21 | 32 | 11 | 10 | 28 | 27 | aconitate hydratase 1 (acnA) |
| SO0344 | 3 | 10 | 18 | 24 | 19 | 3 | 21 | 29 | methylcitrate synthase (prpC) |
| SO0345 | 8 | 9 | 13 | 16 | 11 | 9 | 15 | 17 | methylisocitrate lyase (prpB) |
| SO0346 | 2 | | 3 | 5 | | 3 | 4 | 5 | transcriptional regulator. GntR family |
| SO0347 | 6 | 5 | 6 | 6 | 5 | 3 | 5 | 9 | acyltransferase family protein |
| SO0348 | 12 | 17 | 15 | 15 | 16 | 11 | 16 | 18 | acyltransferase family protein |
| SO0350 | | 2 | 2 | 4 | | | 3 | | hypothetical protein |
| SO0352 | | | | | | 2 | | 2 | sensor histidine kinase, putative |
| SO0355 | 16 | 16 | 17 | 15 | 18 | 13 | 18 | 19 | AMP-binding protein |
| SO0356 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO0358 | 3 | 3 | 5 | 3 | 2 | 4 | 2 | | endoribonuclease L-PSP, putative |
| SO0359 | 7 | 11 | 4 | 14 | 13 | 9 | 16 | 14 | guanosine-3,5-bis(diphosphate) 3-pyrophosphohydrolase (spoT) |
| SO0360 | 7 | 8 | 11 | 8 | 7 | 4 | 10 | 10 | DNA-directed RNA polymerase, omega subunit (rpoZ) |
| SO0361 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | | guanylate kinase (gmk) |
| SO0362 | 15 | 18 | 18 | 27 | 20 | 18 | 23 | 23 | hypothetical protein |
| SO0364 | | 2 | | 3 | | | 2 | 3 | conserved hypothetical protein |
| SO0367 | | | | | | | | 2 | conserved hypothetical protein |
| SO0368 | | | 2 | | | | | | helicase |
| SO0369 | | | 3 | 4 | | 2 | 2 | 2 | transcriptional regulator, LysR family |
| SO0374 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO0375 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO0378 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | |
| SO0379 | | | | | | 2 | | | conserved hypothetical protein |
| SO0380 | | | | 2 | | | | 2 | type I restriction-modification system, R subunit (hsdR-1) |
| SO0382 | 2 | 3 | 2 | 3 | 3 | 3 | 3 | 2 | type I restriction-modification system, S subunit (hsdS-1) |
| SO0383 | 4 | 7 | 2 | 3 | 4 | 3 | 4 | 6 | type I restriction-modification system, M subunit (hsdM-1) |
| SO0388 | 5 | 3 | 3 | 4 | 5 | 2 | 3 | 3 | site-specific recombinase, phage integrase family |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO0391 | | | | | 2 | | | | hypothetical protein |
| SO0393 | 8 | 9 | 10 | 10 | 7 | 11 | 5 | 8 | DNA-binding protein Fis (fis) |
| SO0394 | | 4 | | | | | | 2 | conserved hypothetical protein |
| SO0395 | 6 | 4 | 2 | 6 | 3 | | 8 | 6 | ribosomal protein L11 methyltransferase (prmA) |
| SO0398 | 11 | 5 | 2 | | 23 | 11 | | 2 | fumarate reductase flavoprotein subunit (frdA) |
| SO0401 | | | | 3 | | | 5 | 2 | alcohol dehydrogenase, zinc-containing |
| SO0402 | 2 | 4 | 2 | 6 | 3 | 2 | | 2 | transcriptional regulator, LysR family |
| SO0403 | 2 | | | 3 | 2 | 3 | 2 | 2 | hypothetical protein |
| SO0404 | 101 | 128 | 98 | 99 | 128 | 125 | 68 | 71 | hypothetical protein |
| SO0405 | 46 | 48 | 44 | 47 | 41 | 43 | 43 | 42 | transcription termination factor Rho (rho) |
| SO0406 | 15 | 15 | 14 | 14 | 12 | 21 | 10 | 11 | thioredoxin 1 (trxA) |
| SO0407 | 26 | 28 | 21 | 23 | 27 | 18 | 25 | 16 | ATP-dependent RNA helicase, DEAD box family |
| SO0409 | | | 2 | 2 | 2 | | 3 | 3 | hypothetical protein |
| SO0414 | | | | | 3 | 4 | | | type 4 prepilin-like proteins leader peptide processing enzyme (pilD) |
| SO0415 | 2 | 3 | | 3 | 4 | 3 | 6 | 5 | type IV pilus biogenesis protein PilC |
| SO0416 | 2 | 6 | 5 | 3 | 5 | | 6 | 6 | type IV pilus biogenesis protein PilB |
| SO0417 | 3 | 4 | 4 | 5 | 3 | 2 | 3 | 5 | pilin, putative |
| SO0421 | | | | | | | 2 | 2 | AmpD protein (ampD) |
| SO0423 | | 5 | 3 | 3 | 2 | | 6 | 6 | pyruvate dehydrogenase complex repressor (pdhR) |
| SO0424 | 82 | 96 | 103 | 110 | 93 | 78 | 115 | 114 | pyruvate dehydrogenase complex, E1 component, pyruvate dehydrogenase (aceE) |
| SO0425 | 31 | 38 | 33 | 36 | 30 | 35 | 31 | 31 | pyruvate dehydrogenase complex, E2 component, dihydrolipoamide acetyltransferase (aceF) |
| SO0426 | 56 | 64 | 50 | 56 | 51 | 56 | 60 | 52 | pyruvate dehydrogenase complex, E3 component, lipoamide dehydrogenase (lpdA) |
| SO0427 | | | 2 | 2 | | 2 | 2 | 2 | sensory box protein |
| SO0428 | 9 | 7 | 10 | 5 | 11 | 13 | 7 | 13 | conserved hypothetical protein |
| SO0429 | 80 | 93 | 79 | 80 | 79 | 80 | 89 | 92 | peptidase, M13 family |
| SO0430 | | 2 | | 3 | | | | | conserved hypothetical protein |
| SO0431 | | 2 | 2 | 2 | | | | 3 | HAD-superfamily hydrolase, subfamily IA, variant 3 protein family |
| SO0432 | 78 | 87 | 67 | 74 | 75 | 62 | 63 | 74 | aconitate hydratase 2 (acnB) |
| SO0433 | 3 | 2 | 3 | 6 | 2 | | 4 | 3 | regulator of sigma D (rsd) |
| SO0435 | 8 | 13 | 8 | 10 | 8 | 8 | 7 | 10 | uroporphyrinogen decarboxylase (hemE) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
|---|---|---|---|---|---|---|---|---|---|
| SO0437 | 3 | 4 | 4 | 6 | | 4 | 4 | 6 | sensory box protein |
| SO0438 | 2 | 4 | | 3 | 4 | 3 | 6 | 6 | oxidoreductase, short chain dehydrogenase/reductase family |
| SO0439 | | | | | 2 | | | | hypothetical protein |
| SO0440 | 3 | 6 | 4 | 4 | 3 | 5 | 2 | 6 | conserved hypothetical protein |
| SO0441 | 14 | 22 | 16 | 16 | 16 | 16 | 8 | 12 | phosphoribosylamine--glycine ligase (purD) |
| SO0442 | 23 | 31 | 23 | 26 | 21 | 16 | 13 | 20 | phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase (purH) |
| SO0443 | | | 2 | 2 | | | | | transcriptional regulator, MerR family |
| SO0444 | 7 | 6 | 11 | 13 | 6 | | 10 | 7 | hypothetical protein |
| SO0445 | | | | | 4 | 4 | | | hflC protein, putative |
| SO0449 | 8 | 10 | 8 | 7 | 4 | 4 | 9 | 7 | conserved hypothetical protein |
| SO0452 | 6 | 3 | 5 | 9 | 3 | 3 | 7 | 7 | thioredoxin 2 (trxC) |
| SO0453 | 9 | 8 | 8 | 8 | 6 | 2 | 5 | 8 | peptidyl-prolyl cis-trans isomerase FkbP (fkbP-1) |
| SO0456 | 8 | 8 | 9 | 9 | 7 | 5 | 7 | 10 | immunogenic-related protein |
| SO0459 | | 3 | 2 | | 3 | 2 | 2 | 2 | conserved hypothetical protein |
| SO0463 | | | 2 | | | | | | conserved hypothetical protein |
| SO0466 | | | | | | | | 2 | hypothetical protein |
| SO0467 | 8 | 11 | 3 | 4 | 6 | 2 | 3 | 7 | DNA helicase II (uvrD) |
| SO0470 | | | | 2 | | | | | hypothetical protein |
| SO0471 | | 2 | | | 2 | | | | conserved hypothetical protein |
| SO0474 | | 2 | 2 | 3 | 5 | 2 | 2 | | conserved hypothetical protein |
| SO0481 | | | | 2 | | | | | peptidyl-prolyl cis-trans isomerase, FKBP-type |
| SO0490 | 2 | 4 | | | 9 | 6 | 4 | | transcriptional regulator |
| SO0491 | 32 | 46 | 36 | 32 | 36 | 33 | 31 | 34 | peptidase, M13 family |
| SO0492 | | 2 | | 2 | | 2 | | 2 | conserved hypothetical protein |
| SO0494 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO0495 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO0496 | | | 3 | | 2 | 2 | | | conserved hypothetical protein |
| SO0500 | 4 | 3 | 2 | 4 | 2 | 3 | 2 | 2 | methyl-accepting chemotaxis protein |
| SO0501 | 17 | 19 | 22 | 23 | 18 | 17 | 20 | 24 | conserved hypothetical protein |
| SO0502 | | | | | | | 2 | | transcriptional regulator, ArsR family |
| SO0506 | 8 | 8 | 11 | 10 | 13 | 8 | 9 | 12 | conserved hypothetical protein TIGR00148 |
| SO0508 | 2 | 3 | 4 | 3 | 3 | 3 | | 3 | hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO0510 | | | | 2 | | | | | oxidoreductase, short-chain dehydrogenase/reductase family |
| SO0511 | | 2 | | | | | | | acetyl-CoA carboxylase, biotin carboxyl carrier protein (accB) |
| SO0513 | 2 | | | | | 2 | | | prokaryotic and mitochondrial release factors family protein |
| SO0514 | | | | | 2 | | | | hypothetical protein |
| SO0515 | 2 | 2 | 3 | | | | 2 | 2 | hypothetical protein |
| SO0518 | 9 | 13 | 15 | 14 | 9 | 16 | 15 | 19 | outer membrane efflux family protein, putative |
| SO0519 | 3 | 5 | 9 | 6 | 4 | 2 | 9 | 9 | cation efflux protein, putative |
| SO0520 | 10 | 18 | 11 | 11 | 14 | 12 | 17 | 18 | heavy metal efflux pump, CzcA family |
| SO0521 | 2 | 2 | 4 | | 4 | | 2 | 4 | conserved hypothetical protein |
| SO0526 | | | | | | | | 2 | acetyltransferase, GNAT family |
| SO0527 | | 4 | 3 | 4 | 4 | | 4 | 3 | conserved hypothetical protein |
| SO0528 | 2 | | 2 | | 2 | | 2 | 3 | conserved hypothetical protein |
| SO0532 | 3 | | | | | | | | arsenical resistence operon repressor (arsR) |
| SO0538 | 4 | 2 | 5 | 6 | 5 | 3 | 8 | 6 | glyceraldehyde 3-phosphate dehydrogenase (gapA-1) |
| SO0541 | | | | | | | 2 | | metallo-beta-lactamase family protein |
| SO0542 | 4 | 3 | 4 | | 5 | | 4 | 5 | conserved hypothetical protein |
| SO0543 | | | | | | | | 2 | hypothetical protein |
| SO0544 | | | | | | | 2 | | sensory box histidine kinase |
| SO0546 | | | | | 3 | 2 | 2 | | ribosomal protein S6 modification protein (rimK-1) |
| SO0548 | 7 | 14 | 5 | 10 | 14 | 13 | 4 | 5 | DNA-binding protein, HU family |
| SO0549 | | | | | | 2 | | | chemotaxis protein CheY/response regulator receiver domain protein |
| SO0551 | | | | 2 | | | 2 | | conserved hypothetical protein |
| SO0554 | 2 | 5 | 5 | 5 | 3 | 2 | | 4 | hypothetical protein |
| SO0555 | 2 | | | | | | | | conserved hypothetical protein |
| SO0556 | | | | | 3 | | | 2 | hypothetical protein |
| SO0558 | 2 | 4 | 2 | 2 | 2 | | | 2 | smtA protein (smtA) |
| SO0559 | | 3 | 2 | 3 | 3 | | 3 | 3 | MaoC domain protein |
| SO0560 | 3 | | | | 5 | | | | formate--tetrahydrofolate ligase (fhs) |
| SO0564 | | 3 | 2 | | | 2 | 5 | 3 | conserved hypothetical protein |
| SO0565 | 2 | 3 | 2 | 2 | 2 | 3 | | 3 | adhesion protein, putative |
| SO0567 | 4 | 7 | 2 | 8 | 4 | 3 | 2 | 5 | 1-acyl-sn-glycerol-3-phosphate acyltransferase (plsC) |
| SO0568 | | 3 | 2 | 2 | | | 3 | 4 | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO0570 | | | | | | | | 2 | response regulator |
| SO0572 | | | | 2 | | | | | enoyl-CoA hydratase/isomerase family protein |
| SO0575 | 27 | 27 | 18 | 22 | 21 | 12 | 19 | 17 | RNA polymerase-associated protein HepA (hepA) |
| SO0576 | | | | | | | 5 | 2 | PhoH family protein |
| SO0577 | 14 | 17 | 8 | 11 | 12 | 10 | 16 | 20 | sensory box histidine kinase/response regulator |
| SO0578 | 24 | 27 | 32 | 30 | 26 | 26 | 29 | 35 | hypothetical protein |
| SO0583 | | | | 2 | | | | | bacterioferritin-associated ferredoxin (bfd) |
| SO0584 | 3 | 2 | 2 | 2 | 4 | 2 | 2 | 2 | methyl-accepting chemotaxis protein |
| SO0585 | | | | | | | 2 | | D-isomer specific 2-hydroxyacid dehydrogenase family protein |
| SO0588 | 15 | 22 | 20 | 21 | 15 | 26 | 18 | 20 | transporter, putative |
| SO0591 | 10 | 12 | 9 | 11 | 9 | 5 | 11 | 18 | conserved hypothetical protein TIGR00157 |
| SO0592 | | 5 | 4 | 4 | 3 | | | 3 | oligoribonuclease (orn) |
| SO0595 | 3 | 2 | 2 | 3 | 3 | | | | hypothetical protein |
| SO0596 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO0598 | | | | | 3 | | | | yjeF protein (yjeF) |
| SO0600 | 10 | 9 | 9 | 6 | 8 | 10 | 7 | 5 | N-acetylmuramoyl-L-alanine amidase (amiB) |
| SO0601 | 3 | 4 | 3 | 4 | 4 | 3 | 4 | 4 | DNA mismatch repair protein MutL (mutL) |
| SO0602 | 6 | 10 | 9 | 7 | 8 | 3 | 6 | 12 | tRNA delta(2)-isopentenylpyrophosphate transferase (miaA) |
| SO0603 | 7 | 8 | 8 | 9 | 7 | 6 | 8 | 6 | host factor-I protein (hfq) |
| SO0604 | 3 | 4 | 3 | 3 | 2 | | 5 | 5 | GTP-binding protein HflX (hflX) |
| SO0605 | 29 | 43 | 40 | 37 | 29 | 31 | 35 | 43 | hflK protein (hflK) |
| SO0606 | 42 | 57 | 46 | 46 | 41 | 43 | 51 | 52 | hflC protein (hflC) |
| SO0608 | 14 | 19 | 16 | 14 | 11 | 18 | 14 | 15 | ubiquinol-cytochrome c reductase, iron-sulfur subunit (petA) |
| SO0609 | 4 | 6 | 5 | 4 | 6 | 4 | 5 | 5 | ubiquinol-cytochrome c reductase, cytochrome b (petB) |
| SO0610 | 13 | 15 | 12 | 11 | 13 | 9 | 14 | 15 | ubiquinol-cytochrome c reductase, cytochrome c1 (petC) |
| SO0611 | 5 | 4 | 4 | 5 | 3 | 5 | 6 | 6 | stringent starvation protein a (sspA) |
| SO0612 | 6 | 4 | 5 | 3 | 4 | 2 | 3 | 3 | stringent starvation protein b (sspB) |
| SO0614 | 23 | 24 | 20 | 14 | 18 | 21 | 16 | 17 | dipeptidyl peptidase IV, putative |
| SO0617 | 19 | 22 | 15 | 13 | 14 | 18 | 16 | 19 | acetylornithine aminotransferase (argD) |
| SO0618 | 4 | 3 | 4 | 4 | 8 | 2 | 3 | 4 | arginine N-succinyltransferase (astA) |
| SO0619 | 13 | 13 | 9 | 11 | 10 | 12 | 9 | 11 | succinylglutamic semialdehyde dehydrogenase (astD) |
| SO0620 | | | | | | | | 2 | conserved hypothetical protein |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| **Locus** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Description** |
| SO0621 | | | 2 | | 2 | | 2 | 5 | sensor histidine kinase |
| SO0622 | | | | | 4 | | | 3 | DNA-binding response regulator |
| SO0624 | 13 | 17 | 16 | 17 | 18 | 13 | 12 | 21 | catabolite gene activator (crp) |
| SO0625 | | | 3 | 2 | | 2 | | | conserved domain protein |
| SO0630 | 3 | | 4 | 5 | | 2 | 4 | 7 | TonB-dependent receptor (nosA) |
| SO0632 | | | | | 4 | | 2 | | ATP-dependent helicase HrpB (hrpB) |
| SO0633 | 25 | 27 | 20 | 17 | 28 | 12 | 23 | 20 | penicillin-binding protein 1B (mrcB) |
| SO0635 | 2 | 2 | 2 | 2 | 3 | 7 | 6 | 7 | peptidyl-prolyl cis-trans isomerase C (ppiC-1) |
| SO0636 | | | 2 | 2 | 2 | | 2 | 2 | acetyltransferase, GNAT family |
| SO0640 | 4 | 3 | 2 | 3 | 4 | 3 | 3 | 3 | alcohol dehydrogenase, zinc-containing |
| SO0641 | 2 | | 3 | | 3 | | | | prophage MuSo1, transcriptional regulator, Cro/CI family |
| SO0643 | | | | | | | | 2 | transposase, putative |
| SO0644 | | | | 2 | | | | 2 | prophage MuSo1, DNA transposition protein, putative |
| SO0646 | | | | 2 | | | 2 | | hypothetical protein |
| SO0653 | | | | 2 | | | | | hypothetical protein |
| SO0655 | | | 2 | 2 | | | | | hypothetical protein |
| SO0656 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO0657 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO0667 | 2 | | | | | | | | conserved hypothetical protein |
| SO0670 | 3 | | 3 | 2 | 4 | | 2 | 6 | hypothetical protein |
| SO0672 | | | 2 | | | | | | hypothetical protein |
| SO0673 | 4 | 2 | | 3 | 3 | | 3 | 4 | hypothetical protein |
| SO0680 | | 3 | | 2 | | 4 | 2 | 3 | hypothetical protein |
| SO0684 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO0685 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO0688 | | | | | | | 2 | | hypothetical protein |
| SO0691 | 9 | 15 | 12 | 12 | 10 | 15 | 8 | 10 | hypothetical protein |
| SO0693 | | | | 4 | | | | 2 | aldose 1-epimerase (galM) |
| SO0694 | | | | | | | | 2 | galactokinase (galK) |
| SO0695 | 8 | 9 | 11 | 10 | 8 | 4 | 8 | 8 | glutathione-regulated potassium-efflux system protein KefC, putative |
| SO0696 | 7 | 9 | 6 | 6 | 10 | 6 | 8 | 9 | thiol:disulfide interchange protein DsbD (dsbD) |
| SO0698 | | 2 | | | | 2 | | 4 | fxsA protein (fsxA) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO0701 | | | 3 | | 2 | 2 | | 2 | transcriptional regulator, LysR family |
| SO0703 | 13 | 15 | 18 | 18 | 14 | 19 | 16 | 15 | chaperonin GroES (groES) |
| SO0704 | 122 | 125 | 129 | 133 | 114 | 140 | 132 | 135 | chaperonin GroEL (groEL) |
| SO0708 | | | | 2 | | | | | transposase, mutator family |
| SO0709 | | | 2 | | | | | | hypothetical protein |
| SO0719 | 22 | 27 | 26 | 28 | 21 | 19 | 27 | 29 | TonB-dependent receptor, putative |
| SO0728 | 2 | | | | | | | | conserved hypothetical protein |
| SO0730 | 2 | 2 | | 3 | 6 | 2 | | 3 | hypothetical protein |
| SO0733 | | 3 | | 3 | 3 | 3 | | 2 | cold shock domain family protein |
| SO0740 | 3 | 4 | 5 | 5 | 3 | 4 | 6 | 6 | melanin biosynthesis protein TyrA, putative |
| SO0741 | | 7 | | | 2 | | | | gamma-glutamyltranspeptidase (ggt-1) |
| SO0742 | | 2 | 4 | 4 | | | 4 | | iron(III) ABC transporter, ATP-binding protein |
| SO0744 | 6 | 5 | 5 | | | 2 | 2 | 2 | iron(III) ABC transporter, periplasmic iron(III)-binding protein |
| SO0746 | | | | | | | 2 | | glutathione S-transferase family protein |
| SO0747 | | 2 | | | | | | | ferredoxin--NADP reductase (fpr) |
| SO0749 | | | | | | | 2 | 2 | conserved hypothetical protein |
| SO0750 | 4 | 8 | 5 | 6 | 8 | 7 | 6 | 4 | glutamate synthase, putative |
| SO0752 | 2 | | | 3 | | | | 2 | hypothetical protein |
| SO0754 | 2 | 3 | 2 | 3 | | 3 | 4 | 4 | ABC transporter, ATP-binding protein |
| SO0756 | 12 | 13 | 14 | 12 | 16 | 10 | 14 | 14 | phospho-2-dehydro-3-deoxyheptonate aldolase, phe-sensitive (aroG) |
| SO0764 | 2 | 5 | 7 | 5 | 2 | | 6 | 2 | hypothetical protein |
| SO0768 | 2 | 3 | | 2 | | 3 | 4 | | conserved hypothetical protein |
| SO0769 | 7 | 10 | 7 | 7 | 6 | 9 | 6 | 5 | arginine repressor (argR) |
| SO0770 | 34 | 30 | 31 | 38 | 26 | 29 | 30 | 33 | malate dehydrogenase (mdh) |
| SO0775 | 5 | 4 | | 4 | 3 | | 2 | 2 | conserved hypothetical protein |
| SO0777 | 2 | 4 | 2 | 2 | 4 | 2 | 2 | 2 | 2-octaprenyl-6-methoxyphenol hydroxylase (ubiH) |
| SO0778 | 8 | 10 | 7 | 8 | 10 | 8 | 2 | 7 | oxidoreductase, FAD-binding, UbiH/Coq6 family |
| SO0779 | 12 | 19 | 18 | 22 | 15 | 13 | 14 | 18 | glycine cleavage system T protein (gcvT) |
| SO0781 | 24 | 39 | 27 | 42 | 40 | 21 | 39 | 41 | glycine cleavage system P protein (gcvP) |
| SO0783 | 4 | 7 | 9 | 14 | 10 | 4 | 8 | 12 | hypothetical protein |
| SO0788 | 5 | | | 2 | 2 | | 3 | | conserved hypothetical protein |
| SO0789 | | 2 | 2 | | | | | 2 | conserved hypothetical protein |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO0795 | 3 | 12 | 3 | 8 | 10 | | 11 | 12 | conserved hypothetical protein |
| SO0798 | 3 | 2 | 6 | 12 | | | 11 | 13 | conserved hypothetical protein |
| SO0801 | | 2 | | | | | 2 | | conserved hypothetical protein |
| SO0804 | | 2 | | | | | | | hypothetical protein |
| SO0805 | 7 | 5 | 7 | 5 | 4 | 3 | 5 | 9 | CBS domain protein |
| SO0807 | 5 | 8 | 6 | 5 | 5 | 3 | 3 | 5 | hypoxanthine-guanine phosphoribosyltransferase (hpt-1) |
| SO0808 | 2 | 4 | 2 | 3 | 3 | | 2 | 2 | conserved hypothetical protein |
| SO0809 | 2 | 2 | 3 | 3 | | 2 | 2 | 3 | azurin (azu) |
| SO0810 | 2 | 3 | 3 | 2 | 3 | 3 | 4 | | ribokinase (rbsK) |
| SO0811 | 4 | 2 | 2 | 7 | 4 | 2 | 2 | 5 | inosine-uridine preferring nucleoside hydrolase family protein |
| SO0812 | | | | | 2 | | | | hypothetical protein |
| SO0813 | 16 | 17 | 12 | 11 | 14 | 12 | 15 | 18 | hypothetical protein |
| SO0815 | 28 | 36 | 28 | 33 | 29 | 25 | 43 | 41 | TonB-dependent receptor C-terminal domain protein |
| SO0816 | 2 | 2 | 3 | 3 | 2 | 2 | 2 | 3 | hypothetical protein |
| SO0818 | | | | | | 2 | | | 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (metE) |
| SO0820 | | | | 2 | | | 4 | | HlyD family secretion protein |
| SO0821 | | | 3 | 2 | | 2 | 3 | 3 | ABC transporter, ATP-binding/permease protein |
| SO0822 | | | | 2 | | | | | outer membrane efflux family protein |
| SO0823 | | | | 3 | 2 | 3 | 3 | | conserved hypothetical protein |
| SO0828 | 4 | 5 | 2 | 3 | 3 | 2 | 2 | 5 | conserved hypothetical protein |
| SO0830 | 2 | 6 | 4 | 8 | 3 | 4 | 5 | 4 | alkaline phosphatase |
| SO0831 | 9 | 12 | 8 | 13 | 11 | 3 | 9 | 12 | glutathione synthetase (gshB) |
| SO0832 | 8 | 9 | 7 | 9 | 6 | 4 | 5 | 7 | conserved hypothetical protein TIGR00046 |
| SO0833 | | 3 | | | 2 | | | | endonuclease I (endA) |
| SO0834 | 2 | 2 | 2 | 2 | 2 | 2 | | | sprT protein, putative |
| SO0835 | | 3 | 2 | 2 | | | | | conserved hypothetical protein |
| SO0837 | | | | 2 | | | 2 | 2 | beta-lactamase, putative |
| SO0839 | | | | 2 | 2 | | | | transcriptional regulator, LysR family |
| SO0840 | 148 | 154 | 133 | 136 | 166 | 139 | 156 | 159 | acetyl-CoA carboxylase multifunctional enzyme accADC, carboxyl transferase subunit alpha/carboxyl transferase subunit beta/biotin carboxylase |
| SO0842 | 71 | 82 | 71 | 79 | 70 | 57 | 85 | 81 | translation elongation factor G (fusA-2) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO0845 | 6 | 6 | 3 | 3 | 6 | 6 | 4 | 3 | cytochrome c-type protein NapB (napB) |
| SO0846 | | 2 | | | | 3 | | | iron-sulfur cluster-binding protein napH (napH) |
| SO0847 | 4 | 2 | | | 8 | 6 | | | iron-sulfur cluster-binding protein NapG (napG) |
| SO0848 | 49 | 59 | 33 | 45 | 74 | 66 | 23 | 34 | periplasmic nitrate reductase (napA) |
| SO0853 | | | | | | | 2 | | pilin, putative |
| SO0855 | | 2 | | | 3 | 2 | | 2 | conserved hypothetical protein |
| SO0856 | | | | | 2 | | | | conserved hypothetical protein |
| SO0859 | 2 | 4 | 9 | | 4 | 5 | 3 | 9 | sensory box histidine kinase/response regulator |
| SO0860 | | 2 | | | | | 4 | 2 | response regulator |
| SO0861 | | 2 | 2 | 2 | | | | 2 | conserved hypothetical protein |
| SO0862 | 11 | 15 | 17 | 16 | 17 | 11 | 15 | 15 | D-3-phosphoglycerate dehydrogenase (serA) |
| SO0864 | | | | | | | | 3 | transcriptional regulator, LuxR family |
| SO0869 | 6 | 6 | | 3 | 3 | | 5 | 2 | pantoate--beta-alanine ligase (panC) |
| SO0870 | 5 | 6 | 3 | 3 | 4 | 4 | 4 | 5 | 3-methyl-2-oxobutanoate hydroxymethyltransferase (panB) |
| SO0871 | | 3 | 3 | 4 | 2 | | 2 | 2 | 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase (folK-1) |
| SO0872 | 12 | 13 | 11 | 17 | 18 | 11 | 11 | 16 | polyA polymerase (pcnB) |
| SO0874 | 8 | 12 | 8 | 11 | 10 | 9 | 5 | 9 | DnaK suppressor protein (dksA) |
| SO0876 | 10 | 9 | 7 | 10 | 12 | 4 | 10 | 10 | peptidase B (pepB) |
| SO0880 | 2 | | | | | | | | conserved hypothetical protein |
| SO0881 | 2 | | | | | | | | conserved hypothetical protein |
| SO0882 | 4 | 5 | 2 | 2 | 5 | 6 | 4 | 3 | oxidoreductase, GMC family |
| SO0885 | 2 | 4 | 2 | 4 | 2 | 2 | 3 | 2 | ABC transporter, ATP-binding protein |
| SO0887 | 16 | 15 | 14 | 13 | 7 | 11 | 7 | 18 | conserved hypothetical protein |
| SO0888 | 11 | 19 | 10 | 20 | 8 | 6 | 12 | 17 | amidase family protein |
| SO0891 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO0892 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO0897 | 11 | 11 | 3 | 2 | 7 | 8 | 7 | 6 | ATP-dependent RNA helicase DbpA (dbpA) |
| SO0899 | | | | 3 | | | | | glyoxalase family protein |
| SO0900 | 5 | 5 | 3 | 6 | 8 | 3 | 8 | 10 | oxidoreductase, aldo/keto reductase family |
| SO0902 | 19 | 24 | 10 | 8 | 23 | 18 | 7 | 4 | NADH:ubiquinone oxidoreductase, Na translocating, alpha subunit (nqrA-1) |
| SO0904 | 2 | 4 | 2 | 5 | 7 | 6 | 3 | 4 | NADH:ubiquinone oxidoreductase, gamma subunit (nqrC-1) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
|---|---|---|---|---|---|---|---|---|---|
| SO0907 | 12 | 13 | 7 | 9 | 16 | 11 | 9 | 9 | NADH:ubiquinone oxidoreductase, Na translocating, beta subunit (nqrF-1) |
| SO0912 | | | | | | | 3 | 3 | hypothetical protein |
| SO0915 | | | 2 | | | | | 2 | ankyrin domain protein |
| SO0916 | 2 | | | 2 | | | 3 | | transcriptional regulator, MarR family |
| SO0918 | 11 | 11 | 7 | 7 | 7 | 15 | 5 | 4 | aculeacin A acylase (aac) |
| SO0919 | 2 | 3 | 2 | | 2 | | 2 | 2 | serine transporter, putative |
| SO0923 | 13 | 10 | 13 | 12 | 12 | 9 | 12 | 11 | conserved hypothetical protein |
| SO0929 | 21 | 21 | 21 | 21 | 13 | 17 | 15 | 27 | S-adenosylmethionine synthetase (metK) |
| SO0930 | 28 | 36 | 29 | 33 | 32 | 21 | 32 | 36 | transketolase (tkt) |
| SO0931 | 3 | 5 | 2 | 3 | | 2 | 5 | 6 | D-erythrose-4-phosphate dehydrogenase (epd) |
| SO0932 | 25 | 21 | 21 | 23 | 17 | 19 | 11 | 16 | phosphoglycerate kinase (pgk) |
| SO0933 | 21 | 34 | 29 | 27 | 23 | 33 | 31 | 32 | fructose-bisphosphate aldolase, class II, Calvin cycle subtype (fba) |
| SO0934 | 9 | 10 | 12 | 14 | 2 | 12 | 22 | 24 | conserved hypothetical protein |
| SO0940 | | | | | 2 | | | | transcriptional regulator-related protein |
| SO0942 | | 2 | | 2 | 2 | 2 | | 3 | conserved hypothetical protein |
| SO0943 | 6 | 9 | 6 | 5 | 5 | 5 | 6 | 6 | sensory box protein, putative |
| SO0945 | | 2 | | | | | 2 | | AcrB/AcrD/AcrF family protein |
| SO0946 | 2 | 3 | | | 2 | | | | conserved hypothetical protein |
| SO0947 | 17 | 19 | 12 | 14 | 15 | 11 | 11 | 10 | ATP-dependent RNA helicase SrmB (srmB) |
| SO0951 | 10 | 8 | 13 | 11 | 4 | 4 | 9 | 8 | thiol:disulfide interchange protein DsbC (dsbC) |
| SO0952 | | | | | | | 3 | | single-stranded-DNA-specific exonuclease RecJ (recJ) |
| SO0956 | 4 | | 2 | 4 | | 2 | 2 | 7 | alkyl hydroperoxide reductase, F subunit (ahpF) |
| SO0958 | 14 | 19 | 18 | 19 | 15 | 11 | 12 | 14 | alkyl hydroperoxide reductase, C subunit (ahpC) |
| SO0959 | 22 | 20 | 23 | 18 | 13 | 18 | 17 | 15 | cytosol aminopeptidase (pepA-1) |
| SO0961 | | 2 | 3 | 3 | | | | | conserved hypothetical protein |
| SO0968 | 2 | 3 | 2 | 4 | 3 | | | 3 | D-lactate dehydrogenase (ldhA) |
| SO0970 | 24 | 22 | 31 | 27 | 39 | 37 | 14 | 18 | fumarate reductase flavoprotein subunit precursor |
| SO0976 | | | | | | 2 | | | organic hydroperoxide resistance protein (ohr) |
| SO0977 | | | 2 | 3 | 2 | | | 5 | transcriptional regulator, MarR family |
| SO0978 | | | | | | 2 | | | FAD-dependent glycerol-3-phosphate dehydrogenase, family protein |
| SO0980 | | 2 | | | | | | | RNA methyltransferase, TrmA family |
| SO0982 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO0983 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO0987 | 21 | 23 | 10 | 16 | 24 | 24 | 17 | 21 | methyl-accepting chemotaxis protein |
| SO0988 | 3 | 2 | | | 6 | 3 | 2 | | formate dehydrogenase, alpha subunit |
| SO0992 | 11 | 18 | 9 | 20 | 9 | 11 | 14 | 16 | lysyl-tRNA synthetase (lysS) |
| SO0994 | 11 | 10 | 9 | 11 | 14 | 6 | 7 | 11 | conserved hypothetical protein |
| SO0996 | 2 | 3 | 5 | 4 | 2 | | 2 | 3 | glyoxalase family protein |
| SO1002 | 2 | 3 | | | | 3 | | 2 | hypothetical protein |
| SO1003 | 7 | 13 | 8 | 9 | 13 | 21 | 12 | 8 | hypothetical protein |
| SO1004 | 17 | 26 | 25 | 17 | 21 | 26 | 25 | 22 | hypothetical protein |
| SO1006 | | | | | | | 2 | | dienelactone hydrolase family protein |
| SO1014 | 9 | 14 | 16 | 10 | 12 | 9 | 16 | 11 | NADH dehydrogenase I, I subunit (nuoI) |
| SO1016 | 16 | 29 | 20 | 24 | 23 | 18 | 29 | 29 | NADH dehydrogenase I, G subunit (nuoG) |
| SO1017 | 21 | 25 | 23 | 22 | 22 | 19 | 20 | 14 | NADH dehydrogenase I, F subunit (nuoF) |
| SO1018 | 4 | 2 | 3 | | 2 | | 2 | 3 | NADH dehydrogenase I, E subunit (nuoE) |
| SO1019 | 23 | 34 | 19 | 20 | 27 | 22 | 19 | 22 | NADH dehydrogenase I, C/D subunits (nuoCD) |
| SO1020 | 7 | 9 | 5 | 7 | 7 | 7 | 4 | 6 | NADH dehydrogenase I, B subunit (nuoB) |
| SO1021 | | 3 | | | | | | 2 | NADH dehydrogenase I, A subunit (nuoA) |
| SO1024 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO1025 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO1026 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO1030 | | 5 | | 4 | 4 | | 7 | 3 | 5-methyltetrahydrofolate--homocysteine methyltransferase (metH) |
| SO1033 | | | | | 2 | | | | iron-compound ABC transporter, ATP-binding protein, putative |
| SO1035 | | | 3 | 3 | | | 4 | 4 | nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase (cobT) |
| SO1037 | 2 | | | | | | | | cobinamide kinase/cobinamide phosphate guanylyltransferase (cobU) |
| SO1038 | | 4 | 2 | | | | | | cobyric acid synthase CobQ (cobQ) |
| SO1039 | | | 2 | 2 | | | | | cob(I)alamin adenosyltransferase (cobO) |
| SO1042 | | | | | | | | 2 | amino acid ABC transporter, ATP-binding protein |
| SO1043 | | | | | | | 2 | | amino acid ABC transporter, permease protein |
| SO1044 | | | | 2 | | | | | amino acid ABC transporter, periplasmic amino acid-binding protein |
| SO1045 | | | | 3 | | | 5 | 5 | hypothetical protein |
| SO1051 | 35 | 41 | 34 | 37 | 37 | 27 | 35 | 36 | periplasmic glucan biosynthesis protein, putative |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1052 | 3 | 3 | 2 | 2 | 2 | | | | low-affinity inorganic phosphate transporter (pit) |
| SO1056 | 27 | 31 | 21 | 20 | 31 | 16 | 25 | 25 | methyl-accepting chemotaxis protein |
| SO1059 | 14 | 31 | 16 | 22 | 13 | 21 | 14 | 12 | aminopeptidase N |
| SO1060 | 15 | 20 | 19 | 20 | 22 | 13 | 24 | 24 | conserved hypothetical protein |
| SO1061 | 10 | 17 | 12 | 12 | 10 | 13 | 14 | 11 | TPR domain protein |
| SO1062 | 3 | | 3 | 3 | 2 | | 3 | 2 | polypeptide deformylase (def-2) |
| SO1063 | 2 | | | | | | | | slyX protein (slyX) |
| SO1065 | 33 | 40 | 29 | 32 | 37 | 43 | 39 | 38 | FKBP-type peptidyl-prolyl cis-trans isomerase FkpA (fkpA) |
| SO1066 | 11 | 8 | 8 | 5 | 19 | 24 | 7 | 17 | extracellular nuclease |
| SO1068 | | | 2 | 3 | | | 2 | 2 | conserved hypothetical protein |
| SO1070 | | | | | 3 | | | 2 | catalase (katB) |
| SO1072 | 5 | 6 | 12 | 12 | | 2 | 4 | 4 | chitin-binding protein, putative |
| SO1075 | 24 | 23 | 17 | 23 | 20 | 24 | 14 | 20 | conserved hypothetical protein |
| SO1079 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO1080 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO1082 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO1083 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO1093 | | | 2 | 3 | | | | 3 | ISSo7, transposase |
| SO1094 | | | 2 | | | | | | conserved hypothetical protein |
| SO1095 | 3 | 6 | 8 | 11 | 10 | 3 | 9 | 14 | O-acetylhomoserine (thiol)-lyase, putative |
| SO1096 | | 3 | 3 | | 2 | 2 | 2 | 2 | hypothetical protein |
| SO1097 | 2 | | | 2 | 2 | | | | conserved hypothetical protein |
| SO1099 | | 2 | 2 | | 2 | | | 2 | bolA protein (bolA) |
| SO1101 | 9 | 8 | 6 | 8 | 5 | 4 | 5 | 8 | autoinducer-2 production protein LuxS (luxS) |
| SO1103 | 44 | 55 | 47 | 42 | 45 | 48 | 49 | 55 | NADH:ubiquinone oxidoreductase, Na translocating, alpha subunit (nqrA-2) |
| SO1104 | 3 | 2 | 2 | | 2 | 2 | 2 | | NADH:ubiquinone oxidoreductase, Na translocating, hydrophobic membrane protein NqrB (nqrB-2) |
| SO1105 | 16 | 27 | 18 | 21 | 18 | 16 | 15 | 14 | NADH:ubiquinone oxidoreductase, Na translocating, gamma subunit (nqrC-2) |
| SO1108 | 43 | 42 | 36 | 37 | 42 | 31 | 44 | 43 | NADH:ubiquinone oxidoreductase, Na translocating, beta subunit (nqrF-2) |
| SO1109 | 15 | 15 | 12 | 12 | 13 | 11 | 15 | 12 | thiamin biosynthesis lipoprotein ApbE (apbE) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| **Locus** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Description** |
| SO1111 | 5 | 5 | 2 | 5 | 8 | 12 | 3 | 4 | bacterioferritin subunit 2 (bfr2) |
| SO1112 | 4 | 7 | 3 | 2 | 8 | 10 | 3 | 3 | bacterioferritin subunit 1 (bfr1) |
| SO1114 | | 2 | 6 | 11 | | | 5 | 13 | DNA-damage-inducible protein P (dinP) |
| SO1115 | 7 | 10 | 9 | 10 | 10 | 5 | 16 | 12 | aminoacyl-histidine dipeptidase (pepD) |
| SO1117 | 16 | 16 | 15 | 17 | 22 | 16 | 15 | 20 | cytosol aminopeptidase, putative |
| SO1121 | 7 | 3 | 4 | 3 | 6 | 5 | 4 | 4 | glutamate 5-kinase (proB) |
| SO1122 | 5 | 3 | 4 | 4 | 2 | 4 | 4 | 5 | gamma-glutamyl phosphate reductase (proA) |
| SO1124 | 3 | 3 | | | | 3 | | 3 | conserved hypothetical protein TIGR00011 |
| SO1126 | 108 | 121 | 138 | 147 | 114 | 125 | 137 | 131 | chaperone protein DnaK (dnaK) |
| SO1127 | 39 | 44 | 42 | 50 | 51 | 39 | 47 | 56 | chaperone protein DnaJ (dnaJ) |
| SO1129 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO1130 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO1133 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO1134 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO1136 | 2 | 3 | 3 | 2 | 2 | | 2 | 2 | ATP-dependent RNA helicase, DEAD box family |
| SO1137 | 8 | 7 | 13 | 6 | 7 | 6 | 13 | 11 | conserved hypothetical protein |
| SO1139 | 7 | 10 | 8 | 8 | 7 | 7 | 6 | 9 | peptidyl-prolyl cis-trans isomerase FklB (fklB) |
| SO1140 | 5 | 4 | 7 | 6 | 5 | 5 | 9 | 8 | dihydrodipicolinate reductase (dapB) |
| SO1141 | 25 | 25 | 21 | 23 | 21 | 28 | 14 | 16 | carbamoyl-phosphate synthase, small subunit (carA) |
| SO1142 | 71 | 66 | 52 | 44 | 67 | 47 | 43 | 44 | carbamoyl-phosphate synthase, large subunit (carB) |
| SO1144 | 15 | 13 | 16 | 15 | 15 | 12 | 13 | 16 | methyl-accepting chemotaxis protein |
| SO1149 | 5 | 7 | 4 | 5 | 5 | 6 | 6 | 6 | conserved hypothetical protein |
| SO1150 | 9 | 8 | 7 | 8 | 8 | 10 | 8 | 7 | ribose 5-phosphate isomerase (rpiA) |
| SO1154 | 2 | | | | | | | | hypothetical protein |
| SO1156 | | | | | | 2 | | | TonB-dependent receptor |
| SO1158 | 2 | | | 2 | | | | | Dps family protein |
| SO1161 | 14 | 19 | 12 | 20 | 14 | 15 | 17 | 17 | lipoic acid synthetase (lipA) |
| SO1162 | | | | 5 | 5 | | 6 | 6 | lipoate-protein ligase B (lipB) |
| SO1163 | 7 | 8 | 9 | 12 | 8 | 2 | 9 | 10 | conserved hypothetical protein |
| SO1164 | 51 | 61 | 56 | 46 | 51 | 41 | 50 | 50 | D-alanyl-D-alanine carboxypeptidase (dacA-1) |
| SO1165 | 10 | 6 | 7 | 6 | 7 | 2 | 8 | 7 | rare lipoprotein A |
| SO1166 | 3 | 4 | 4 | 7 | 4 | 2 | 4 | 7 | membrane-bound lytic transglycosylase, putative |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1168 | 9 | 15 | 18 | 11 | 15 | 23 | 15 | 11 | penicillin-binding protein 2 (mrdA) |
| SO1169 | 3 | 3 | 5 | 2 | 4 | | 2 | | conserved hypothetical protein TIGR00246 |
| SO1170 | | 3 | 4 | 8 | 3 | 4 | 6 | 6 | iojap domain protein |
| SO1173 | | 2 | 4 | 5 | | | 2 | 2 | rare lipoprotein B |
| SO1174 | 32 | 29 | 26 | 24 | 24 | 14 | 21 | 24 | leucyl-tRNA synthetase (leuS) |
| SO1175 | 2 | | 2 | 3 | 2 | 2 | 3 | 2 | conserved hypothetical protein |
| SO1177 | | 4 | | 2 | 3 | 5 | | 4 | apolipoprotein N-acyltransferase (cutE) |
| SO1178 | 4 | 7 | 11 | 10 | 4 | 2 | 8 | 7 | magnesium and cobalt efflux protein CorC (corC) |
| SO1180 | 23 | 25 | 34 | 32 | 27 | 22 | 33 | 39 | PhoH family protein |
| SO1181 | 9 | 17 | 13 | 13 | 16 | 6 | 11 | 17 | tRNA-i(6)A37 modification enzyme MiaB (miaB) |
| SO1183 | 5 | 5 | 2 | 4 | 5 | 4 | | 5 | oxidoreductase, FAD-binding |
| SO1184 | 7 | 5 | 6 | 8 | 3 | 8 | 3 | 5 | peptidyl-tRNA hydrolase (pth) |
| SO1185 | 8 | 15 | 11 | 10 | 11 | 8 | 6 | 10 | conserved hypothetical protein TIGR00092 |
| SO1188 | | | 3 | 2 | | | | | conserved hypothetical protein |
| SO1189 | | | 4 | 3 | | | 3 | 4 | conserved hypothetical protein |
| SO1190 | 17 | 14 | 28 | 28 | 9 | 6 | 24 | 22 | conserved hypothetical protein |
| SO1191 | 8 | 13 | 10 | 12 | 12 | 12 | 6 | 11 | transcription elongation factor GreA (greA) |
| SO1193 | 5 | 7 | 3 | 4 | 3 | 2 | 7 | 4 | protein-export membrane protein SecD (secD-1) |
| SO1194 | | 2 | | | | | | | protein-export membrane protein SecF (secF-1) |
| SO1195 | | | | | 3 | | | 2 | conserved hypothetical protein TIGR00253 |
| SO1196 | 12 | 13 | 15 | 15 | 10 | 12 | 16 | 21 | ribosomal RNA large subunit methyltransferase J (rrmJ) |
| SO1197 | 54 | 67 | 63 | 71 | 66 | 42 | 72 | 84 | cell division protein FtsH (ftsH) |
| SO1198 | | 7 | 2 | 5 | 4 | 2 | 3 | 2 | dihydropteroate synthase (folP) |
| SO1199 | 8 | 10 | 9 | 10 | 14 | 6 | 14 | 17 | phosphoglucosamine mutase (glmM) |
| SO1200 | 22 | 20 | 17 | 16 | 17 | 22 | 18 | 15 | triosephosphate isomerase (tpiA) |
| SO1201 | | 3 | 4 | | | 3 | 5 | 7 | preprotein translocase, SecG subunit (secG) |
| SO1202 | 7 | 3 | 4 | 7 | 8 | 6 | 2 | 4 | conserved hypothetical protein |
| SO1203 | 35 | 45 | 35 | 43 | 38 | 43 | 33 | 32 | N utilization substance protein A (nusA) |
| SO1204 | 82 | 109 | 82 | 90 | 87 | 91 | 74 | 83 | translation initiation factor IF-2 (infB) |
| SO1205 | 12 | 8 | 5 | 4 | 7 | 7 | 4 | 3 | ribosome-binding factor A (rbfA) |
| SO1206 | 11 | 12 | 6 | 5 | 9 | 7 | 4 | 9 | tRNA pseudouridine synthase B (truB) |
| SO1207 | 12 | 12 | 13 | 15 | 14 | 10 | 14 | 17 | ribosomal protein S15 (rpsO) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1208 | 4 | 4 | 2 | 2 | 8 | 2 | 6 | 7 | GGDEF domain protein |
| SO1209 | 73 | 76 | 66 | 76 | 81 | 67 | 66 | 81 | polyribonucleotide nucleotidyltransferase (pnp) |
| SO1210 | 4 | 6 | 5 | 5 | 7 | 4 | 4 | 2 | TPR domain protein |
| SO1214 | 2 | 2 | | | 2 | | 2 | 2 | NupC family protein |
| SO1215 | 8 | 7 | 11 | 10 | 7 | 10 | 16 | 13 | outer membrane protein OmpK, putative |
| SO1217 | 9 | 14 | 11 | 10 | 8 | 9 | 13 | 12 | deoxyribose-phosphate aldolase (deoC) |
| SO1218 | 13 | 11 | 13 | 12 | 15 | 8 | 12 | 12 | thymidine phosphorylase (deoA) |
| SO1219 | 7 | 12 | 6 | 6 | 12 | 13 | 7 | 12 | phosphopentomutase (deoB) |
| SO1221 | 17 | 21 | 19 | 22 | 21 | 16 | 19 | 24 | purine nucleoside phosphorylase (deoD-2) |
| SO1222 | 4 | 5 | 3 | 5 | 5 | 2 | 11 | 11 | hypothetical protein |
| SO1223 | | | | 2 | | | 2 | | phosphoserine phosphatase (serB) |
| SO1224 | | | 2 | | | | | 2 | conserved hypothetical protein |
| SO1225 | | 2 | | | | | | | conserved hypothetical protein |
| SO1226 | 4 | 4 | | 2 | | 5 | 2 | 2 | DNA repair protein RadA (radA) |
| SO1230 | 2 | | 3 | | 3 | | 2 | 2 | sensor histidine kinase/response regulator TorS (torS) |
| SO1236 | | | | | 2 | | | | xanthine/uracil permease family protein |
| SO1242 | 2 | 2 | | | 2 | | | 3 | hypothetical protein |
| SO1250 | | | | | | 2 | | | conserved hypothetical protein |
| SO1252 | 9 | 19 | 18 | 8 | 19 | 7 | 5 | 17 | peptidase, U32 family |
| SO1254 | | | | | 2 | | | | conserved hypothetical protein |
| SO1255 | | | | | | | | 2 | cyclic nucleotide phosphodiesterase, putative |
| SO1258 | 7 | 8 | 5 | 3 | 6 | 4 | 2 | 3 | adenylosuccinate synthetase, putative |
| SO1259 | | | | | | | 2 | 2 | transcriptional regulator, LysR family |
| SO1262 | | | | | | 2 | | | hypothetical protein |
| SO1264 | 5 | 3 | | 3 | 8 | 2 | 2 | 2 | conserved hypothetical protein |
| SO1265 | 4 | 4 | 2 | 3 | 2 | 4 | | | transcriptional regulator, putative |
| SO1267 | 2 | 2 | | | | | | | conserved hypothetical protein |
| SO1268 | | 4 | 2 | | 4 | | 4 | 2 | glutamine synthetase |
| SO1269 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO1270 | 3 | 3 | 10 | 5 | 5 | 5 | | 5 | polyamine ABC transporter, periplasmic polyamine-binding protein |
| SO1271 | 9 | 9 | 9 | 8 | 8 | 5 | 6 | 6 | polyamine ABC transporter, ATP-binding protein |
| SO1274 | 3 | | 5 | 6 | 6 | 3 | 4 | 6 | conserved hypothetical protein |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
|---|---|---|---|---|---|---|---|---|---|
| SO1275 | 5 | 7 | 3 | 2 | 4 | 3 | 6 | 8 | succinate-semialdehyde dehydrogenase (gabD) |
| SO1276 | 8 | 13 | 10 | 6 | 9 | 7 | 8 | 6 | 4-aminobutyrate aminotransferase (gabT) |
| SO1278 | 4 | 5 | 5 | 7 | 7 | 5 | 8 | 9 | methyl-accepting chemotaxis protein |
| SO1284 | 22 | 25 | 25 | 31 | 23 | 17 | 18 | 36 | RNA polymerase sigma-70 factor (rpoD) |
| SO1286 | | 5 | | 4 | 2 | 4 | 3 | 7 | DNA primase (dnaG) |
| SO1287 | 4 | 4 | 6 | 4 | 3 | 6 | 2 | 4 | conserved hypothetical protein |
| SO1288 | 13 | 17 | 13 | 15 | 11 | 16 | 11 | 14 | ribosomal protein S21 (rpsU) |
| SO1289 | 4 | 6 | 5 | 5 | 4 | 3 | 2 | 8 | O-sialoglycoprotein endopeptidase (gcp) |
| SO1295 | 9 | 15 | 11 | 10 | 12 | 15 | 18 | 17 | major outer membrane lipoprotein, putative |
| SO1297 | 10 | 12 | 12 | 4 | 11 | 7 | 10 | 11 | general secretion pathway protein a (gspA) |
| SO1298 | 3 | 3 | 4 | 3 | 2 | 6 | 4 | 4 | general secretion pathway protein B (gspB) |
| SO1300 | 16 | 22 | 12 | 15 | 18 | 15 | 8 | 11 | glutamate-1-semialdehyde-2,1-aminomutase (hemL) |
| SO1301 | 5 | 14 | 12 | 11 | 13 | 4 | 10 | 15 | aspartate carbamoyltransferase (pyrB) |
| SO1303 | 9 | 10 | 10 | 7 | 12 | 5 | 11 | 12 | hypothetical protein |
| SO1304 | 6 | 6 | 7 | 6 | 5 | 5 | | 2 | HesB/YadR/YfhF family protein |
| SO1306 | | | | 2 | | | | 2 | conserved hypothetical protein |
| SO1309 | | | | 2 | | | | | conserved hypothetical protein |
| SO1313 | | 4 | 3 | | | | 2 | 3 | conserved hypothetical protein |
| SO1314 | | 5 | 2 | 2 | 2 | 2 | | | peptidase, M23/M37 family |
| SO1315 | 23 | 28 | 28 | 31 | 27 | 15 | 24 | 22 | tyrosyl-tRNA synthetase (tyrS) |
| SO1322 | 2 | 2 | 2 | | | | | | 5-methylthioadenosine nucleosidase/S-adenosylhomocysteine nucleosidase (pfs) |
| SO1324 | 4 | 3 | 2 | | 5 | 4 | | | glutamate synthase, small subunit (gltD) |
| SO1325 | | 4 | | 2 | 2 | 2 | 2 | | glutamate synthase, large subunit (gltB) |
| SO1327 | 4 | 5 | 4 | 6 | 7 | | 6 | 12 | sensor histidine kinase-related protein |
| SO1328 | 6 | 7 | 7 | 6 | 9 | 4 | 2 | 7 | transcriptional regulator, LysR family |
| SO1329 | 5 | 2 | 3 | 3 | 4 | | 2 | | adenylate cyclase-related protein |
| SO1330 | | | | | 2 | | 2 | | DNA mismatch repair protein MutH (mutH) |
| SO1332 | 4 | 6 | 4 | 5 | 7 | 3 | 4 | 2 | phosphoenolpyruvate-protein phosphotransferase PtsP (ptsP) |
| SO1334 | | | | | 2 | 2 | | 2 | prolipoprotein diacylglyceryl transferase (lgt) |
| SO1335 | 4 | | | 4 | 2 | 2 | | | thymidylate synthase (thyA) |
| SO1339 | 4 | 2 | 2 | 2 | 2 | 2 | | | conserved hypothetical protein |
| SO1341 | 8 | 8 | 7 | 7 | 8 | 4 | 9 | 9 | L-aspartate oxidase (nadB) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1342 | 3 | 4 | 4 | 5 | 4 | 7 | 8 | 6 | RNA polymerase sigma-24 factor (rpoE) |
| SO1343 | 2 | 4 | 3 | 3 | 3 | | | 4 | sigma-E factor negative regulatory protein (rseA) |
| SO1344 | 4 | 5 | 3 | 7 | 5 | 5 | 2 | 2 | sigma-E factor regulatory protein RseB (rseB) |
| SO1345 | | | 2 | 4 | | | | | sigma-E factor regulatory protein RseC (rseC) |
| SO1346 | 9 | 11 | 6 | 12 | 17 | 8 | 17 | 14 | GTP-binding protein LepA (lepA) |
| SO1347 | 21 | 29 | 25 | 19 | 22 | 23 | 23 | 19 | signal peptidase I (lepB) |
| SO1348 | 3 | 6 | | 2 | 5 | 4 | 4 | 4 | ribonuclease III (rnc) |
| SO1349 | 8 | 12 | 10 | 10 | 6 | 6 | 11 | 9 | GTP-binding protein Era (era) |
| SO1351 | 7 | 7 | 6 | 4 | 5 | 7 | 7 | 7 | pyridoxal phosphate biosynthetic protein PdxJ (pdxJ) |
| SO1353 | | 2 | | 3 | 2 | 3 | | | conserved hypothetical protein |
| SO1354 | 2 | 4 | 3 | 4 | 3 | 9 | 4 | 3 | hemolysin protein, putative |
| SO1356 | 36 | 52 | 39 | 46 | 38 | 39 | 38 | 42 | signal recognition particle protein Ffh (ffh) |
| SO1357 | 8 | 11 | 11 | 10 | 10 | 12 | 9 | 10 | ribosomal protein S16 (rpsP) |
| SO1358 | 5 | 5 | 2 | 3 | 3 | | | 3 | 16S rRNA processing protein RimM (rimM) |
| SO1359 | | 3 | 2 | | | | | | tRNA (guanine-N1)-methyltransferase (trmD) |
| SO1360 | 23 | 22 | 20 | 17 | 19 | 23 | 18 | 19 | ribosomal protein L19 (rplS) |
| SO1361 | | 3 | | 2 | | | | | phospho-2-dehydro-3-deoxyheptonate aldolase, tyr-sensitive (aroF) |
| SO1362 | | | | 2 | | | | 2 | chorismate mutase/prephenate dehydrogenase (tyrA) |
| SO1366 | | | | | | 2 | | | sodium/hydrogen exchanger family protein |
| SO1367 | 4 | 4 | 2 | 4 | 4 | 2 | | 3 | chorismate mutase/prephenate dehydratase (pheA) |
| SO1368 | 26 | 28 | 23 | 21 | 27 | 19 | 22 | 21 | cytosol aminopeptidase (pepA-2) |
| SO1369 | 6 | 7 | 4 | 5 | 3 | 4 | 6 | 6 | conserved hypothetical protein |
| SO1370 | 3 | 4 | | 5 | 4 | 2 | 2 | 2 | conserved hypothetical protein |
| SO1371 | 2 | | | 2 | | | | | conserved hypothetical protein |
| SO1372 | 2 | | 2 | | 3 | | 3 | 2 | conserved hypothetical protein |
| SO1375 | | 4 | 5 | 3 | 4 | 2 | 9 | 3 | carboxypeptidase |
| SO1377 | 35 | 46 | 40 | 41 | 46 | 50 | 41 | 54 | conserved hypothetical protein |
| SO1378 | | | | | | | 2 | 3 | ThiJ/PfpI family protein |
| SO1380 | | | | | | 2 | | | conserved hypothetical protein |
| SO1383 | 20 | 22 | 10 | 11 | 17 | 23 | 11 | 13 | ATP-dependent RNA helicase, DEAD box family |
| SO1385 | 5 | 8 | 3 | 9 | 4 | 6 | 6 | 8 | methyl-accepting chemotaxis protein |
| SO1388 | 15 | 24 | 13 | 12 | 15 | 14 | 11 | 10 | aminopeptidase P, putative |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1389 | 2 | | | | 2 | | | | ROK family protein |
| SO1390 | 12 | 11 | 10 | 6 | 11 | 12 | 8 | 7 | peptidyl-prolyl cis-trans isomerase, FKBP-type |
| SO1401 | 13 | 12 | 9 | 8 | 14 | 13 | 11 | 15 | hypothetical protein |
| SO1402 | 6 | 3 | 2 | | 3 | 4 | | 2 | conserved hypothetical protein |
| SO1403 | | 2 | 2 | | | | | | conserved hypothetical protein |
| SO1404 | 4 | 7 | 4 | 4 | 4 | 3 | 2 | 3 | endoribonuclease L-PSP, putative |
| SO1405 | 3 | 4 | 4 | 2 | 10 | 7 | 2 | | transglutaminase family protein |
| SO1407 | | | | | 2 | | | | mercuric transport periplasmic protein MerP, putative |
| SO1408 | 2 | 5 | 2 | 2 | 2 | | | | helicase, putative |
| SO1410 | 2 | 2 | | 4 | 3 | 3 | | | hypothetical protein |
| SO1412 | | | 3 | | | | | 2 | conserved hypothetical protein |
| SO1417 | 2 | | 2 | | 3 | | | | sensor histidine kinase |
| SO1424 | 18 | 27 | 7 | 20 | 21 | 26 | 16 | 23 | hypothetical protein |
| SO1425 | 13 | 18 | 13 | 20 | 19 | 14 | 17 | 12 | hypothetical protein |
| SO1429 | 2 | 3 | | 4 | 24 | 23 | | | anaerobic dimethyl sulfoxide reductase, A subunit (dmaA-1) |
| SO1430 | 2 | | | | 8 | 7 | | | anaerobic dimethyl sulfoxide reductase, B subunit (dmsB-1) |
| SO1434 | 3 | 4 | 3 | 3 | 4 | 3 | 3 | 4 | methyl-accepting chemotaxis protein |
| SO1438 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO1441 | | 2 | 4 | 3 | 3 | | 4 | 4 | hypothetical protein |
| SO1457 | 13 | 15 | 15 | 14 | 15 | 8 | 17 | 18 | type I restriction-modification system, M subunit, putative |
| SO1458 | | 2 | | | | | | | hypothetical protein |
| SO1460 | 17 | 16 | 14 | 10 | 16 | 12 | 14 | 12 | type I restriction-modification system, S subunit, putative |
| SO1461 | 10 | 8 | 11 | 11 | 14 | 11 | 12 | 13 | protein kinase, putative |
| SO1462 | 6 | 2 | 4 | | 5 | 4 | 2 | 2 | hypothetical protein |
| SO1463 | 2 | 3 | 4 | | | 2 | | | hypothetical protein |
| SO1464 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO1465 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO1468 | | | 2 | | | 2 | | | conserved hypothetical protein |
| SO1471 | 2 | 3 | 4 | 4 | 7 | 2 | 5 | 6 | site-specific recombinase, phage integrase family |
| SO1473 | 8 | 8 | 9 | 11 | 10 | 7 | 4 | 6 | SsrA-binding protein (smpB) |
| SO1474 | 2 | | | 2 | | | | | conserved hypothetical protein |
| SO1475 | 7 | 10 | 5 | 4 | 8 | 8 | 4 | 4 | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1476 | 6 | 11 | 7 | 8 | 8 | 9 | 8 | 8 | small protein A (smpA) |
| SO1481 | 2 | 3 | 2 | 2 | 5 | 4 | 5 | 6 | glutathione-regulated potassium-efflux system protein |
| SO1482 | 34 | 36 | 66 | 61 | 14 | 16 | 87 | 88 | TonB-dependent receptor, putative |
| SO1483 | | | | 5 | | | 2 | 5 | malate synthase A (aceB) |
| SO1484 | 5 | 6 | 12 | 12 | 12 | 3 | 9 | 16 | isocitrate lyase (aceA) |
| SO1487 | 3 | | | | 3 | 4 | 2 | 3 | hypothetical protein |
| SO1489 | 3 | 5 | 4 | 7 | 10 | 5 | 4 | 4 | hypothetical protein |
| SO1490 | 16 | 16 | 9 | 10 | 23 | 27 | 5 | 11 | alcohol dehydrogenase II (adhB) |
| SO1496 | | | 2 | | | | | | glycogen phosphorylase family protein |
| SO1500 | 3 | | | 3 | 3 | | 4 | | sensory box protein |
| SO1501 | 2 | 2 | | | | | 2 | 3 | ATP-dependent RNA helicase, DEAD box family |
| SO1502 | 4 | 3 | 2 | 3 | 2 | 3 | 3 | 3 | cobalamin synthesis protein/P47K family protein |
| SO1507 | 8 | 6 | 11 | 11 | 5 | 2 | 4 | 7 | hypothetical protein |
| SO1511 | | | | | | 2 | | | transposase, IS91 family, putative |
| SO1512 | 2 | | | 2 | | 2 | | | hypothetical protein |
| SO1514 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO1515 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO1518 | 12 | 14 | 13 | 13 | 16 | 15 | 7 | 8 | conserved hypothetical protein |
| SO1519 | 35 | 39 | 23 | 27 | 36 | 36 | 21 | 27 | iron-sulfur cluster-binding protein |
| SO1520 | 8 | 7 | 6 | 9 | 13 | 5 | 4 | 9 | conserved hypothetical protein |
| SO1521 | 89 | 95 | 80 | 78 | 99 | 92 | 71 | 70 | iron-sulfur cluster-binding protein |
| SO1522 | | | | | 2 | | | | L-lactate permease, putative |
| SO1523 | 2 | | | | | | | | conserved hypothetical protein |
| SO1524 | 22 | 21 | 25 | 26 | 18 | 25 | 21 | 26 | heat shock protein GrpE (grpE) |
| SO1525 | 8 | 10 | 10 | 11 | 13 | 2 | 10 | 6 | deoxyxylulose-5-phosphate synthase (dxs) |
| SO1526 | | 2 | | | | | | | geranyltranstransferase (ispA) |
| SO1529 | 6 | 6 | 6 | 2 | 3 | 2 | | 4 | chemotaxis motA protein (pomA) |
| SO1530 | 5 | 3 | 2 | 6 | 3 | 5 | 3 | 7 | sodium-driven polar flagellar protein PomB (pomB) |
| SO1531 | 14 | 10 | 9 | 9 | 16 | 11 | 10 | 10 | thiamine biosynthesis protein ThiI (thiI) |
| SO1533 | 2 | | 2 | | 2 | | | | glycine cleavage system transcriptional activator, putative |
| SO1536 | 9 | 13 | 11 | 6 | 10 | 7 | 7 | 12 | conserved hypothetical protein |
| SO1538 | 13 | 13 | 15 | 17 | 13 | 7 | 17 | 19 | isocitrate dehydrogenase, NAD-dependent |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1539 | 11 | 21 | 15 | 22 | 19 | 15 | 21 | 22 | conserved hypothetical protein |
| SO1548 | 4 | 3 | | 4 | 5 | 5 | 3 | 7 | hypothetical protein |
| SO1549 | 4 | 3 | 5 | 6 | | | 4 | 5 | exodeoxyribonuclease IX (xni) |
| SO1550 | 5 | 10 | 11 | 5 | 7 | 5 | 8 | 5 | conserved hypothetical protein |
| SO1551 | 12 | 10 | 9 | 12 | 17 | 12 | 15 | 7 | GGDEF domain protein |
| SO1552 | 7 | 9 | 5 | 6 | 8 | 9 | 6 | 7 | TPR domain protein |
| SO1556 | 19 | 29 | 18 | 18 | 21 | 9 | 13 | 23 | conserved hypothetical protein |
| SO1559 | 4 | 7 | 5 | 5 | 7 | 6 | 4 | 5 | phosphate regulon sensor protein PhoR (phoR) |
| SO1560 | 2 | | | 2 | | | | | phosphate-binding protein |
| SO1563 | | | | 2 | | | 2 | | glutathione peroxidase, putative |
| SO1565 | | 2 | 3 | | 4 | | 2 | 2 | magnesium transporter, putative |
| SO1568 | 7 | 11 | 10 | 9 | 5 | 6 | 11 | 10 | hypothetical protein |
| SO1571 | 3 | 10 | 4 | 5 | 10 | 4 | 11 | 12 | hypothetical protein |
| SO1575 | 5 | | | 3 | 2 | | | 2 | NOL1/NOP2/sun family putative RNA methylase |
| SO1576 | 2 | 5 | 7 | 6 | 3 | | 6 | 11 | glutathione S-transferase family protein |
| SO1579 | | | | | 2 | | | 3 | conserved hypothetical protein |
| SO1580 | 11 | 13 | 22 | 21 | 3 | 7 | 19 | 26 | TonB-dependent heme receptor |
| SO1581 | 7 | 6 | 6 | 5 | 4 | | 3 | 4 | phnA protein (phnA) |
| SO1582 | 2 | 3 | | 2 | 2 | | | | transcriptional regulator, MarR family |
| SO1589 | 4 | 3 | 3 | 3 | 5 | 2 | 2 | 4 | hypothetical protein |
| SO1595 | | | | | | | 2 | | hypothetical protein |
| SO1597 | 3 | 3 | 4 | 6 | 3 | | 3 | 3 | conserved hypothetical protein |
| SO1599 | 18 | 14 | 24 | 19 | 18 | 5 | 12 | 17 | beta-ketoacyl synthase |
| SO1602 | 37 | 35 | 44 | 49 | 35 | 33 | 29 | 36 | multi-domain beta-ketoacyl synthase |
| SO1603 | | | 2 | | | | | | transcriptional regulator, putative |
| SO1605 | 2 | | | 2 | | | 3 | 2 | lipoprotein, putative |
| SO1606 | 2 | | 3 | 3 | 2 | | | 6 | metallo-beta-lactamase superfamily protein |
| SO1608 | 3 | 7 | 4 | 4 | 5 | 2 | 3 | 4 | conserved hypothetical protein |
| SO1610 | 4 | 6 | 4 | 3 | 2 | 2 | 5 | 4 | conserved hypothetical protein |
| SO1611 | 3 | 6 | 2 | 4 | 2 | 3 | 4 | 4 | hypothetical protein |
| SO1616 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO1617 | 3 | | | | | | | | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1620 | | | | | 2 | | 2 | | RNA pseudouridylate synthase family protein |
| SO1624 | | | | | 2 | | | | formyltetrahydrofolate deformylase (purU) |
| SO1625 | 10 | 9 | 15 | 18 | 10 | 9 | 18 | 15 | 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (dapD) |
| SO1626 | 7 | 4 | 4 | 5 | 7 | 5 | 6 | 4 | protein-P-II uridylyltransferase (glnD) |
| SO1627 | 15 | 17 | 13 | 17 | 11 | 12 | 10 | 12 | methionine aminopeptidase, type I (map) |
| SO1629 | 36 | 43 | 39 | 42 | 39 | 41 | 44 | 41 | ribosomal protein S2 (rpsB) |
| SO1630 | 44 | 43 | 40 | 49 | 37 | 47 | 29 | 34 | translation elongation factor Ts (tsf) |
| SO1631 | 12 | 17 | 15 | 17 | 15 | 9 | 8 | 10 | uridylate kinase (pyrH) |
| SO1632 | 22 | 18 | 17 | 24 | 14 | 19 | 11 | 19 | ribosome recycling factor (frr) |
| SO1633 | 2 | | | | | 2 | 2 | 2 | undecaprenyl diphosphate synthase (uppS) |
| SO1635 | 2 | 2 | 2 | 2 | 3 | | | 3 | 1-deoxy-D-xylulose 5-phosphate reductoisomerase (dxr) |
| SO1636 | 2 | 6 | 6 | 3 | 5 | 2 | 4 | 5 | membrane-associated zinc metalloprotease, putative |
| SO1637 | 46 | 55 | 47 | 55 | 38 | 42 | 46 | 46 | bacterial surface antigen |
| SO1638 | 22 | 21 | 25 | 21 | 17 | 25 | 19 | 22 | outer membrane protein OmpH (ompH) |
| SO1639 | 5 | 6 | 6 | 7 | 6 | 10 | 7 | 7 | UDP-3-O-(3-hydroxymyristoyl) glucosamine n-acyltransferase (lpxD) |
| SO1641 | 8 | 10 | 5 | 7 | 8 | 5 | 4 | 5 | acyl-(acyl-carrier-protein)--UDP-N-acetylglucosamine O-acyltransferase (lpxA) |
| SO1642 | 6 | 7 | 4 | 8 | 8 | 3 | 13 | 10 | lipid A disaccharide synthase (lpxB) |
| SO1643 | 2 | 2 | 3 | 4 | 2 | 2 | 2 | 2 | ribonuclease HII (rnhB) |
| SO1644 | 2 | 3 | 4 | 6 | 6 | 4 | 4 | 9 | DNA polymerase III, alpha subunit (dnaE) |
| SO1648 | 5 | 7 | 4 | 5 | 5 | 5 | 4 | 3 | cold shock domain family protein |
| SO1651 | 2 | 5 | 5 | 2 | 3 | | 4 | 3 | Snf2 family protein |
| SO1652 | 2 | | 2 | 2 | 4 | | 3 | 2 | conserved hypothetical protein |
| SO1655 | 2 | 2 | 3 | 4 | | | 4 | 4 | cysQ protein (cysQ-2) |
| SO1656 | | 3 | 3 | 5 | 2 | 2 | 3 | 8 | ROK family protein |
| SO1657 | 4 | 3 | 3 | 2 | 3 | 4 | 3 | 2 | conserved hypothetical protein |
| SO1659 | | 3 | 2 | | | | | | decaheme cytochrome c |
| SO1662 | 2 | 3 | 3 | 2 | 2 | 3 | 3 | 4 | conserved hypothetical protein |
| SO1664 | 5 | 8 | 7 | 6 | 4 | 2 | 8 | 7 | UDP-glucose 4-epimerase (galE) |
| SO1665 | 24 | 22 | 21 | 22 | 18 | 22 | 18 | 25 | UTP-glucose-1-phosphate uridylyltransferase (galU) |
| SO1669 | 12 | 12 | 9 | 10 | 15 | 9 | 9 | 11 | transcriptional regulatory protein TyrR (tyrR) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1670 | 9 | 9 | 6 | 7 | 8 | 7 | 10 | 10 | fumarylacetoacetate hydrolase family protein |
| SO1674 | 6 | 13 | 4 | 5 | 8 | 5 | 4 | 8 | oxidoreductase, short chain dehydrogenase/reductase family |
| SO1675 | 14 | 15 | 12 | 12 | 13 | 9 | 9 | 11 | conserved hypothetical protein |
| SO1676 | 2 | 3 | 2 | 2 | | | 3 | 4 | homoserine O-succinyltransferase (metA) |
| SO1677 | 29 | 38 | 36 | 38 | 32 | 31 | 28 | 33 | acetyl-CoA acetyltransferase (atoB) |
| SO1678 | 8 | 10 | 11 | 11 | 17 | 10 | 11 | 16 | methylmalonate-semialdehyde dehydrogenase (mmsA) |
| SO1679 | 19 | 18 | 24 | 22 | 31 | 23 | 19 | 21 | acyl-CoA dehydrogenase family protein |
| SO1680 | 10 | 17 | 14 | 15 | 17 | 19 | 15 | 16 | enoyl-CoA hydratase/isomerase family protein |
| SO1681 | | 3 | 3 | 5 | 4 | | 2 | 2 | enoyl-CoA hydratase/isomerase family protein |
| SO1682 | 4 | 6 | 3 | 6 | 4 | 2 | 2 | 3 | 3-hydroxyisobutyrate dehydrogenase (mmsB) |
| SO1683 | 11 | 11 | 9 | 7 | 13 | 21 | 10 | 13 | 3-oxoacyl-(acyl-carrier-protein) reductase, putative |
| SO1686 | 3 | 4 | 3 | | 2 | 6 | 3 | 4 | prolyl oligopeptidase family protein |
| SO1689 | 4 | 2 | 2 | 8 | 3 | 2 | 4 | 3 | cation transport ATPase, E1-E2 family |
| SO1690 | 17 | 18 | 21 | 15 | 16 | 12 | 23 | 16 | ABC transporter, ATP-binding protein |
| SO1691 | 3 | 3 | 3 | 4 | 2 | 2 | 3 | 5 | lipoprotein Blc (blc) |
| SO1700 | 5 | 9 | 12 | 15 | 4 | 5 | 9 | 8 | hypothetical protein |
| SO1701 | | | 5 | | 3 | 2 | 3 | 2 | hypothetical protein |
| SO1717 | 3 | 5 | 6 | 7 | 6 | | 7 | 9 | hypothetical protein |
| SO1718 | 2 | | | 2 | 2 | 2 | | 5 | conserved hypothetical protein |
| SO1723 | 4 | 5 | 3 | 3 | 2 | 4 | 2 | 2 | phosphate ABC transporter, permease protein, putative |
| SO1724 | 6 | 6 | 8 | 3 | 4 | 4 | 6 | 3 | phosphate ABC transporter, permease protein, putative |
| SO1725 | 10 | 12 | 16 | 19 | 12 | 7 | 11 | 18 | phosphate ABC transporter, ATP-binding protein (pstB-1) |
| SO1726 | 3 | 4 | | 4 | 3 | 2 | 2 | 3 | phosphate transport system regulatory protein PhoU (phoU) |
| SO1732 | | 3 | | 3 | 3 | | | 2 | cold shock domain family protein |
| SO1734 | | | 2 | | | | | 3 | glyoxalase family protein |
| SO1738 | | | | | | | 4 | | conserved hypothetical protein |
| SO1742 | | | | 4 | 2 | | | 3 | conserved hypothetical protein |
| SO1743 | 6 | 9 | 3 | 4 | 5 | 6 | 4 | 5 | hydrolase, alpha/beta hydrolase fold family |
| SO1744 | 4 | 2 | | 2 | 4 | 4 | 2 | 3 | AMP-binding protein |
| SO1745 | 7 | 6 | 5 | 4 | 6 | 5 | 3 | 5 | 3-beta hydroxysteroid dehydrogenase/isomerase family protein |
| SO1750 | 22 | 25 | 20 | 20 | 25 | 20 | 20 | 23 | ABC transporter, ATP-binding protein |
| SO1751 | 12 | 12 | 9 | 9 | 19 | 15 | 11 | 13 | membrane protein, putative |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO1755 | 7 | 16 | 12 | 14 | 6 | 4 | 16 | 21 | phosphoglucomutase/phosphomannomutase family protein |
| SO1756 | | | | 2 | | | | 3 | glyoxalase family protein |
| SO1767 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO1769 | | 2 | | | | | 3 | 2 | glutamate decarboxylase, putative |
| SO1770 | | 2 | | | | | | | glycerate kinase, putative |
| SO1776 | 13 | 20 | 14 | 9 | 32 | 26 | 12 | 18 | outer membrane protein precursor MtrB (mtrB) |
| SO1777 | | 2 | | 2 | 4 | 3 | | | decaheme cytochrome c MtrA (mtrA) |
| SO1778 | 30 | 34 | 29 | 29 | 42 | 44 | 14 | 12 | decaheme cytochrome c (omcB) |
| SO1779 | 35 | 39 | 21 | 29 | 56 | 55 | 18 | 18 | decaheme cytochrome c (omcA) |
| SO1783 | 7 | 7 | 8 | 8 | 2 | 3 | 2 | 8 | ferrous iron transport protein A (feoA) |
| SO1784 | 20 | 19 | 21 | 20 | 14 | 13 | 23 | 26 | ferrous iron transport protein B (feoB) |
| SO1786 | 11 | 18 | 11 | 17 | 21 | 6 | 21 | 13 | glutaminyl-tRNA synthetase (glnS) |
| SO1788 | | | | | | | 4 | | tRNA-(MS[2]IO[6]A)-hydroxylase (miaE) |
| SO1790 | 12 | 12 | 11 | 7 | 9 | 11 | 8 | 11 | peptidyl-prolyl cis-trans isomerase B (ppiB-1) |
| SO1791 | 7 | 13 | 6 | 12 | 9 | 4 | 13 | 9 | cysteinyl-tRNA synthetase (cysS) |
| SO1792 | 4 | 2 | 4 | 4 | 4 | 4 | 5 | 5 | methylenetetrahydrofolate dehydrogenase/methylenetetrahydrofolate cyclohydrolase (folD) |
| SO1793 | 69 | 76 | 78 | 77 | 69 | 81 | 70 | 70 | trigger factor (tig) |
| SO1794 | 5 | 9 | 7 | 6 | 6 | 2 | 6 | 9 | ATP-dependent Clp protease, proteolytic subunit (clpP) |
| SO1795 | 19 | 20 | 24 | 25 | 13 | 14 | 18 | 22 | ATP-dependent Clp protease, ATP-binding subunit ClpX (clpX) |
| SO1796 | 51 | 65 | 60 | 71 | 56 | 48 | 69 | 74 | ATP-dependent protease La (lon) |
| SO1797 | 22 | 23 | 22 | 23 | 23 | 31 | 23 | 21 | DNA-binding protein, HU family |
| SO1798 | 88 | 112 | 100 | 95 | 77 | 99 | 85 | 100 | peptidyl-prolyl cis-trans isomerase D (ppiD) |
| SO1800 | 18 | 29 | 26 | 23 | 19 | 12 | 19 | 22 | conserved hypothetical protein |
| SO1801 | 9 | 10 | 8 | 8 | 10 | 8 | 8 | 5 | peptide ABC transporter, ATP-binding protein (sapF) |
| SO1802 | 7 | 9 | 7 | | 10 | 7 | 6 | 6 | peptide ABC transporter, ATP-binding protein (sapD) |
| SO1805 | 8 | 10 | 7 | 5 | 4 | 3 | 6 | 8 | peptide ABC transporter, periplasmic peptide-binding protein (sapA) |
| SO1806 | 2 | | | | | | 2 | 2 | psp operon transcriptional activator (pspF) |
| SO1807 | 36 | 39 | 33 | 40 | 39 | 33 | 41 | 40 | phage shock protein A (pspA) |
| SO1808 | 2 | 2 | 2 | 3 | 3 | 2 | 2 | 2 | phage shock protein B (pspB) |
| SO1809 | | 4 | 3 | 5 | 4 | 5 | 4 | 2 | phage shock protein C (pspC) |
| SO1810 | | | 2 | | 2 | 3 | 2 | | conserved hypothetical protein |
| SO1811 | 3 | 4 | 4 | 4 | 6 | 3 | 8 | 11 | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1812 | | | 2 | 2 | | | | 2 | methionine gamma-lyase (mdeA) |
| SO1816 | 8 | 10 | 8 | 10 | 7 | 5 | 5 | 9 | conserved hypothetical protein |
| SO1817 | | 4 | 3 | 3 | 3 | | | 3 | primosomal replication protein n, putative |
| SO1819 | 4 | 3 | | 4 | 3 | | 3 | 2 | ATP-dependent helicase DinG (dinG) |
| SO1821 | 10 | 14 | 16 | 19 | 14 | 18 | 22 | 22 | outer membrane porin, putative |
| SO1824 | 49 | 61 | 52 | 51 | 51 | 51 | 52 | 52 | conserved hypothetical protein |
| SO1825 | 92 | 119 | 87 | 93 | 97 | 120 | 97 | 112 | MotA/TolQ/ExbB proton channel family protein |
| SO1826 | 5 | 6 | 4 | 4 | 6 | 6 | 2 | 4 | TonB system transport protein ExbB2 (exbB2) |
| SO1827 | 12 | 10 | 13 | 11 | 10 | 8 | 11 | 11 | TonB system transport protein ExbD2 (exbD2) |
| SO1828 | 14 | 18 | 13 | 17 | 17 | 12 | 15 | 15 | TonB2 protein (tonB2) |
| SO1829 | 74 | 85 | 73 | 66 | 77 | 64 | 70 | 72 | TPR domain protein |
| SO1831 | 34 | 44 | 30 | 28 | 38 | 21 | 33 | 31 | conserved hypothetical protein |
| SO1832 | 2 | 5 | 7 | 4 | 4 | 3 | 5 | 8 | peptidyl-prolyl cis-trans isomerase C (ppiC-2) |
| SO1833 | | | 2 | 3 | | | | 3 | ISSo7, transposase |
| SO1839 | | 2 | | | 2 | | | | hypothetical protein |
| SO1844 | | 4 | | 4 | 4 | 4 | 2 | 3 | extracellular nuclease, putative |
| SO1846 | | | | | | | | 3 | hypothetical protein |
| SO1848 | 6 | 5 | 6 | 5 | 4 | 3 | 4 | 5 | hypothetical protein |
| SO1849 | 7 | 8 | 4 | 3 | 6 | 2 | 2 | 5 | hypothetical protein |
| SO1850 | 2 | 3 | 2 | 3 | 2 | 2 | 3 | 4 | DnaJ domain protein |
| SO1851 | 7 | 5 | 6 | 6 | 8 | 2 | 3 | 6 | conserved hypothetical protein |
| SO1853 | 10 | 15 | 6 | 9 | 6 | 10 | 7 | 13 | ABC transporter, ATP-binding protein |
| SO1854 | 27 | 37 | 29 | 29 | 29 | 29 | 32 | 36 | hypothetical protein |
| SO1855 | | | 2 | | | | 3 | 2 | ribosome modulation factor (rmf) |
| SO1856 | 9 | 5 | 3 | 5 | 4 | 2 | 4 | 3 | 3-hydroxydecanoyl-(acyl-carrier-protein) dehydratase (fabA) |
| SO1857 | 7 | 9 | 3 | 7 | 6 | 6 | 4 | 9 | conserved hypothetical protein |
| SO1860 | 5 | 3 | 3 | 6 | 4 | 2 | 7 | 7 | DNA-binding response regulator, LuxR family |
| SO1861 | 2 | 2 | 3 | 2 | 4 | 2 | | 4 | excinuclease ABC, C subunit (uvrC) |
| SO1863 | 4 | 6 | 3 | 2 | 3 | 2 | | 2 | DNA-binding protein, HU family |
| SO1865 | 8 | 3 | 5 | 8 | 5 | 4 | 6 | 4 | ABC transporter, ATP-binding protein |
| SO1867 | 7 | 6 | 5 | 5 | 6 | 3 | 2 | 7 | conserved hypothetical protein |
| SO1868 | 3 | 7 | | 4 | 4 | | 4 | 3 | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO1870 | 19 | 18 | 15 | 15 | 19 | 9 | 17 | 21 | biosynthetic arginine decarboxylase (speA) |
| SO1875 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO1877 | 6 | 8 | 3 | 7 | 9 | 3 | | 3 | bacterioferritin comigratory protein (bcp) |
| SO1878 | 3 | | 2 | 3 | 3 | | 2 | 2 | glycine cleavage system transcriptional repressor, putative |
| SO1879 | 13 | 17 | 11 | 13 | 11 | 10 | 11 | 11 | dihydrodipicolinate synthase (dapA) |
| SO1880 | 18 | 36 | 26 | 28 | 24 | 29 | 25 | 30 | lipoprotein-34 NlpB (nlpB) |
| SO1881 | 7 | 11 | 6 | 6 | 3 | 4 | 2 | 6 | HlyD family-related protein |
| SO1882 | 5 | 2 | 4 | 3 | 4 | 2 | 2 | 2 | AcrB/AcrD/AcrF family protein |
| SO1891 | 5 | 11 | 10 | 9 | 8 | 6 | 8 | 14 | 3-oxoadipate CoA-succinyl transferase, beta subunit |
| SO1892 | 9 | 6 | 6 | 6 | 9 | 10 | 6 | 8 | acetate CoA-transferase, subunit A (atoD) |
| SO1893 | 11 | 6 | 5 | 5 | 11 | 10 | 3 | 8 | hydroxymethylglutaryl-CoA lyase (mvaB) |
| SO1894 | 34 | 39 | 34 | 33 | 49 | 32 | 32 | 33 | acetyl-CoA carboxylase, biotin carboxylase, putative |
| SO1895 | 7 | 3 | 3 | 5 | 8 | 5 | 6 | 7 | enoyl-CoA hydratase/isomerase family protein |
| SO1896 | 23 | 24 | 14 | 21 | 28 | 35 | 19 | 25 | 3-methylcrotonyl CoA carboxylase, beta subunit (pccB-1) |
| SO1897 | 20 | 19 | 19 | 27 | 28 | 26 | 20 | 24 | isovaleryl-CoA dehydrogenase (ivd) |
| SO1898 | 7 | 9 | 11 | 8 | 7 | 7 | 9 | 11 | transcriptional regulator, putative |
| SO1899 | | | 2 | 3 | | | | 2 | conserved hypothetical protein |
| SO1902 | 6 | 7 | 4 | 3 | 8 | 6 | 3 | 2 | 6-phosphogluconate dehydrogenase, decarboxylating (gnd) |
| SO1904 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO1905 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO1912 | 2 | 3 | 2 | 4 | 3 | 2 | 3 | 3 | acyl-CoA thioesterase II (tesB) |
| SO1913 | 3 | 4 | 3 | 4 | | 5 | 4 | 3 | conserved hypothetical protein |
| SO1915 | 3 | 2 | | | | 2 | | | serine protease, subtilase family |
| SO1916 | 2 | 2 | 2 | | 2 | | | | transcriptional regulator, LysR family |
| SO1921 | 3 | 5 | 2 | 2 | 3 | 3 | 4 | 2 | hypothetical protein |
| SO1923 | | 3 | | | 2 | | | | AcrB/AcrD/AcrF family protein |
| SO1924 | | 2 | | | | | 3 | | AcrB/AcrD/AcrF family protein |
| SO1925 | | | 2 | | | | | | HlyD family secretion protein |
| SO1926 | 34 | 33 | 29 | 34 | 42 | 19 | 40 | 41 | citrate synthase (gltA) |
| SO1928 | 67 | 69 | 66 | 73 | 63 | 56 | 55 | 59 | succinate dehydrogenase, flavoprotein subunit (sdhA) |
| SO1929 | 32 | 32 | 29 | 27 | 31 | 27 | 28 | 35 | succinate dehydrogenase, iron-sulfur protein (sdhB) |
| SO1930 | 63 | 76 | 68 | 67 | 75 | 61 | 92 | 91 | 2-oxoglutarate dehydrogenase, E1 component (sucA) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO1931 | 45 | 48 | 45 | 38 | 42 | 44 | 43 | 39 | 2-oxoglutarate dehydrogenase, E2 component, dihydrolipoamide succinyltransferase (sucB) |
| SO1932 | 41 | 49 | 38 | 46 | 34 | 43 | 27 | 40 | succinyl-CoA synthase, beta subunit (sucC) |
| SO1933 | 17 | 25 | 17 | 16 | 19 | 18 | 14 | 18 | succinyl-CoA synthase, alpha subunit (sucD) |
| SO1935 | 3 | 2 | 3 | 2 | 3 | | 3 | 2 | regulator of nucleoside diphosphate kinase (rnk) |
| SO1936 | | | | 2 | | | 2 | | acetyltransferase, GNAT family |
| SO1937 | 6 | 13 | 12 | 12 | 9 | 8 | 9 | 12 | ferric uptake regulation protein (fur) |
| SO1940 | | | | | 2 | | | | hypothetical protein |
| SO1941 | | 2 | 2 | 2 | 2 | 2 | 3 | | magnesium and cobalt transport protein CorA (corA) |
| SO1942 | | | | 2 | | | | | HDIG domain protein |
| SO1945 | 5 | 10 | 5 | 6 | 7 | 7 | 7 | 8 | sensor protein PhoQ (phoQ) |
| SO1946 | | 2 | | 3 | 2 | | 4 | 3 | transcriptional regulatory protein PhoP (phoP) |
| SO1948 | 3 | 4 | 3 | 2 | 2 | 4 | | | sodium:dicarboxylate symporter family protein |
| SO1949 | 18 | 29 | 13 | 16 | 26 | 32 | 10 | 18 | invasin domain protein |
| SO1952 | 3 | 6 | 5 | 10 | 2 | | 5 | 2 | gamma-glutamyltranspeptidase (ggt-2) |
| SO1959 | | | 2 | | | | | | ABC transporter, periplasmic substrate-binding protein, putative |
| SO1961 | | | | | | | | 3 | maltose O-acetyltransferase (maa) |
| SO1962 | | | 2 | 2 | 4 | | | 2 | 4-hydroxyphenylpyruvate dioxygenase |
| SO1966 | | 2 | 6 | 2 | 3 | 4 | 4 | 2 | conserved hypothetical protein TIGR00266 |
| SO1968 | | | | | 2 | | | | hypothetical protein |
| SO1970 | | | | 2 | | | | | hypothetical protein |
| SO1975 | | | | | | | | 2 | Zinc carboxypeptidase-related protein |
| SO1980 | | | | | | | | 2 | phosphoribosyl transferase domain protein |
| SO1981 | 2 | | | 5 | 2 | 2 | 2 | | conserved hypothetical protein |
| SO1986 | | 2 | | | | | 2 | | RNA polymerase sigma-70 factor, ECF subfamily |
| SO1988 | | | | | | | | 2 | conserved hypothetical protein |
| SO1989 | | | | 2 | 2 | | 2 | 2 | chemotaxis protein CheV (cheV-1) |
| SO1994 | 16 | 17 | 15 | 16 | 17 | 17 | 18 | 18 | membrane-bound lytic transglycolase-related protein |
| SO1995 | 7 | 8 | 4 | 4 | 9 | 6 | 6 | 7 | peptidyl-prolyl cis-trans isomerase, FkbP family |
| SO2001 | 45 | 47 | 48 | 44 | 55 | 60 | 39 | 44 | 5-nucleotidase (ushA) |
| SO2006 | 2 | 2 | | 2 | | | | 2 | NifR3/Smm1 family protein |
| SO2007 | | | | | | | | 2 | conserved hypothetical protein |
| SO2008 | | 4 | | | | | | | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO2012 | 2 | 2 | 2 | 2 | | 2 | | 2 | adenine phosphoribosyltransferase (apt) |
| SO2013 | 9 | 8 | 2 | 6 | 5 | 5 | 4 | 7 | DNA polymerase III, gamma and tau subunits (dnaX) |
| SO2014 | 8 | 9 | 12 | 11 | 9 | 11 | 6 | 8 | conserved hypothetical protein TIGR00103 |
| SO2015 | | 2 | | 3 | | | | | recombination protein RecR (recR) |
| SO2016 | 63 | 75 | 92 | 100 | 82 | 72 | 113 | 124 | heat shock protein HtpG (htpG) |
| SO2017 | | | 4 | 5 | 3 | | 6 | 8 | conserved hypothetical protein |
| SO2018 | 24 | 32 | 27 | 26 | 25 | 19 | 18 | 22 | adenylate kinase (adk) |
| SO2019 | 4 | 5 | 2 | 2 | 6 | 7 | 3 | 3 | ferrochelatase (hemH-1) |
| SO2020 | 4 | 6 | 2 | 2 | 7 | 2 | 4 | 2 | inosine-guanosine kinase (gsk) |
| SO2021 | 6 | 7 | 5 | 9 | 3 | 4 | 3 | 7 | NH(3)-dependent NAD(+) synthetase (nadE) |
| SO2025 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO2026 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO2027 | 3 | 4 | 3 | 2 | 2 | 2 | 5 | 2 | hypothetical protein |
| SO2031 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO2032 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO2034 | | | 2 | | | | 4 | | hypothetical protein |
| SO2035 | | | | | | 2 | | | transposase, putative |
| SO2037 | | 2 | | | | | | | site-specific recombinase, phage integrase family |
| SO2040 | 6 | 6 | 5 | 13 | 8 | 5 | 10 | 9 | soluble lytic murein transglycosylase, putative |
| SO2041 | 14 | 24 | 19 | 25 | 21 | 26 | 20 | 26 | conserved hypothetical protein |
| SO2042 | | | 3 | 2 | | | | 2 | conserved hypothetical protein |
| SO2044 | 2 | 4 | 2 | 4 | 2 | | | 4 | lactoylglutathione lyase (gloA) |
| SO2045 | 2 | 5 | | | 3 | 3 | 3 | | cation efflux family protein |
| SO2046 | | | | 2 | | | 2 | | transcriptional regulator, MarR family |
| SO2047 | 3 | | 4 | 4 | | | 2 | 3 | prolyl oligopeptidase family protein |
| SO2048 | 9 | 9 | 8 | 2 | 10 | 11 | 6 | 7 | membrane protein, putative |
| SO2052 | | 2 | | | 2 | | 2 | | conserved hypothetical protein |
| SO2056 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO2057 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO2062 | 4 | 8 | 6 | 6 | 3 | 13 | 7 | 8 | conserved hypothetical protein |
| SO2064 | | | 2 | | 2 | | 4 | | conserved domain protein |
| SO2067 | | | | | | | 2 | 2 | phosphoribosyl-ATP pyrophosphatase/phosphoribosyl-AMP cyclohydrolase (hisI) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO2069 | | | | 2 | | | | | phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase (hisA) |
| SO2071 | | | | | | | 2 | 2 | imidazoleglycerol-phosphate dehydratase/histidinol-phosphatase (hisB) |
| SO2072 | | | | | | | | 2 | histidinol-phosphate aminotransferase (hisC) |
| SO2074 | | 3 | | | | | | | ATP phosphoribosyltransferase (hisG) |
| SO2078 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO2082 | | 5 | 3 | 4 | 3 | 2 | 2 | 2 | rod shape-determining-related protein |
| SO2083 | 5 | 6 | 5 | 5 | 5 | 7 | 5 | 6 | methyl-accepting chemotaxis protein |
| SO2085 | 24 | 26 | 21 | 24 | 27 | 23 | 16 | 18 | phenylalanyl-tRNA synthetase, alpha subunit (pheS) |
| SO2086 | 35 | 37 | 28 | 30 | 37 | 31 | 33 | 38 | phenylalanyl-tRNA synthetase, beta subunit (pheT) |
| SO2087 | 12 | 16 | 13 | 16 | 13 | 12 | 16 | 13 | integration host factor, alpha subunit (ihfA) |
| SO2088 | 13 | 8 | 8 | 5 | 11 | 10 | 5 | 9 | lipid A biosynthesis acyltransferase, putative |
| SO2098 | | | | | 4 | 2 | | | quinone-reactive Ni/Fe hydrogenase, large subunit (hyaB) |
| SO2099 | | | | | 2 | | | | quinone-reactive Ni/Fe hydrogenase, small subunit precursor (hoxK) |
| SO2107 | 5 | 12 | 11 | 11 | 8 | 6 | 10 | 10 | periplasmic glucans biosynthesis protein MdoG (mdoG-1) |
| SO2108 | 14 | 17 | 8 | 13 | 19 | 17 | 14 | 16 | periplasmic glucans biosynthesis protein MdoH (mdoH) |
| SO2110 | | | 2 | | | | | | conserved hypothetical protein |
| SO2112 | 4 | 5 | 3 | 3 | 3 | 4 | 7 | 4 | ribosomal protein L25 (rplY) |
| SO2113 | 11 | 11 | 10 | 4 | 15 | 18 | 6 | 10 | conserved hypothetical protein |
| SO2114 | 5 | 4 | 3 | 4 | 5 | 5 | 7 | 4 | conserved hypothetical protein |
| SO2116 | 3 | 3 | | 5 | 4 | | 5 | 2 | acetyltransferase, GNAT family |
| SO2117 | 3 | 4 | 3 | 4 | 4 | 3 | 3 | 4 | methyl-accepting chemotaxis protein |
| SO2123 | | | 2 | 3 | | | | | methyl-accepting chemotaxis protein |
| SO2130 | | | 2 | | | | | | ISSo1, transposase OrfB |
| SO2131 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO2133 | 2 | | | | | | | | hypothetical protein |
| SO2134 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO2135 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO2136 | | 4 | | | | | | 2 | aldehyde-alcohol dehydrogenase (adhE) |
| SO2147 | 2 | | | | | | | | exodeoxyribonuclease V, alpha subunit (recD) |
| SO2148 | 2 | | 2 | | 2 | | 2 | | exodeoxyribonuclease V, beta subunit (recB) |
| SO2149 | 3 | 2 | | | 2 | | | 2 | exodeoxyribonuclease V, gamma subunit (recC) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO2150 | 2 | | | | | 2 | | 2 | transglutaminase family protein |
| SO2151 | 2 | 3 | 3 | | 2 | 2 | 2 | 3 | conserved hypothetical protein |
| SO2165 | | | | | | 2 | | | transposase, IS91 family, putative |
| SO2168 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO2169 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO2171 | | | | 2 | | | | | ISSo1, transposase OrfB |
| SO2172 | 5 | 3 | 7 | 7 | 10 | 2 | 8 | 12 | ISSo1, transposase OrfA |
| SO2175 | 7 | 8 | 9 | 8 | 8 | 3 | 11 | 8 | conserved hypothetical protein |
| SO2176 | | | | 2 | | | 2 | | conserved hypothetical protein |
| SO2177 | 10 | 14 | 5 | 13 | 8 | 13 | 10 | 12 | conserved hypothetical protein |
| SO2178 | 2 | 2 | 4 | 3 | 3 | 3 | | | cytochrome c551 peroxidase (ccpA) |
| SO2180 | 4 | 5 | 9 | 7 | | 3 | | 4 | peptidase, M23/M37 family |
| SO2183 | 2 | 3 | 3 | 3 | 3 | | 3 | 2 | conserved hypothetical protein |
| SO2189 | 2 | 5 | | 3 | 3 | | | | conserved hypothetical protein |
| SO2190 | 3 | 3 | 4 | | 3 | 5 | 3 | 2 | creA protein (creA) |
| SO2191 | | 2 | 2 | 4 | 2 | | 2 | 2 | cystathionine beta-lyase (metC) |
| SO2192 | 5 | 4 | 3 | | 2 | 3 | 2 | 4 | sensor histidine kinase |
| SO2193 | | | | | 2 | | | | DNA-binding response regulator |
| SO2195 | | | | | | 2 | | | inter-alpha-trypsin inhibitor domain protein |
| SO2197 | | | | | | | | 2 | GGDEF family protein |
| SO2198 | 4 | 6 | 7 | 6 | 6 | | 2 | 6 | conserved hypothetical protein |
| SO2200 | 4 | 5 | 2 | 5 | 4 | 5 | 2 | | ribosomal protein S6 modification protein, putative |
| SO2201 | | | | 4 | | | 4 | 3 | conserved hypothetical protein |
| SO2202 | | | | | | | 2 | | transcriptional regulator, LysR family |
| SO2212 | | | | | | 2 | | | transposase, IS91 family, putative |
| SO2215 | 8 | 13 | 5 | 5 | 7 | 7 | | 6 | sun protein, putative |
| SO2217 | 2 | 2 | 2 | 2 | 2 | | 2 | 2 | D-alanine--D-alanine ligase (ddlA) |
| SO2218 | 17 | 21 | 24 | 24 | 22 | 9 | 21 | 33 | asparaginyl-tRNA synthetase (asnS) |
| SO2220 | | | | | | | | 2 | MutT/nudix family protein |
| SO2221 | | | 2 | | | | 2 | 2 | para-aminobenzoate synthase, component I (pabB) |
| SO2222 | 26 | 38 | 26 | 30 | 25 | 17 | 23 | 22 | fumarate hydratase, class I, anaerobic, putative |
| SO2223 | 11 | 14 | 14 | 14 | 15 | 9 | 19 | 10 | peptidase, putative |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO2228 | | | | | | | 2 | | CBS domain protein |
| SO2229 | 11 | 14 | 9 | 5 | 9 | 8 | 5 | 12 | ATP-dependent helicase HrpA (hrpA) |
| SO2236 | 3 | | | 3 | | | | 3 | PTS system, glucose-specific IIA component (crr) |
| SO2237 | | | 3 | | 3 | | | 2 | phosphoenolpyruvate-protein phosphotransferase (ptsI) |
| SO2240 | 12 | 12 | 13 | 12 | 19 | 12 | 10 | 12 | methyl-accepting chemotaxis protein |
| SO2244 | 2 | | | | | | | | transcriptional regulator, LacI family |
| SO2247 | | 4 | 5 | 9 | 5 | 5 | 6 | 7 | hypothetical protein |
| SO2248 | 5 | 16 | 13 | 14 | 11 | 3 | 12 | 15 | L-serine dehydratase 1 (sdaA) |
| SO2250 | 2 | | 2 | | 2 | 2 | 2 | 2 | beta-hexosaminidase (nagZ) |
| SO2251 | | | | | 4 | 2 | | | conserved hypothetical protein |
| SO2254 | 11 | 11 | 14 | 11 | 13 | 7 | 9 | 12 | conserved hypothetical protein |
| SO2255 | 3 | 8 | 5 | | 5 | 4 | 6 | 8 | transcription-repair coupling factor (mfd) |
| SO2257 | 2 | 5 | | 2 | 3 | 4 | 5 | 4 | lipoprotein releasing system transmembrane protein LolE (lolE) |
| SO2258 | 6 | 8 | 5 | 11 | 7 | 3 | 8 | 5 | lipoprotein releasing system ATP-binding protein LolD (lolD) |
| SO2259 | 4 | 4 | 3 | 3 | 4 | 2 | 3 | 5 | lipoprotein releasing system transmembrane protein LolC, putative |
| SO2260 | 16 | 17 | 14 | 15 | 15 | 10 | 15 | 10 | extragenic suppressor protein SuhB (suhB) |
| SO2261 | 7 | 4 | 7 | 5 | 5 | 3 | 2 | 5 | RNA methyltransferase, TrmH family, group 1 |
| SO2262 | 15 | 19 | 20 | 16 | 17 | 20 | 18 | 22 | serine acetyltransferase (cysE) |
| SO2263 | 6 | 4 | 10 | 6 | 3 | | 8 | 9 | Rrf2 family protein |
| SO2264 | 27 | 35 | 38 | 40 | 29 | 29 | 36 | 39 | cysteine desulfurase (iscS) |
| SO2265 | 8 | 10 | 11 | 13 | 5 | 6 | 10 | 13 | NifU family protein |
| SO2267 | 2 | 5 | 4 | 3 | 3 | | 4 | 6 | co-chaperone Hsc20 (hscB) |
| SO2268 | 12 | 16 | 16 | 14 | 15 | 9 | 19 | 20 | chaperone protein HscA (hscA) |
| SO2269 | 4 | 3 | 3 | 2 | 2 | 3 | | | ferredoxin, 2Fe-2S |
| SO2270 | 9 | 9 | 4 | 5 | 6 | 5 | | 7 | ribosomal protein S6 modification protein (rimK-2) |
| SO2272 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO2274 | 13 | 16 | 16 | 16 | 17 | 16 | 16 | 16 | nucleoside diphosphate kinase (ndk) |
| SO2277 | 11 | 12 | 20 | 19 | 10 | 14 | 18 | 19 | 16 kDa heat shock protein A (ibpA) |
| SO2278 | 4 | 2 | | | | 3 | | | acetolactate synthase III, small subunit (ilvH) |
| SO2279 | 5 | 4 | 4 | 4 | 6 | 3 | 8 | 8 | acetolactate synthase III, large subunit (ilvI) |
| SO2286 | 4 | 5 | 4 | 2 | 5 | 2 | 3 | 7 | sulfate permease family protein |
| SO2290 | 6 | 10 | 12 | 13 | 5 | 3 | 17 | 14 | rhodanese domain protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO2292 | 15 | 19 | 13 | 13 | 15 | 16 | 14 | 14 | ISSo4, transposase |
| SO2299 | 18 | 23 | 19 | 22 | 29 | 9 | 30 | 28 | threonyl-tRNA synthetase (thrS) |
| SO2300 | 15 | 19 | 18 | 14 | 16 | 23 | 13 | 13 | translation initiation factor IF-3 (infC) |
| SO2301 | 4 | 5 | 3 | 4 | 4 | 2 | | | ribosomal protein L35 (rpmI) |
| SO2302 | 19 | 20 | 16 | 20 | 19 | 25 | 16 | 18 | ribosomal protein L20 (rplT) |
| SO2303 | 11 | 17 | 14 | 17 | 15 | 14 | 14 | 14 | thioredoxin reductase (trxB) |
| SO2304 | 20 | 21 | 11 | 9 | 13 | 18 | 10 | 10 | alanine dehydrogenase, authentic point mutation (ald) |
| SO2305 | 11 | 13 | 10 | 6 | 10 | 12 | 8 | 9 | leucine-responsive regulatory protein (lrp) |
| SO2306 | 20 | 25 | 24 | 21 | 28 | 21 | 27 | 31 | cell division protein FtsK, putative |
| SO2307 | | 3 | | 2 | | | 2 | | outer membrane lipoprotein carrier protein LolA (lolA) |
| SO2308 | 3 | | | 2 | 3 | 5 | 6 | 4 | ATPase, AAA family |
| SO2310 | 22 | 28 | 22 | 20 | 26 | 26 | 20 | 21 | seryl-tRNA synthetase (serS) |
| SO2317 | 2 | 2 | 3 | 3 | 2 | 3 | 3 | 3 | methyl-accepting chemotaxis protein, truncation |
| SO2320 | | | | | | | | 2 | chemotaxis protein CheA, interruption-N |
| SO2321 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO2323 | 6 | 6 | 7 | 7 | 7 | 3 | 3 | 7 | methyl-accepting chemotaxis protein |
| SO2328 | 14 | 15 | 13 | 14 | 14 | 13 | 16 | 14 | translation elongation factor P (efp) |
| SO2330 | 3 | 5 | 3 | 7 | 5 | 3 | 5 | 6 | flavodoxin |
| SO2331 | 5 | | 2 | | | | | | conserved hypothetical protein |
| SO2333 | | 2 | | 2 | 3 | 2 | 4 | 2 | hydrolase, alpha/beta fold family |
| SO2335 | 4 | 7 | 6 | 3 | 7 | 5 | 2 | 3 | seqA protein (seqA) |
| SO2336 | 4 | 4 | 4 | 5 | 6 | 2 | 4 | 5 | phosphoglucomutase, alpha-D-glucose phosphate-specific (pgm) |
| SO2338 | 8 | 6 | 8 | 8 | 9 | 5 | 5 | 5 | succinylglutamate desuccinylase (astE) |
| SO2339 | 45 | 54 | 43 | 47 | 50 | 53 | 45 | 50 | alpha keto acid dehydrogenase complex, E1 component, alpha subunit |
| SO2340 | 30 | 31 | 25 | 28 | 33 | 36 | 24 | 27 | alpha keto acid dehydrogenase complex, E1 component, beta subunit |
| SO2341 | 35 | 33 | 35 | 36 | 34 | 33 | 36 | 33 | alpha keto acid dehydrogenase complex, E2 component |
| SO2342 | 4 | 6 | 5 | 6 | 9 | 3 | 2 | 4 | quinolinate synthetase complex, subunit A (nadA) |
| SO2344 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO2345 | 22 | 26 | 29 | 29 | 24 | 21 | 25 | 25 | glyceraldehyde 3-phosphate dehydrogenase (gapA-2) |
| SO2346 | | | | | | 3 | | 3 | conserved hypothetical protein |
| SO2347 | 37 | 43 | 34 | 42 | 34 | 31 | 33 | 37 | glyceraldehyde 3-phosphate dehydrogenase (gapA-3) |
| SO2350 | 15 | 14 | 8 | 16 | 18 | 4 | 6 | 13 | aspartate aminotransferase (aspC-1) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO2352 | | 3 | 2 | 2 | 2 | 4 | | | bax protein, putative |
| SO2353 | 2 | 4 | | 4 | 2 | 2 | | 2 | hypothetical protein |
| SO2354 | | 2 | | 3 | 2 | 2 | | | conserved hypothetical protein |
| SO2355 | 4 | 9 | 7 | 7 | 9 | 4 | 3 | 7 | universal stress protein family |
| SO2356 | 2 | | | | | | | 2 | electron transport regulator A (etrA) |
| SO2359 | 6 | 4 | 6 | 5 | 6 | 9 | 6 | 3 | cation transport ATPase, E1-E2 family |
| SO2360 | 6 | 10 | 6 | 6 | 6 | 10 | 6 | 6 | conserved hypothetical protein |
| SO2361 | 32 | 38 | 28 | 31 | 31 | 27 | 24 | 29 | cytochrome c oxidase, cbb3-type, subunit III (ccoP) |
| SO2363 | 17 | 24 | 18 | 20 | 24 | 11 | 16 | 20 | cytochrome c oxidase, cbb3-type, subunit II (ccoO) |
| SO2365 | | | 3 | 3 | | | 2 | 2 | conserved hypothetical protein |
| SO2366 | | | | | | | | 2 | response regulator |
| SO2374 | 5 | 4 | 2 | | 4 | 3 | 4 | 4 | transcriptional regulator, LysR family |
| SO2379 | | | | | 2 | | | | conserved hypothetical protein |
| SO2380 | | | | | 3 | 2 | | 3 | RecQ domain protein |
| SO2387 | | 2 | | | | | | | conserved hypothetical protein |
| SO2388 | | 2 | 3 | 3 | | | 3 | 2 | beta-lactamase |
| SO2390 | 8 | 8 | 8 | 6 | 12 | 6 | 13 | 14 | CDP-diacylglycerol--serine O-phosphatidyltransferase (pssA) |
| SO2392 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO2394 | | | | | | | 2 | | penicillin-binding protein 4 (dacB) |
| SO2395 | 6 | 12 | 7 | 8 | 10 | 3 | 6 | 7 | acyl-CoA dehydrogenase family protein |
| SO2396 | 3 | 3 | 5 | 3 | 4 | 4 | 5 | 4 | conserved hypothetical protein |
| SO2398 | 11 | 12 | 9 | 12 | 7 | 7 | 7 | 9 | orotidine 5`-phosphate decarboxylase (pyrF) |
| SO2399 | 4 | 4 | 9 | | 4 | 3 | 4 | 8 | conserved hypothetical protein |
| SO2400 | | | | | | 2 | | | conserved hypothetical protein |
| SO2401 | 9 | 6 | 8 | 8 | 6 | 9 | 9 | 6 | integration host factor, beta subunit (ihfB) |
| SO2402 | 83 | 119 | 96 | 110 | 97 | 99 | 100 | 108 | ribosomal protein S1 (rpsA) |
| SO2403 | 8 | 9 | 8 | 9 | 7 | 7 | 10 | 8 | cytidylate kinase (cmk) |
| SO2404 | 6 | 7 | 5 | 5 | 5 | | 5 | 4 | 3-phoshoshikimate 1-carboxyvinyltransferase (aroA) |
| SO2406 | 12 | 11 | 15 | 16 | 16 | 4 | 6 | 13 | aspartate aminotransferase (aspC-2) |
| SO2407 | 2 | 2 | 3 | 2 | 3 | 2 | | | conserved hypothetical protein |
| SO2410 | 3 | | | | | | 3 | 3 | phosphoserine aminotransferase (serC) |
| SO2411 | 42 | 49 | 37 | 48 | 39 | 38 | 33 | 42 | DNA gyrase, A subunit (gyrA) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO2413 | 7 | 11 | 9 | 8 | 4 | 9 | 8 | 13 | 3-demethylubiquinone-9 3-methyltransferase (ubiG) |
| SO2415 | 24 | 37 | 24 | 21 | 32 | 31 | 37 | 35 | ribonucleoside-diphosphate reductase, alpha subunit (nrdA) |
| SO2416 | | 4 | 3 | 3 | 2 | | 4 | 7 | ribonucleoside-diphosphate reductase, beta subunit (nrdB) |
| SO2418 | 3 | | 3 | | 2 | 3 | 2 | 2 | ISSo5, transposase |
| SO2419 | 25 | 33 | 30 | 17 | 25 | 22 | 19 | 27 | 2,4-dienoyl-CoA reductase, putative |
| SO2420 | 10 | 17 | 10 | 13 | 13 | 10 | 9 | 15 | signal peptide peptidase SppA, 67K type (sppA) |
| SO2421 | 3 | | | | | | | | L-asparaginase I (ansA) |
| SO2422 | | | 2 | 5 | | | 3 | | conserved hypothetical protein |
| SO2423 | | | | | | | | 5 | hypothetical protein |
| SO2424 | 6 | 6 | 5 | 8 | 8 | 4 | 5 | 7 | zinc carboxypeptidase domain protein |
| SO2426 | | 4 | 9 | 9 | | | 9 | 12 | DNA-binding response regulator |
| SO2427 | 55 | 62 | 56 | 59 | 51 | 50 | 51 | 51 | TonB-dependent receptor, putative |
| SO2429 | 3 | 7 | 3 | 5 | 2 | 5 | 5 | 7 | Holliday junction DNA helicase RuvB (ruvB) |
| SO2430 | | | | 2 | | | | | Holliday junction DNA helicase RuvA (ruvA) |
| SO2431 | 2 | | | | | | | | crossover junction endodeoxyribonuclease RuvC (ruvC) |
| SO2432 | | 4 | 2 | 2 | | 5 | | 2 | conserved hypothetical protein TIGR01033 |
| SO2433 | 34 | 50 | 32 | 30 | 39 | 23 | 42 | 39 | aspartyl-tRNA synthetase (aspS) |
| SO2434 | | | | | 2 | | | | extracellular solute-binding proteins, family 3/GGDEF domain protein |
| SO2435 | | 3 | 3 | | | | 2 | 3 | methyltransferase, putative |
| SO2436 | 3 | 3 | 5 | | 6 | 4 | 9 | 9 | methyltransferase, putative |
| SO2437 | 5 | 9 | 10 | 12 | 10 | 11 | 3 | 4 | hypothetical protein |
| SO2438 | 2 | | | 3 | | | | 2 | transcriptional regulator, LysR family |
| SO2441 | | 2 | | | | | | 2 | thiG protein (thiG) |
| SO2446 | | | 2 | | | | 2 | | hypothetical protein |
| SO2453 | | 2 | | | 3 | | 2 | 2 | N-ethylmaleimide reductase, putative |
| SO2460 | 2 | | | | 3 | | | | hypothetical protein |
| SO2469 | 5 | 8 | 7 | 5 | 17 | 16 | 4 | 7 | conserved hypothetical protein |
| SO2470 | | | 2 | | 4 | 2 | 2 | 4 | conserved hypothetical protein |
| SO2471 | 2 | 5 | | | 2 | | | 3 | succinyl-diaminopimelate desuccinylase (dapE) |
| SO2474 | 2 | 3 | 7 | 4 | 2 | | 5 | 7 | carbonic anhydrase family protein |
| SO2476 | 17 | 20 | 18 | 22 | 15 | 15 | 20 | 22 | polysaccharide biosynthesis protein |
| SO2477 | 11 | 9 | 12 | 9 | 8 | 4 | 12 | 9 | alcohol dehydrogenase, iron-containing |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
|---|---|---|---|---|---|---|---|---|---|
| SO2478 | 5 | 2 | 2 | 2 | 3 | 2 | | 2 | 3-deoxy-D-manno-octulosonate cytidylyltransferase (kdsB) |
| SO2479 | | | | | | 2 | | 2 | hypothetical protein |
| SO2481 | | | | | | 2 | | | conserved hypothetical protein |
| SO2483 | 19 | 13 | 8 | 14 | 18 | 9 | 12 | 15 | aspartate aminotransferase, putative |
| SO2484 | 3 | | 3 | 4 | | | 3 | | conserved hypothetical protein |
| SO2485 | 2 | | 3 | 4 | 5 | | 3 | 3 | deoxyguanosinetriphosphate triphosphohydrolase, putative |
| SO2486 | 12 | 13 | 5 | 7 | 6 | 5 | 8 | 8 | 2-deydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase (eda) |
| SO2487 | 14 | 19 | 14 | 17 | 10 | 4 | 12 | 15 | 6-phosphogluconate dehydratase (edd) |
| SO2488 | 7 | 4 | 3 | 4 | 2 | 3 | 2 | 5 | 6-phosphogluconolactonase (pgl) |
| SO2489 | 6 | 18 | 12 | 9 | 11 | 3 | 10 | 18 | glucose-6-phosphate 1-dehydrogenase (zwf) |
| SO2490 | 14 | 13 | 11 | 17 | 12 | 20 | 6 | 7 | transcriptional regulator, RpiR family |
| SO2491 | 21 | 24 | 24 | 26 | 25 | 15 | 16 | 24 | pyruvate kinase II (pykA) |
| SO2492 | 43 | 36 | 43 | 40 | 45 | 24 | 28 | 40 | oxidoreductase, acyl-CoA dehydrogenase family |
| SO2493 | 4 | 2 | 2 | 3 | | 3 | 2 | 3 | transcriptional regulator, TetR family |
| SO2494 | 23 | 35 | 28 | 24 | 24 | 26 | 19 | 31 | zinc-dependent metallopeptidase |
| SO2495 | | 2 | 2 | 5 | | | 3 | 3 | Smr domain protein |
| SO2497 | | | 2 | | 2 | | | | conserved hypothetical protein |
| SO2501 | | 3 | | 3 | 3 | | | 2 | radical activating enzyme |
| SO2503 | | | 2 | | 2 | | | | exsB protein (exsB) |
| SO2506 | | 3 | 2 | 4 | 6 | 2 | 6 | 5 | excinuclease ABC, B subunit (uvrB) |
| SO2507 | 4 | 4 | 4 | 7 | 5 | 4 | 5 | 7 | GGDEF domain protein |
| SO2509 | 5 | 4 | 4 | 3 | 3 | 2 | 3 | 2 | iron-sulfur cluster-binding protein |
| SO2510 | 24 | 27 | 15 | 24 | 21 | 21 | 18 | 22 | iron-sulfur cluster-binding protein |
| SO2512 | 8 | 7 | 4 | 4 | 5 | 10 | 12 | 8 | conserved hypothetical protein |
| SO2514 | | | | 2 | | | 3 | 4 | Endonuclease III (nth) |
| SO2525 | 4 | 3 | 3 | 3 | 5 | 3 | 3 | 3 | ABC transporter, ATP-binding protein |
| SO2527 | | | | | | 2 | | | transposase, IS91 family, putative |
| SO2529 | 4 | 7 | 3 | 3 | 6 | 7 | 5 | 8 | hypothetical protein |
| SO2530 | | | | | | | | 2 | polypeptide deformylase (def-3) |
| SO2533 | 3 | 4 | 9 | 7 | 4 | 2 | 4 | 9 | conserved hypothetical protein |
| SO2535 | | | 2 | | | | | | conserved hypothetical protein |
| SO2536 | 28 | 30 | 30 | 29 | 36 | 27 | 34 | 35 | oxidoreductase, acyl-CoA dehydrogenase family |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO2537 | 4 | 6 | 5 | 6 | 7 | 10 | 7 | 11 | sodium/hydrogen exchanger family protein |
| SO2540 | | 2 | | | | | | | response regulator |
| SO2543 | | | | 2 | | | | | sensor histidine kinase |
| SO2553 | | | | | | | | 2 | conserved hypothetical protein |
| SO2557 | | 4 | | 2 | 2 | 4 | 3 | 6 | conserved hypothetical protein |
| SO2559 | | 2 | 4 | 4 | 3 | 2 | 3 | 3 | DNA polymerase III, epsilon subunit (dnaQ-2) |
| SO2560 | | 3 | 2 | | | | | | ribonuclease HI (rnhA) |
| SO2564 | 6 | 10 | 6 | 3 | 8 | 4 | 3 | 2 | transglycosylase, Slt family |
| SO2566 | 12 | 19 | 22 | 23 | 25 | 27 | 38 | 33 | asmA protein (asmA) |
| SO2567 | 3 | 2 | 2 | 3 | 2 | | | | S-adenosylmethionine:2-demethylmenaquinone methyltransferase (menG-1) |
| SO2569 | 3 | 5 | 6 | 9 | 5 | 7 | 7 | 7 | hypothetical protein |
| SO2570 | 17 | 22 | 15 | 18 | 17 | 21 | 18 | 19 | lipoprotein, putative |
| SO2571 | 14 | 10 | 10 | 9 | 12 | 12 | 6 | 5 | ATP-dependent RNA helicase, DEAD box family |
| SO2572 | | | | 2 | 2 | | | 2 | peptidyl-prolyl cis-trans isomerase B (ppiB-2) |
| SO2573 | | | | | 2 | | | | conserved hypothetical protein |
| SO2575 | | | 2 | | | | | | conserved hypothetical protein |
| SO2576 | | | | | | | | 2 | septum site-determining protein MinC (minC) |
| SO2577 | 3 | 6 | 8 | 10 | 4 | 4 | 12 | 12 | septum site-determining protein MinD (minD) |
| SO2578 | 5 | 2 | 6 | 3 | 4 | | | 4 | cell division topological specificity factor MinE (minE) |
| SO2580 | 2 | 2 | 3 | 7 | 4 | | | 2 | ribonuclease D (rnd) |
| SO2581 | 24 | 20 | 22 | 27 | 31 | 28 | 27 | 20 | long-chain-fatty-acid--CoA ligase (fadD-1) |
| SO2583 | 9 | 9 | 10 | 15 | 8 | 3 | 6 | 9 | hypothetical protein |
| SO2587 | 8 | 10 | 8 | 10 | 19 | 13 | 10 | 11 | delta-aminolevulinic acid dehydratase (hemB-1) |
| SO2588 | | 4 | 2 | | 2 | 2 | 2 | 2 | protein-methionine-S-oxide reductase, PilB family |
| SO2589 | | | | | 2 | | | | oxidoreductase, iron/ascorbate family |
| SO2590 | 10 | 9 | 6 | 11 | 4 | 7 | 3 | 6 | GTP-binding protein |
| SO2592 | 24 | 21 | 24 | 25 | 22 | 23 | 17 | 24 | dihydroorotate dehydrogenase (pyrD) |
| SO2593 | 80 | 99 | 80 | 99 | 101 | 71 | 79 | 100 | conserved hypothetical protein |
| SO2594 | | 2 | | | 2 | 2 | | | conserved hypothetical protein |
| SO2596 | 2 | 2 | | 3 | 2 | | 2 | | conserved hypothetical protein |
| SO2600 | 3 | 3 | 3 | 6 | 4 | | 3 | 6 | aminopeptidase N (pepN) |
| SO2601 | 20 | 26 | 26 | 30 | 22 | 15 | 18 | 25 | carboxyl-terminal protease |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO2602 | 13 | 19 | 10 | 15 | 9 | 17 | 12 | 13 | conserved hypothetical protein |
| SO2603 | | | | | | | | 3 | conserved hypothetical protein |
| SO2604 | 2 | | 2 | 2 | 2 | | | 2 | conserved hypothetical protein |
| SO2606 | 29 | 39 | 32 | 31 | 29 | 31 | 26 | 24 | PqiB family protein |
| SO2610 | | 8 | 7 | 9 | 4 | | 3 | 7 | hydrolase, TatD family |
| SO2612 | | | | 2 | | | | | DNA polymerase III, delta prime subunit (holB) |
| SO2613 | 2 | | 2 | | 3 | | | 2 | thymidylate kinase (tmk) |
| SO2614 | 7 | 11 | 5 | 9 | 6 | 9 | 7 | 5 | conserved hypothetical protein TIGR00247 |
| SO2616 | 3 | 5 | 3 | 6 | 2 | 2 | 4 | 4 | deoxycytidine triphosphate deaminase (dcd) |
| SO2617 | 9 | 4 | 6 | 4 | 6 | 4 | 6 | 7 | uridine kinase (udk) |
| SO2618 | 9 | 9 | 7 | 8 | 7 | 4 | 8 | 6 | ATP-binding protein, Mrp/Nbp35 family |
| SO2619 | 14 | 27 | 23 | 27 | 22 | 9 | 22 | 28 | methionyl-tRNA synthetase (metG) |
| SO2621 | 2 | 2 | 2 | 3 | | | | 3 | conserved hypothetical protein TIGR00486 |
| SO2622 | | 4 | | 6 | | | 2 | | conserved hypothetical protein |
| SO2625 | 6 | 4 | 5 | 3 | 3 | 4 | 2 | 4 | translation initiation factor IF-1 (infA) |
| SO2626 | 8 | 11 | 10 | 19 | 10 | 10 | 12 | 19 | ATP-dependent Clp protease, ATP-binding subunit ClpA (clpA) |
| SO2628 | 2 | | | | | | | | stress response protein CspD (cspD) |
| SO2629 | 54 | 60 | 57 | 68 | 51 | 44 | 51 | 55 | isocitrate dehydrogenase, NADP-dependent (icd) |
| SO2630 | 2 | 2 | 3 | | | 2 | | 2 | RNA pseudouridylate synthase family protein |
| SO2633 | 5 | 9 | 9 | 10 | 6 | 2 | 8 | 7 | tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (trmU) |
| SO2634 | 9 | 10 | 10 | 9 | 12 | 7 | 8 | 12 | conserved hypothetical protein |
| SO2635 | 16 | 27 | 16 | 19 | 23 | 12 | 21 | 18 | adenylosuccinate lyase (purB) |
| SO2636 | 9 | 7 | 6 | 4 | 4 | 4 | 7 | 4 | conserved hypothetical protein |
| SO2637 | 7 | 7 | 5 | 9 | 3 | 12 | 4 | 5 | hypothetical protein |
| SO2638 | 17 | 14 | 16 | 14 | 16 | 20 | 19 | 20 | leucine dehydrogenase (ldh) |
| SO2643 | 30 | 35 | 22 | 19 | 36 | 34 | 21 | 16 | oxidoreductase, FAD-binding, putative |
| SO2644 | 56 | 68 | 58 | 62 | 69 | 50 | 51 | 63 | phosphoenolpyruvate synthase (ppsA) |
| SO2645 | 4 | 3 | 4 | 4 | 3 | 3 | 3 | 4 | conserved hypothetical protein |
| SO2646 | 10 | 11 | 7 | 8 | 8 | 3 | 8 | 9 | phospho-2-dehydro-3-deoxyheptonate aldolase, trp-sensitive (aroH) |
| SO2649 | 9 | 12 | 10 | 8 | 13 | 6 | 13 | 7 | cys regulon transcriptional activator (cysB) |
| SO2650 | 5 | 10 | 7 | 7 | 10 | 9 | 9 | 9 | conserved hypothetical protein |
| SO2652 | 10 | 6 | 8 | 4 | 12 | 6 | 5 | 8 | prophage MuSo2, transcriptional regulator, Cro/CI family |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO2654 | | 2 | | | | | | | transposase, putative |
| SO2655 | 2 | 2 | 2 | | 4 | | | | prophage MuSo2, DNA transposition protein, putative |
| SO2660 | | | | | | | 2 | | conserved hypothetical protein |
| SO2663 | | | | | | 2 | 2 | | conserved hypothetical protein |
| SO2681 | | 2 | | | | | | | prophage MuSo2, F protein, putative |
| SO2682 | 4 | 3 | 3 | 3 | 3 | 3 | | 3 | hypothetical protein |
| SO2683 | | | | 2 | | 2 | | | hypothetical protein |
| SO2697 | | 2 | | | | | | | conserved hypothetical protein |
| SO2698 | 2 | | | | | | | | prophage MuSo2, DNA circulation protein, putative |
| SO2705 | 34 | 40 | 35 | 31 | 33 | 35 | 21 | 31 | DNA topoisomerase I (topA) |
| SO2706 | 12 | 8 | 6 | 9 | 4 | 4 | 3 | 7 | succinylarginine dihydrolase (astB) |
| SO2707 | 2 | | | | | | | | acetyltransferase, GNAT family |
| SO2708 | 3 | 4 | 2 | 3 | 2 | 2 | 3 | 3 | nitroreductase family protein |
| SO2714 | 2 | 2 | | 2 | | 2 | | | hypothetical protein |
| SO2715 | | 3 | | | | | | 3 | TonB-dependent receptor |
| SO2720 | 3 | 3 | 2 | 3 | 2 | 4 | | 2 | conserved hypothetical protein |
| SO2721 | | | | | | | 3 | 2 | conserved hypothetical protein |
| SO2722 | 2 | 3 | | | | | 3 | 3 | conserved hypothetical protein |
| SO2723 | 2 | 5 | | 3 | 3 | 4 | 2 | 4 | HIT family protein |
| SO2725 | 2 | | | | | | 2 | 3 | transcriptional regulator, LuxR family |
| SO2726 | | | | | | 3 | | | cytochrome b, putative |
| SO2728 | 6 | 5 | 8 | 9 | 6 | | 15 | 10 | peptidase HtpX (htpX) |
| SO2730 | 4 | 6 | 5 | 6 | 3 | 3 | 4 | 8 | peptidase E (pepE) |
| SO2731 | 19 | 22 | 24 | 28 | 21 | 10 | 15 | 17 | periplasmic glucans biosynthesis protein MdoG (mdoG-2) |
| SO2737 | 3 | 2 | 2 | 3 | 2 | | 3 | | dethiobiotin synthase (bioD) |
| SO2739 | 2 | | 4 | | 2 | 2 | 3 | 5 | 8-amino-7-oxononanoate synthase (bioF) |
| SO2740 | 7 | 10 | 10 | 9 | 7 | 7 | 6 | 12 | biotin synthase (bioB) |
| SO2741 | 2 | | | 2 | 2 | | | 2 | adenosylmethionine--8-amino-7-oxononanoate aminotransferase (bioA) |
| SO2742 | 4 | 7 | 3 | 7 | 9 | 9 | 4 | 6 | sensor histidine kinase/response regulator |
| SO2743 | 13 | 15 | 13 | 15 | 10 | 10 | 15 | 18 | acetyl-coenzyme A synthetase (acs) |
| SO2744 | | | | | | | | 2 | helicase |
| SO2745 | 7 | 7 | 7 | 9 | 6 | 3 | 8 | 13 | glutaredoxin |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO2746 | 13 | 15 | 12 | 15 | 11 | 8 | 7 | 9 | conserved hypothetical protein |
| SO2747 | 13 | 21 | 15 | 18 | 12 | 16 | 17 | 17 | peptidoglycan-associated lipoprotein (pal) |
| SO2748 | 25 | 15 | 21 | 17 | 16 | 16 | 9 | 14 | tolB protein (tolB) |
| SO2749 | 12 | 23 | 14 | 19 | 10 | 19 | 15 | 21 | tolA protein (tolA) |
| SO2750 | | 2 | | | 2 | | | 2 | tolr protein (tolR) |
| SO2751 | 9 | 20 | 12 | 11 | 15 | 14 | 10 | 16 | MotA/TolQ/ExbB proton channel family protein |
| SO2752 | | | | 2 | | | | | conserved hypothetical protein TIGR00051 |
| SO2753 | 28 | 32 | 24 | 28 | 38 | 27 | 22 | 23 | prolyl endopeptidase |
| SO2755 | | 2 | 3 | 3 | 3 | | 2 | 3 | ribonuclease T (rnt) |
| SO2756 | 21 | 21 | 25 | 29 | 23 | 26 | 26 | 25 | antioxidant, AhpC/Tsa family, authentic frameshift |
| SO2759 | 10 | 14 | 11 | 12 | 14 | 12 | 8 | 9 | uracil phosphoribosyltransferase (upp) |
| SO2760 | 13 | 18 | 11 | 17 | 13 | 9 | 13 | 16 | phosphoribosylformylglycinamidine cyclo-ligase (purM) |
| SO2761 | 9 | 11 | 12 | 12 | 11 | 5 | 8 | 10 | phosphoribosylglycinamide formyltransferase (purN) |
| SO2762 | 4 | 7 | 5 | 8 | 8 | 3 | 5 | 8 | nagD protein |
| SO2763 | 11 | 6 | 5 | 8 | 4 | 12 | 4 | 5 | conserved hypothetical protein |
| SO2766 | 30 | 35 | 34 | 26 | 31 | 31 | 35 | 26 | conserved hypothetical protein |
| SO2768 | | | | | 2 | | 2 | | acyl-CoA dehydrogenase family protein |
| SO2769 | | | | | | | 2 | | conserved hypothetical protein |
| SO2771 | 2 | 5 | 4 | 3 | 7 | | | 7 | 2-hydroxy-3-oxopropionate reductase (garR) |
| SO2774 | 12 | 14 | 11 | 13 | 9 | 7 | 7 | 10 | 3-oxoacyl-(acyl-carrier-protein) synthase II (fabF-1) |
| SO2775 | 6 | 5 | | 4 | 4 | 4 | 3 | 4 | acyl carrier protein (acpP) |
| SO2776 | 16 | 16 | 14 | 16 | 11 | 16 | 14 | 8 | 3-oxoacyl-(acyl-carrier-protein) reductase (fabG-1) |
| SO2777 | 9 | 10 | 7 | 6 | 5 | 9 | 2 | 4 | malonyl CoA-acyl carrier protein transacylase (fabD) |
| SO2778 | 5 | 9 | 9 | 7 | 7 | 3 | 6 | 9 | 3-oxoacyl-(acyl-carrier-protein) synthase III (fabH-1) |
| SO2779 | 4 | 3 | 3 | 2 | 5 | 3 | 2 | 2 | fatty acid/phospholipid synthesis protein PlsX (plsX) |
| SO2780 | 6 | 8 | 10 | 9 | 9 | 7 | 10 | 9 | ribosomal protein L32 (rpmF) |
| SO2781 | 3 | 3 | 5 | 3 | 2 | 2 | 2 | 4 | conserved hypothetical protein |
| SO2782 | 2 | 2 | 2 | 2 | 3 | | 4 | 3 | maf protein, putative |
| SO2784 | 10 | 12 | 3 | 8 | 6 | 9 | 6 | 6 | ribosomal large subunit pseudouridine synthase C (rluC) |
| SO2785 | 82 | 94 | 87 | 95 | 91 | 93 | 87 | 100 | ribonuclease E (rne) |
| SO2787 | 20 | 18 | 19 | 17 | 18 | 14 | 19 | 18 | cold shock domain family protein |
| SO2788 | | 4 | 2 | 3 | 5 | 2 | | | ribosomal RNA large subunit methyltransferase A (rrmA) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
|---|---|---|---|---|---|---|---|---|---|
| SO2790 | 4 | 5 | 4 | 8 | 3 | 3 | 6 | 6 | exodeoxyribonuclease I (sbcB) |
| SO2794 | | 3 | | 3 | 2 | 3 | | 2 | conserved hypothetical protein |
| SO2796 | 56 | 65 | 49 | 56 | 61 | 57 | 51 | 56 | conserved hypothetical protein |
| SO2797 | | 2 | 4 | 2 | | | 2 | 2 | conserved hypothetical protein |
| SO2799 | 3 | 5 | | 5 | 3 | 2 | 2 | 6 | lipoprotein, putative |
| SO2800 | 3 | 2 | | | 4 | 3 | 8 | 7 | conserved hypothetical protein |
| SO2801 | 3 | 3 | 2 | 6 | | 4 | 4 | 6 | tetraacyldisaccharide 4-kinase (lpxK) |
| SO2802 | 15 | 13 | 16 | 16 | 12 | 8 | 15 | 15 | ABC transporter, ATP-binding protein MsbA (msbA) |
| SO2807 | | | | 2 | | | 3 | | hypothetical protein |
| SO2811 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO2813 | 2 | 2 | 2 | 2 | 2 | | | | oxidoreductase, short chain dehydrogenase/reductase family |
| SO2815 | 3 | | 3 | 2 | 3 | 2 | 3 | 5 | CBS domain protein |
| SO2817 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO2822 | | | | | | | 2 | 2 | sensor histidine kinase |
| SO2827 | 4 | 2 | 3 | 2 | 3 | | | | conserved hypothetical protein |
| SO2831 | 3 | 4 | 2 | 2 | 2 | 2 | | 2 | GTP cyclohydrolase II (ribA) |
| SO2832 | 7 | 12 | 7 | 9 | 5 | 8 | 3 | 7 | conserved hypothetical protein |
| SO2834 | | | | | 5 | 5 | | 2 | anaerobic ribonucleoside-triphosphate reductase (nrdD) |
| SO2836 | | | | | | | | 2 | conserved hypothetical protein |
| SO2838 | 15 | 8 | 3 | 6 | 9 | 6 | 5 | 12 | ATP-dependent RNA helicase, DEAD box family |
| SO2839 | 3 | 6 | 5 | 7 | 7 | 5 | 7 | 6 | hypothetical protein |
| SO2840 | | | | | | | 2 | | conserved hypothetical protein |
| SO2841 | | | 3 | 2 | | | 2 | 5 | hypothetical protein |
| SO2842 | 3 | 4 | 2 | 4 | 4 | 6 | 3 | 3 | peptidase, M23/M37 family |
| SO2843 | 8 | 10 | 7 | 7 | 9 | 8 | 7 | 7 | exonuclease SbcC, putative |
| SO2844 | | | 2 | 3 | | | | 4 | exonuclease SbcD, putative |
| SO2847 | | | 2 | | | | | | transcriptional regulator, LysR family |
| SO2848 | | | 4 | 2 | 2 | | | | hypothetical protein |
| SO2849 | | | | | 2 | | | | acetyltransferase, GNAT family |
| SO2851 | 3 | 3 | 4 | 7 | 4 | 4 | 4 | 7 | histidinol phosphatase domain protein |
| SO2852 | 2 | 4 | 2 | 5 | 2 | 3 | 5 | 4 | transcriptional regulator, GntR family |
| SO2853 | 8 | 7 | 5 | 5 | 4 | 5 | 3 | 4 | 3-oxoacyl-(acyl-carrier-protein) synthase III, putative |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO2856 | 2 | | | | | | | | CBS domain protein |
| SO2857 | | 2 | 2 | 2 | | | 2 | | sodium/solute symporter family protein |
| SO2860 | | | | | | | | 3 | thiol:disulfide interchange protein, DsbA family |
| SO2861 | | | 3 | 2 | | | | | conserved hypothetical protein |
| SO2862 | | | | | | | 2 | | HDIG domain protein |
| SO2866 | 3 | 2 | 2 | | 3 | 2 | | | chromosome initiation inhibitor (iciA) |
| SO2869 | | 4 | 3 | 7 | 2 | | | 5 | conserved hypothetical protein |
| SO2871 | | | 3 | 4 | | | | 2 | arsenate reductase, putative |
| SO2877 | | | 2 | 2 | | | | | conserved hypothetical protein |
| SO2878 | 9 | 10 | 11 | 8 | 11 | 11 | 6 | 8 | conserved hypothetical protein |
| SO2880 | | | | | | 2 | | | glutaredoxin domain protein |
| SO2881 | 5 | 3 | 4 | 7 | 3 | | 3 | 7 | superoxide dismutase, Fe (sodB) |
| SO2882 | | | | | | 2 | 2 | 2 | conserved hypothetical protein |
| SO2885 | 3 | | | | | 2 | 4 | 2 | fatty acid metabolism regulator protein (fadR) |
| SO2886 | | | | | 2 | | | | Na+/H+ antiporter (nhaB) |
| SO2889 | 3 | 5 | 3 | 3 | | 2 | 2 | | sensory box histidine kinase |
| SO2893 | | | | | | 2 | | | conserved hypothetical protein |
| SO2894 | 3 | | 2 | 3 | | | 2 | | YaiI/YqxD family protein |
| SO2895 | 3 | 3 | 6 | | 2 | 4 | | | pyridoxamine 5-phosphate oxidase (pdxH) |
| SO2896 | 23 | 16 | 18 | 20 | 19 | 16 | 8 | 15 | DNA ligase, NAD-dependent (ligA) |
| SO2897 | 16 | 23 | 18 | 17 | 16 | 15 | 13 | 19 | cell division protein ZipA (zipA) |
| SO2898 | 2 | | | 3 | 3 | 4 | 2 | 4 | SMC family protein |
| SO2901 | | | | | | | 2 | 3 | 3-oxoacyl-(acyl-carrier-protein) synthase III (fabH-2) |
| SO2902 | | | | | | | 2 | | hypothetical protein |
| SO2903 | 37 | 50 | 52 | 58 | 37 | 41 | 52 | 53 | cysteine synthase A (cysK) |
| SO2907 | 40 | 63 | 34 | 37 | 41 | 48 | 32 | 32 | TonB-dependent receptor domain protein |
| SO2912 | 13 | 13 | 33 | 45 | 23 | 13 | 48 | 45 | formate acetyltransferase (pflB) |
| SO2915 | 6 | 7 | 9 | 10 | 5 | 3 | 16 | 17 | acetate kinase (ackA) |
| SO2916 | 24 | 31 | 32 | 35 | 32 | 25 | 30 | 42 | phosphate acetyltransferase (pta) |
| SO2917 | 2 | | | 2 | 2 | | | 2 | hypothetical protein |
| SO2919 | | 2 | | | | | | | conserved hypothetical protein |
| SO2921 | | | | | 2 | | 2 | | D-erythro-7,8-dihydroneopterin triphosphate epimerase (folX) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO2923 | | | 2 | | 2 | 2 | | 2 | sodium/glutamate symporter (gltS) |
| SO2926 | 9 | 10 | 6 | 9 | 12 | 8 | 7 | 11 | ABC transporter, permease, putative |
| SO2927 | 3 | 6 | 5 | 4 | 4 | 7 | 3 | 4 | ABC transporter, ATP-binding protein |
| SO2929 | 30 | 31 | 8 | 13 | 36 | 34 | 14 | 12 | hypothetical protein |
| SO2933 | 14 | 20 | 14 | 13 | 12 | 11 | 7 | 12 | sohB protein, peptidase U7 family |
| SO2934 | 3 | 5 | 2 | 4 | 4 | 7 | | 4 | conserved hypothetical protein |
| SO2935 | 5 | 6 | 5 | 8 | 6 | 5 | 5 | 9 | oxidoreductase, short-chain dehydrogenase/reductase family |
| SO2937 | 13 | 15 | 12 | 15 | 16 | 12 | 10 | 12 | RNA pseudouridylate synthase family protein |
| SO2938 | 6 | 7 | 8 | 9 | 5 | 6 | 8 | 7 | hypothetical protein |
| SO2939 | | | | | | 2 | | | hypothetical protein |
| SO2942 | | | | 2 | | | | | hypothetical protein |
| SO2945 | | 2 | | | 2 | 3 | | 2 | hypothetical protein |
| SO2951 | | | | | 4 | 6 | | | hypothetical protein |
| SO2953 | | | | | 4 | 2 | | | prophage LambdaSo, tail length tape meausure protein (H) |
| SO2963 | 4 | 5 | | 3 | 8 | 6 | 4 | 7 | prophage LambdaSo, major capsid protein, HK97 family |
| SO2967 | 2 | | 2 | | | | 2 | | conserved hypothetical protein |
| SO2968 | | 2 | 2 | 2 | | 2 | | 2 | conserved hypothetical protein |
| SO2969 | | | | | | | | 2 | prophage LambdaSo, holin, putative |
| SO2988 | | | 2 | | 3 | | 3 | 2 | conserved hypothetical protein |
| SO2990 | 2 | 3 | 2 | | | 4 | | | prophage LambdaSo, transcriptional regulator, Cro/CI family |
| SO2991 | 4 | 8 | 6 | 4 | 3 | 4 | 6 | 5 | conserved hypothetical protein |
| SO2992 | | 2 | 2 | 3 | | | 2 | 3 | hypothetical protein |
| SO2993 | | | 2 | | 3 | | 9 | 8 | prophage LambdaSo, type II DNA modification methyltransferase, putative, truncation |
| SO3000 | | | | | | | 3 | 2 | conserved hypothetical protein |
| SO3002 | | | | | | | | 2 | conserved hypothetical protein |
| SO3004 | | | | | | | 2 | 2 | prophage LambdaSo, DNA modification methyltransferase, putative |
| SO3006 | 2 | 2 | | | | | 3 | | prophage LambdaSo, type II DNA modification methyltransferase, putative |
| SO3008 | 2 | | | | | | | | hypothetical protein |
| SO3013 | | 2 | 2 | 2 | | 2 | 2 | | site-specific recombinase, phage integrase family |
| SO3016 | 4 | | | 2 | | 3 | 2 | | Sua5/YciO/YrdC/YwlC family protein |
| SO3017 | 2 | 2 | | | | | 3 | | TrpH family protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3019 | | 2 | | | 3 | | | | anthranilate synthase component I (trpE) |
| SO3022 | | | 2 | | 2 | | | | indole-3-glycerol phosphate synthase/phosphoribosylanthranilate isomerase (trpC/F) |
| SO3023 | 4 | 7 | 6 | 6 | 5 | 2 | 6 | 7 | tryptophan synthase, beta subunit (trpB) |
| SO3024 | 7 | 8 | 6 | 5 | 2 | 8 | 2 | 2 | tryptophan synthase, alpha subunit (trpA) |
| SO3025 | | | | 2 | | | 2 | 2 | conserved hypothetical protein |
| SO3030 | 10 | 8 | 28 | 28 | 3 | | 33 | 36 | siderophore biosynthesis protein (alcA) |
| SO3031 | | | 3 | 3 | | | | 2 | siderophore biosynthesis protein, putative |
| SO3032 | 5 | 5 | 12 | 16 | 3 | | 19 | 17 | siderophore biosynthesis protein, putative |
| SO3033 | 23 | 19 | 44 | 43 | 4 | 2 | 54 | 52 | ferric alcaligin siderophore receptor |
| SO3034 | | | | | | | 2 | 3 | ferric iron reductase protein, putative |
| SO3036 | 2 | 2 | 2 | | 2 | 2 | 2 | | conserved hypothetical protein |
| SO3037 | 8 | 13 | 6 | 5 | 7 | 9 | 8 | 5 | exodeoxyribonuclease III (xth) |
| SO3043 | | | | | | | 2 | | hypothetical protein |
| SO3044 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO3048 | | 2 | | | | | 2 | | isoquinoline 1-oxidoreductase, beta subunit, putative |
| SO3049 | | | | | | 2 | | | isoquinoline 1-oxidoreductase, alpha subunit, putative |
| SO3052 | 8 | 7 | 10 | 10 | 9 | 6 | 17 | 13 | methyl-accepting chemotaxis protein |
| SO3054 | | | | | 3 | 2 | | | metallo-beta-lactamase family protein |
| SO3055 | | | | | | | 2 | 2 | proline iminopeptidase (pip) |
| SO3057 | | | | | 2 | | 3 | 2 | Pal/histidase family protein |
| SO3061 | 6 | 2 | 5 | 8 | 3 | 2 | 11 | 13 | DNA topoisomerase III (topB) |
| SO3063 | | | | 2 | | | | 3 | sodium:alanine symporter family protein |
| SO3064 | 23 | 28 | 22 | 29 | 28 | 18 | 23 | 26 | amidophosphoribosyltransferase (purF) |
| SO3066 | 10 | 25 | 15 | 18 | 13 | 23 | 14 | 22 | conserved hypothetical protein |
| SO3067 | | 2 | 2 | 4 | 3 | | | 4 | FolC bifunctional protein (folC) |
| SO3068 | 2 | 2 | | | | 2 | | | tRNA pseudouridine synthase A (truA) |
| SO3069 | 14 | 18 | 18 | 21 | 14 | 17 | 16 | 24 | conserved hypothetical protein |
| SO3070 | 18 | 18 | 15 | 19 | 18 | 14 | 17 | 17 | aspartate semialdehyde dehydrogenese (asd) |
| SO3071 | 2 | 2 | | | | | | | erythronate-4-phosphate dehydrogenase (pdxB) |
| SO3072 | 21 | 31 | 30 | 27 | 25 | 19 | 25 | 21 | 3-oxoacyl-(acyl-carrier-protein) synthase I (fabB) |
| SO3073 | 2 | 2 | | | | | | | conserved hypothetical protein |
| SO3077 | 3 | 5 | 3 | 5 | 2 | | | | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3080 | 2 | | | | | | | 2 | hemK family protein |
| SO3081 | | | | | | | | 2 | Smr domain protein |
| SO3083 | | | | | | | 2 | | peptidase, M16 family |
| SO3084 | 2 | 3 | 5 | 4 | 4 | 4 | 4 | 6 | sensory box protein |
| SO3088 | 14 | 17 | 23 | 11 | 15 | 6 | 21 | 23 | fatty oxidation complex, alpha subunit |
| SO3089 | 9 | 12 | 10 | 11 | 11 | 5 | 11 | 13 | fatty oxidation complex, beta subunit |
| SO3090 | 14 | 14 | 8 | 17 | 13 | 13 | 12 | 17 | MoxR domain protein |
| SO3091 | 5 | | 2 | | 3 | | | | conserved hypothetical protein |
| SO3092 | 6 | 9 | 5 | 9 | 10 | 12 | 8 | 8 | hypothetical protein |
| SO3093 | 3 | 3 | 3 | 4 | | 2 | 3 | 8 | von Willebrand factor type A domain protein |
| SO3094 | 6 | 11 | 5 | 4 | 8 | 10 | 7 | 12 | TPR domain protein |
| SO3095 | 6 | 15 | 16 | 14 | 17 | 6 | 16 | 11 | conserved hypothetical protein |
| SO3096 | 5 | 5 | 7 | 4 | 7 | 3 | 2 | 6 | RNA polymerase sigma-70 factor, ECF subfamily |
| SO3097 | 4 | 4 | 4 | 2 | 4 | | 5 | 2 | conserved hypothetical protein |
| SO3099 | 17 | 21 | 16 | 22 | 14 | 19 | 23 | 26 | long-chain fatty acid transport protein, putative |
| SO3101 | 14 | 21 | 17 | 14 | 19 | 18 | 17 | 19 | conserved hypothetical protein |
| SO3102 | 3 | | | | | | | 2 | AcrA/AcrE family protein |
| SO3103 | 6 | 11 | 7 | 9 | 16 | 12 | 15 | 11 | AcrB/AcrD/AcrF family protein |
| SO3105 | 2 | 3 | 4 | 2 | | 3 | | 2 | phage shock protein E (pspE-1) |
| SO3108 | | 4 | 6 | 7 | 5 | | 4 | 9 | siroheme synthase, N-terminal component, putative |
| SO3110 | 11 | 10 | 11 | 12 | 12 | 11 | 9 | 7 | protein-export membrane protein SecF (secF-2) |
| SO3111 | 38 | 52 | 45 | 52 | 39 | 37 | 44 | 45 | protein-export membrane protein SecD (secD-2) |
| SO3112 | 4 | 6 | 6 | 9 | 2 | 4 | 5 | 7 | preprotein translocase, YajC subunit (yajC) |
| SO3113 | 17 | 23 | 19 | 21 | 13 | 17 | 16 | 21 | queuine tRNA-ribosyltransferase (tgt) |
| SO3114 | 5 | 9 | 4 | 6 | 6 | 2 | 4 | 5 | S-adenosylmethionine:tRNA ribosyltransferase-isomerase (queA) |
| SO3116 | | | 2 | | 2 | | | | conserved hypothetical protein |
| SO3118 | 5 | 4 | 4 | 2 | 3 | 2 | 2 | 6 | ribosomal large subunit pseudouridine synthase A (rluA-1) |
| SO3120 | | 4 | 6 | 7 | 6 | 2 | 8 | 7 | oxidoreductase, Gfo/Idh/MocA family |
| SO3122 | 3 | 3 | 4 | 3 | 4 | | 2 | | sodium/dicarboxylate symporter |
| SO3124 | | | 3 | 4 | 2 | | | 4 | tyrosine-specific protein phosphatase, putative |
| SO3125 | 9 | 12 | 3 | 7 | 4 | 3 | 7 | 6 | ATP-dependent RNA helicase, DEAD box family |
| SO3126 | | | | | 2 | | | | methylated-DNA--protein-cysteine methyltransferase (ogt) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3127 | | | | | | | 2 | | Ada regulatory protein, putative |
| SO3128 | 8 | 9 | 10 | 6 | 7 | 7 | 9 | 6 | ABC transporter, ATP-binding/permease protein, putative |
| SO3133 | | | 2 | | | | | | hypothetical protein |
| SO3134 | 2 | | | | | | | | C4-dicarboxylate-binding periplasmic protein (dctP) |
| SO3138 | | | 2 | 3 | 3 | | | 3 | C4-dicarboxylate transport transcriptional regulatory protein (dctD) |
| SO3140 | 3 | | 2 | | | | | | thymidine kinase |
| SO3142 | 43 | 45 | 43 | 54 | 46 | 41 | 51 | 52 | peptidyl-dipeptidase Dcp (dcp-1) |
| SO3144 | 8 | 9 | 8 | 5 | 9 | 6 | 6 | 7 | electron transfer flavoprotein, alpha subunit (etfA) |
| SO3145 | 12 | 9 | 10 | 9 | 7 | 12 | 11 | 9 | electron transfer flavoprotein, beta subunit (etfB) |
| SO3146 | 28 | 37 | 33 | 33 | 32 | 36 | 34 | 36 | DNA-binding protein, H-NS family |
| SO3148 | | 2 | | | | | | | conserved hypothetical protein |
| SO3149 | | 3 | | 3 | | | | | conserved hypothetical protein |
| SO3150 | 3 | 7 | 8 | 8 | 4 | 7 | 4 | 7 | lipoprotein, putative |
| SO3154 | 22 | 32 | 21 | 29 | 29 | 11 | 27 | 38 | prolyl-tRNA synthetase (proS) |
| SO3155 | 2 | 2 | 4 | | | 4 | | 3 | conserved hypothetical protein |
| SO3157 | 5 | 5 | 4 | 6 | 5 | 5 | 3 | 5 | lipoprotein, putative |
| SO3158 | 8 | 9 | 7 | 6 | 5 | 9 | 6 | 4 | polysaccharide synthesis-related protein |
| SO3159 | 14 | 20 | 19 | 13 | 15 | 10 | 10 | 17 | conserved hypothetical protein |
| SO3163 | 2 | 2 | 3 | | 5 | 5 | 4 | 4 | lipoprotein |
| SO3166 | | | | 2 | | | | | conserved hypothetical protein |
| SO3168 | 2 | 2 | 4 | 4 | 5 | 2 | 7 | 5 | DnaJ domain protein |
| SO3171 | 14 | 19 | 13 | 11 | 13 | 10 | 7 | 11 | polysaccharide biosynthesis protein |
| SO3172 | 8 | 14 | 12 | 12 | 13 | 12 | 13 | 13 | galactosyl transferase |
| SO3173 | 6 | 3 | 6 | 6 | 6 | 5 | 6 | 7 | UDP-galactose 4-epimerase, putative |
| SO3174 | 13 | 6 | 12 | 11 | 11 | 10 | 12 | 16 | glycosyl transferase, group 1 family protein |
| SO3175 | 18 | 23 | 17 | 23 | 20 | 15 | 19 | 18 | asparagine synthetase, glutamine-hydrolyzing (asnB-2) |
| SO3176 | 3 | 7 | | | 2 | 6 | | | glycosyl transferase, group 1 family protein |
| SO3177 | 12 | 11 | 10 | 7 | 9 | 12 | 5 | 4 | formyl transferase domain protein |
| SO3178 | 18 | 23 | 21 | 18 | 18 | 17 | 22 | 19 | hypothetical protein |
| SO3180 | 4 | 5 | 3 | 2 | 5 | 2 | 3 | 3 | glycosyl transferase, group 2 family protein |
| SO3182 | 2 | | | | | 2 | | 2 | acetyltransferase, GNAT family |
| SO3183 | 3 | 3 | 5 | 4 | 4 | 3 | | 5 | perosamine synthetase-related protein |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO3184 | 5 | 2 | | 4 | 6 | | 7 | 7 | conserved hypothetical protein |
| SO3185 | 5 | 7 | 6 | 8 | 10 | 3 | 5 | 6 | polysaccharide biosynthesis protein |
| SO3186 | 6 | 9 | 6 | 8 | 11 | 4 | 9 | 9 | glucose-1-phosphate-thymidylyltransferase (rfbA) |
| SO3188 | 5 | 7 | 7 | 9 | 9 | 3 | 8 | 11 | dTDP-glucose 4,6-dehydratase (rfbB) |
| SO3189 | 9 | 11 | 8 | 9 | 6 | 5 | 7 | 10 | polysaccharide biosynthesis protein |
| SO3190 | 22 | 34 | 33 | 28 | 28 | 20 | 26 | 33 | polysaccharide biosynthesis protein |
| SO3191 | 9 | 15 | 7 | 14 | 15 | 13 | 11 | 13 | chain length determinant protein |
| SO3193 | 53 | 73 | 62 | 54 | 49 | 59 | 63 | 56 | polysaccharide biosynthesis protein |
| SO3194 | 3 | 5 | 5 | 5 | 4 | 5 | 5 | 2 | transcriptional activator rfaH, putative |
| SO3197 | 10 | 11 | 9 | 9 | 9 | 8 | 9 | 9 | vacJ lipoprotein, putative |
| SO3198 | 3 | 4 | 3 | | 5 | 2 | 2 | | hypothetical protein |
| SO3200 | | 2 | 3 | 5 | | | 2 | 3 | conserved hypothetical protein |
| SO3202 | 2 | | | 2 | 4 | 2 | 5 | 4 | purine-binding chemotaxis protein CheW (cheW-3) |
| SO3203 | 2 | | | | 3 | | | | CheW domain protein |
| SO3204 | 9 | 7 | 6 | 3 | 3 | 4 | 7 | 6 | ParA family protein |
| SO3205 | 2 | 2 | 5 | 5 | 2 | 3 | 4 | 3 | hypothetical protein |
| SO3206 | 4 | 10 | 4 | 8 | | 7 | 2 | 3 | protein-glutamate methylesterase CheB (cheB-3) |
| SO3207 | 20 | 27 | 22 | 20 | 22 | 20 | 21 | 30 | chemotaxis protein |
| SO3208 | 3 | 3 | 5 | 5 | 4 | 2 | 7 | 4 | chemotaxis protein CheZ (cheZ) |
| SO3209 | 2 | 5 | | 5 | 3 | 4 | 10 | 8 | chemotaxis protein CheY (cheY-3) |
| SO3210 | 4 | 4 | 3 | 4 | 3 | | 3 | 6 | RNA polymerase sigma-27 factor (fliA) |
| SO3211 | 8 | 10 | 6 | 9 | 10 | 9 | 8 | 6 | flagellar biosynthetic protein FlhG (flhG) |
| SO3212 | 8 | 13 | 17 | 19 | 8 | 13 | 9 | 17 | flagellar biosynthetic protein FlhF (flhF) |
| SO3213 | 3 | | 2 | 2 | 6 | 2 | 4 | 3 | flagellar biosynthesis protein FlhA (flhA) |
| SO3215 | | 2 | | | | | | 3 | flagellar biosynthetic protein FlhB (flhB) |
| SO3219 | | | | 3 | | | 2 | | flagellar protein FliO (fliO) |
| SO3220 | | | | 2 | | | | 2 | flagellar motor switch protein FliN (fliN) |
| SO3221 | 2 | | | 2 | 2 | | 2 | 2 | flagellar motor switch protein FliM (fliM) |
| SO3222 | 3 | 4 | 4 | 2 | 3 | 12 | 4 | 6 | flagellar protein FliL (fliL) |
| SO3223 | | | 5 | | | | 4 | | flagellar hook-length control protein FliK (fliK) |
| SO3224 | | | | | | | | 2 | flagellar protein FliJ (fliJ) |
| SO3225 | | | | 3 | | | | 2 | flagellum-specific ATP synthase FliI (fliI) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3226 | | 5 | 2 | 4 | 4 | 4 | | 6 | flagellar assembly protein FliH (fliH) |
| SO3227 | 9 | 13 | 12 | 10 | 9 | 4 | 14 | 12 | flagellar motor switch protein FliG (fliG) |
| SO3228 | 9 | 11 | 12 | 13 | 8 | 14 | 9 | 5 | flagellar M-ring protein FliF (fliF) |
| SO3230 | | 4 | | | 4 | 3 | | 2 | flagellar regulatory protein C (flrC) |
| SO3231 | 2 | 3 | 2 | 6 | 5 | 7 | 5 | 8 | flagellar regulatory protein B (flrB) |
| SO3232 | 5 | 8 | 5 | 10 | 8 | 7 | 8 | 11 | flagellar regulatory protein A (flrA) |
| SO3233 | 6 | 9 | 9 | 8 | 5 | 6 | 8 | 8 | flagellar protein FliS (fliS) |
| SO3235 | 4 | 2 | 5 | 5 | 2 | 2 | | 7 | flagellar hook-associated protein FliD (fliD) |
| SO3237 | 24 | 21 | 23 | 23 | 21 | 16 | 18 | 18 | flagellin |
| SO3238 | 22 | 20 | 20 | 22 | 22 | 15 | 18 | 18 | flagellin |
| SO3239 | 4 | 2 | | 3 | 3 | 2 | | | flagellar hook-associated protein FlgL (flgL) |
| SO3241 | | | | 2 | | | | | flagellar protein FlgJ (flgJ) |
| SO3242 | 3 | 6 | 3 | 4 | 3 | | 2 | 4 | flagellar P-ring protein FlgI (flgI) |
| SO3243 | 3 | 5 | 10 | 5 | | | 5 | 5 | flagellar L-ring protein FlgH (flgH) |
| SO3244 | 2 | 2 | | | | | | 4 | flagellar basal-body rod protein FlgG (flgG) |
| SO3245 | | | | 3 | | | | | flagellar basal-body rod protein FlgF (flgF) |
| SO3247 | 8 | 12 | 8 | 8 | 6 | 7 | 5 | 3 | flagellar hook protein FlgE (flgE) |
| SO3248 | | 3 | | 2 | 2 | | | | basal-body rod modification protein FlgD (flgD) |
| SO3251 | 6 | 5 | 5 | 3 | 3 | 3 | 2 | 7 | chemotaxis protein methyltransferase CheR (cheR-2) |
| SO3252 | | | | | | 2 | | | chemotaxis protein CheV (cheV-3) |
| SO3254 | | | 4 | 4 | | 4 | 2 | | negative regulator of flagellin synthesis FlgM (flgM) |
| SO3255 | | | 2 | 4 | | | 2 | 2 | flagellar biosynthetic protein FlgN (flgN) |
| SO3256 | 3 | 3 | 2 | | | 2 | 2 | 2 | conserved hypothetical protein |
| SO3257 | 2 | | 2 | 3 | 4 | | 2 | | conserved hypothetical protein |
| SO3259 | | | 2 | | 3 | 3 | | 2 | conserved domain protein |
| SO3261 | 2 | | | | | | | | polysaccharide biosynthesis related-protein |
| SO3262 | 4 | 6 | 9 | 4 | 9 | 2 | 4 | 6 | acetolactate synthase isozyme I, large subunit (ilvB) |
| SO3263 | | | | | 3 | | | 2 | 3-oxoacyl-(acyl-carrier-protein) reductase, putative |
| SO3264 | | | | | | | | 2 | hypothetical protein |
| SO3265 | 5 | 6 | 7 | 6 | 3 | | | 4 | conserved hypothetical protein |
| SO3266 | | | 3 | 2 | | | 2 | 3 | conserved domain protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3267 | 5 | 2 | 4 | | 5 | | 3 | 2 | conserved domain protein |
| SO3268 | | | | 2 | | | | | alpha amylase domain protein |
| SO3271 | 10 | 11 | 9 | 6 | 13 | 9 | 7 | 8 | polysaccharide biosynthesis protein |
| SO3273 | | | | | | | 2 | | conserved hypothetical protein |
| SO3275 | | 3 | 2 | 2 | 2 | | 2 | 2 | hypothetical protein |
| SO3279 | | | | | 2 | | | 2 | AcrB/AcrD/AcrF family protein |
| SO3282 | 8 | 12 | 3 | 6 | 18 | 20 | 6 | 11 | methyl-accepting chemotaxis protein |
| SO3285 | | 2 | | | 3 | | | | cytochrome d ubiquinol oxidase, subunit II (cydB) |
| SO3286 | 19 | 19 | 16 | 19 | 26 | 21 | 14 | 17 | cytochrome d ubiquinol oxidase, subunit I (cydA) |
| SO3287 | 42 | 57 | 38 | 47 | 52 | 26 | 33 | 49 | phosphoribosylformylglycinamidine synthase (purL) |
| SO3288 | 2 | 3 | 2 | | 4 | 3 | 6 | 2 | transglycosylase, Slt family |
| SO3291 | 2 | | | | | | | | cytidine/deoxycytidylate deaminase family protein |
| SO3292 | 11 | 17 | 16 | 13 | 19 | 4 | 16 | 16 | GMP synthase (guaA) |
| SO3293 | 51 | 62 | 50 | 49 | 48 | 55 | 47 | 47 | inosine-5-monophosphate dehydrogenase (guaB) |
| SO3294 | 3 | | 3 | | 2 | | | | exodeoxyribonuclease VII, large subunit (xseA) |
| SO3297 | | 2 | | | 2 | 2 | | | transcriptional regulator, LysR family |
| SO3298 | | | | | | | | 2 | conserved hypothetical protein |
| SO3299 | 2 | 5 | 5 | 8 | 6 | 3 | 3 | 8 | Pal/histidase family protein |
| SO3306 | | | | | | | 2 | | sensor histidine kinase |
| SO3308 | 19 | 23 | 17 | 14 | 20 | 14 | 14 | 18 | GTP-binding protein EngA (engA) |
| SO3309 | 14 | 13 | 11 | 14 | 13 | 12 | 12 | 13 | PQQ enzyme repeat domain protein |
| SO3310 | 33 | 42 | 32 | 37 | 35 | 35 | 33 | 36 | conserved hypothetical protein |
| SO3311 | 11 | 12 | 12 | 11 | 16 | 11 | 4 | 16 | histidyl-tRNA synthetase (hisS) |
| SO3312 | 9 | 7 | 4 | 7 | 5 | 6 | 7 | 9 | 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (ispG) |
| SO3313 | 13 | 12 | 11 | 11 | 12 | 11 | 13 | 14 | conserved hypothetical protein |
| SO3314 | 13 | 7 | 11 | 11 | 8 | 5 | 8 | 10 | fimbrial biogenesis and twitching motility protein, putative |
| SO3315 | 4 | 4 | 3 | 3 | 5 | | 3 | 5 | conserved hypothetical protein TIGR00048 |
| SO3316 | 13 | 20 | 16 | 21 | 14 | 11 | 11 | 13 | conserved hypothetical protein |
| SO3317 | 2 | 2 | 2 | 5 | 3 | | 4 | 5 | 5-nucleotidase, putative |
| SO3325 | 3 | 2 | 2 | 4 | 5 | 5 | 2 | 4 | nrfJ-related protein |
| SO3335 | | | | 2 | | 2 | | 2 | conserved hypothetical protein |
| SO3338 | 7 | 2 | 6 | 6 | 4 | 7 | 2 | 7 | L-allo-threonine aldolase (ybjU) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3339 | | 2 | 3 | 5 | | | | | OmpA family protein |
| SO3340 | 5 | 6 | 7 | 5 | 10 | | 8 | 11 | conserved hypothetical protein |
| SO3341 | 9 | 12 | 13 | 17 | 9 | 5 | 8 | 10 | antioxidant, AhpC/TSA family |
| SO3343 | 17 | 19 | 19 | 21 | 17 | 10 | 15 | 15 | conserved hypothetical protein |
| SO3345 | 3 | 4 | 6 | 4 | 2 | 4 | 3 | 4 | translation initiation factor, putative |
| SO3346 | | 2 | 3 | 2 | 3 | | 4 | 6 | conserved hypothetical protein |
| SO3347 | 2 | 3 | 3 | 2 | | | | | conserved hypothetical protein TIGR00250 |
| SO3348 | 8 | 9 | 7 | 6 | 4 | 4 | 9 | 7 | ferrochelatase (hemH-2) |
| SO3349 | | | 2 | | | | | 2 | glutathione peroxidase, putative |
| SO3350 | | 3 | | | 6 | | 4 | 3 | twitching motility protein PilU (pilU) |
| SO3351 | 15 | 7 | 9 | 6 | 11 | 3 | 18 | 14 | twitching motility protein PilT (pilT) |
| SO3352 | 3 | 2 | | 2 | | 2 | 2 | | conserved hypothetical protein TIGR00044 |
| SO3354 | 8 | 5 | 5 | 7 | 3 | 5 | 3 | 3 | pyrroline-5-carboxylate reductase (proC) |
| SO3355 | | | | | | | 2 | | conserved hypothetical protein |
| SO3358 | 5 | 3 | 2 | 3 | | 2 | | 4 | HAM1 protein |
| SO3360 | 4 | 3 | 3 | | 2 | 4 | 3 | 2 | ISSo5, transposase |
| SO3361 | 9 | 9 | 11 | 11 | 6 | 7 | | 4 | conserved hypothetical protein |
| SO3363 | 4 | 2 | 3 | 5 | 5 | 2 | 4 | 4 | transcriptional regulator, LysR family |
| SO3364 | | | | | 2 | | | | conserved hypothetical protein |
| SO3365 | | | | | 2 | | 3 | | glutaminase A (glsA) |
| SO3367 | 7 | 9 | 7 | 9 | 4 | 4 | 9 | 12 | conserved hypothetical protein TIGR00091 |
| SO3368 | 2 | 5 | 2 | 3 | 3 | | 4 | 3 | A/G-specific adenine glycosylase (mutY) |
| SO3369 | 5 | 7 | 7 | 8 | 6 | 4 | 5 | 9 | conserved hypothetical protein |
| SO3370 | 2 | 4 | 4 | 6 | | | 4 | 5 | conserved hypothetical protein |
| SO3374 | 2 | 2 | | 2 | 3 | 3 | 3 | 3 | conserved hypothetical protein |
| SO3379 | | | 3 | 5 | | | | 2 | cyclopropane-fatty-acyl-phospholipid synthase (cfa) |
| SO3380 | | | 2 | | | | | 2 | conserved hypothetical protein |
| SO3381 | | | | | 2 | 2 | 2 | 2 | conserved hypothetical protein |
| SO3384 | | | 3 | 2 | 2 | | 2 | | deoxyribodipyrimidine photolyase (phrB) |
| SO3388 | 5 | 7 | 4 | 5 | 7 | 2 | 5 | 6 | ATP-dependent RNA helicase, DEAD box family |
| SO3389 | | | | 2 | | | | | sensory box protein |
| SO3391 | 7 | 10 | 5 | 7 | 7 | 5 | 7 | 7 | ATP-dependent protease, putative |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3392 | 2 | 4 | 3 | 3 | 3 | | 4 | 6 | oxidoreductase, FMN-binding |
| SO3396 | 4 | 4 | 3 | 6 | 5 | 4 | 4 | 5 | methyl-accepting chemotaxis protein, truncation |
| SO3399 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO3401 | 10 | 6 | 7 | 11 | 7 | 8 | 4 | 4 | conserved hypothetical protein TIGR01033 |
| SO3403 | 4 | 8 | 11 | 10 | 10 | 10 | 12 | 14 | ribosomal subunit interface protein (yfiA-1) |
| SO3404 | 3 | 3 | 3 | 2 | 3 | 5 | 2 | 3 | methyl-accepting chemotaxis protein |
| SO3407 | 5 | 8 | 8 | 10 | 4 | 3 | 13 | 14 | conserved hypothetical protein |
| SO3409 | 2 | 4 | 2 | | 3 | 3 | 4 | 3 | OsmC/Ohr family protein |
| SO3411 | 26 | 21 | 18 | 23 | 19 | 22 | 20 | 25 | protease, putative |
| SO3413 | 3 | | 2 | | | | | | threonine synthase (thrC) |
| SO3415 | | | 2 | 3 | 2 | | | 5 | aspartokinase I/homoserine dehydrogenase, threonine-sensitive (thrA) |
| SO3417 | 6 | 6 | 7 | 5 | 5 | 6 | 8 | 5 | peptidyl-prolyl cis-trans isomerase SlyD (slyD) |
| SO3420 | 11 | 16 | 19 | 21 | 13 | 10 | 15 | 17 | cytochrome c |
| SO3422 | 4 | 4 | 8 | 5 | 5 | 4 | 5 | 5 | ribosomal subunit interface protein (yfiA-2) |
| SO3424 | 23 | 32 | 24 | 34 | 27 | 17 | 27 | 40 | valyl-tRNA synthetase (valS) |
| SO3426 | 4 | 3 | 3 | 5 | 4 | 4 | 6 | 3 | carbon storage regulator (csrA) |
| SO3427 | 3 | 2 | | | 3 | 3 | 2 | | aspartokinase |
| SO3428 | 48 | 41 | 34 | 43 | 42 | 36 | 33 | 41 | alanyl-tRNA synthetase (alaS) |
| SO3430 | 40 | 41 | 47 | 47 | 38 | 42 | 51 | 52 | recA protein (recA) |
| SO3431 | 13 | 11 | 12 | 14 | 20 | 7 | 12 | 13 | DNA mismatch repair protein MutS (mutS) |
| SO3432 | | | | | 3 | | | | RNA polymerase sigma-38 factor (rpoS) |
| SO3433 | 2 | 2 | 4 | | 5 | 7 | 6 | 5 | lipoprotein NlpD (nlpD) |
| SO3434 | 3 | 3 | | 2 | 2 | | | 2 | protein-L-isoaspartate O-methyltransferase (pcm) |
| SO3435 | 2 | | | | | | | | stationary-phase survival protein SurE (surE) |
| SO3436 | 7 | 5 | 5 | 7 | 6 | 4 | 11 | 9 | conserved hypothetical protein |
| SO3437 | | 2 | | | | | 3 | | 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (ispF) |
| SO3438 | 2 | 7 | 5 | 5 | 5 | 2 | 8 | 7 | 4-diphosphocytidyl-2C-methyl-D-erythritol synthase (ispD) |
| SO3439 | 2 | 3 | 2 | 2 | 3 | | 2 | 4 | conserved hypothetical protein |
| SO3440 | 24 | 29 | 36 | 37 | 22 | 29 | 34 | 32 | enolase (eno) |
| SO3441 | 37 | 43 | 29 | 35 | 33 | 23 | 26 | 34 | CTP synthase (pyrG) |
| SO3442 | 3 | | 2 | 6 | | 3 | 4 | 2 | MazG family protein |
| SO3443 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3444 | | | | | | | 2 | | hypothetical protein |
| SO3449 | 2 | 6 | 2 | 4 | 4 | | 2 | 4 | conserved domain protein |
| SO3451 | 15 | 22 | 15 | 15 | 16 | 18 | 15 | 15 | ISSo4, transposase |
| SO3455 | 16 | 15 | 21 | 19 | 28 | 20 | 21 | 26 | GTP pyrophosphokinase (relA) |
| SO3456 | 9 | 11 | 3 | 5 | 7 | 4 | 2 | 6 | RNA methyltransferase, TrmA family |
| SO3457 | 6 | 6 | 5 | 5 | 7 | 7 | 5 | 8 | sensor histidine kinase/response regulator |
| SO3458 | | | | | 3 | | 2 | 2 | conserved hypothetical protein |
| SO3462 | 4 | | 6 | 12 | | 2 | 12 | 15 | DNA repair protein RecN (recN) |
| SO3463 | 2 | | | | | | | | phosphatidylglycerophosphatase A (pgpA) |
| SO3464 | 2 | 3 | 2 | 4 | | 2 | 7 | 6 | thiamin-monophosphate kinase (thiL) |
| SO3465 | 3 | 8 | 8 | 8 | 7 | 7 | 9 | 9 | N utilization substance protein B (nusB) |
| SO3466 | 19 | 23 | 18 | 18 | 21 | 22 | 19 | 20 | riboflavin synthase, beta subunit (ribH) |
| SO3467 | 19 | 20 | 14 | 19 | 19 | 15 | 15 | 12 | 3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II, putative (ribBA) |
| SO3468 | | 2 | 3 | 2 | 5 | 3 | 3 | 3 | riboflavin synthase, alpha subunit (ribE-2) |
| SO3469 | 2 | 4 | 2 | 2 | 2 | | | 4 | riboflavin biosynthesis protein RibD (ribD) |
| SO3470 | 4 | 4 | 6 | 7 | 2 | | | 4 | conserved hypothetical protein TIGR00244 |
| SO3471 | 26 | 32 | 31 | 28 | 29 | 20 | 22 | 28 | serine hydroxymethyltransferase (glyA) |
| SO3472 | 41 | 32 | 39 | 46 | 30 | 32 | 31 | 33 | ABC transporter, ATP-binding protein |
| SO3480 | | | | 2 | | | | | conserved hypothetical protein |
| SO3483 | 13 | 9 | 13 | 17 | 8 | 9 | 7 | 13 | HlyD family secretion protein |
| SO3484 | 10 | 11 | 12 | 6 | 9 | 8 | 8 | 9 | AcrB/AcrD/AcrF family protein |
| SO3494 | 2 | 5 | 4 | 2 | | | | 2 | transcriptional regulator, TetR family |
| SO3496 | | 3 | | 4 | 4 | 2 | 3 | 2 | aldehyde dehydrogenase |
| SO3506 | | | 2 | 4 | | | 2 | 2 | SIS domain protein |
| SO3509 | 11 | 14 | 14 | 13 | 18 | 12 | 16 | 18 | beta-hexosaminidase b precursor (hex) |
| SO3516 | 6 | 2 | 4 | 4 | 4 | 4 | | 4 | transcriptional regulator, LacI family |
| SO3517 | 13 | 12 | 12 | 9 | 12 | 14 | 9 | 8 | NADH dehydrogenase (ndh) |
| SO3518 | | 3 | 2 | 2 | | 2 | | | ISSo4, transposase |
| SO3519 | 6 | | | 4 | | 3 | 4 | 2 | nitrogen regulatory protein P-II 1 (glnB-2) |
| SO3521 | | | 2 | 2 | 2 | | 2 | | type IV pilus biogenesis protein, putative |
| SO3524 | 2 | | | 4 | | | 2 | 2 | type IV pilus biogenesis protein PilE (pilE) |
| SO3529 | 9 | 10 | 7 | 9 | 5 | 3 | 7 | 10 | penicillin tolerance protein LytB (lytB) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| **Locus** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Description** |
| SO3532 | 17 | 27 | 23 | 29 | 27 | 13 | 24 | 35 | isoleucyl-tRNA synthetase (ileS) |
| SO3533 | | | 2 | 3 | | | 4 | | riboflavin biosynthesis protein RibF (ribF) |
| SO3537 | 7 | 6 | 7 | 8 | 6 | 7 | 6 | 6 | ribosomal protein S20 (rpsT) |
| SO3538 | 5 | 3 | | | 6 | 3 | | | transcriptional regulator HlyU (hlyU) |
| SO3539 | 23 | 30 | 25 | 24 | 28 | 23 | 25 | 26 | peptidase, M28D family |
| SO3540 | 2 | | 2 | | | | | | conserved hypothetical protein |
| SO3541 | 4 | 2 | 3 | 3 | 3 | 2 | 3 | 3 | sodium:alanine symporter family protein |
| SO3542 | 2 | 5 | | 2 | 6 | | | | conserved hypothetical protein |
| SO3543 | | | | | | | | 2 | ISSo13, transposase |
| SO3544 | | | 2 | 3 | | | | 3 | ISSo7, transposase |
| SO3545 | 48 | 56 | 50 | 52 | 51 | 48 | 59 | 59 | OmpA family protein |
| SO3546 | 11 | 11 | 10 | 9 | 10 | 9 | 7 | 7 | transaldolase (tal) |
| SO3547 | 10 | 7 | 6 | 6 | 6 | 6 | 2 | 5 | glucose-6-phosphate isomerase (pgi) |
| SO3550 | 2 | 4 | 3 | 3 | 3 | | 7 | 8 | hypothetical protein |
| SO3551 | | | 2 | | 2 | | | 3 | RNA polymerase sigma-70 factor, ECF subfamily |
| SO3552 | | | | | | | | 2 | von Willebrand factor type A domain protein |
| SO3553 | 2 | 4 | 3 | | 5 | 2 | | | sulfate permease family protein |
| SO3554 | 9 | 13 | 6 | 6 | 6 | 4 | 6 | 6 | phosphoribosylaminoimidazole carboxylase, catalytic subunit (purE) |
| SO3555 | 3 | 5 | 3 | 6 | 3 | 2 | 2 | | phosphoribosylaminoimidazole carboxylase, ATPase subunit (purK) |
| SO3556 | 4 | 5 | 2 | 3 | 8 | 6 | 6 | 3 | GAF/GGDEF domain protein |
| SO3557 | | | 2 | 7 | 4 | 2 | 4 | 3 | conserved hypothetical protein |
| SO3558 | | | 2 | | | | | 2 | hypothetical protein |
| SO3559 | 5 | 6 | 3 | 6 | 4 | | | 3 | glutamate--cysteine ligase (gshA) |
| SO3560 | 62 | 66 | 63 | 58 | 74 | 56 | 61 | 59 | peptidase, M16 family |
| SO3562 | 2 | 5 | 5 | 4 | 2 | 3 | 2 | 3 | proton/glutamate symporter, putative |
| SO3564 | 54 | 64 | 46 | 54 | 58 | 57 | 46 | 57 | peptidyl-dipeptidase Dcp (dcp-2) |
| SO3565 | 34 | 47 | 29 | 31 | 44 | 44 | 22 | 21 | 2,3-cyclic-nucleotide 2-phosphodiesterase (cpdB) |
| SO3571 | 4 | 5 | 4 | 3 | | 3 | 3 | 3 | conserved hypothetical protein |
| SO3575 | | 3 | | | 2 | 2 | 2 | 2 | CDP-diacylglycerol--serine O-phosphatidyltransferase, putative |
| SO3577 | 30 | 37 | 55 | 71 | 39 | 41 | 66 | 67 | clpB protein (clpB) |
| SO3578 | | | | | | | 3 | 2 | conserved hypothetical protein TIGR00726 |
| SO3579 | 4 | 7 | 5 | 4 | 3 | 4 | 3 | 6 | ribosomal large subunit pseudouridine synthase D (rluD) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3580 | 14 | 21 | 13 | 16 | 12 | 15 | 12 | 12 | conserved hypothetical protein |
| SO3582 | 8 | 12 | 8 | 7 | 8 | 11 | 9 | 12 | methyl-accepting chemotaxis protein |
| SO3584 | | 2 | | 2 | | | | 5 | conserved hypothetical protein |
| SO3585 | | | 10 | 9 | | | 5 | 4 | Azoreductase, putative |
| SO3586 | | | 5 | 6 | | | 8 | 3 | glyoxalase family protein |
| SO3587 | | 2 | 5 | 7 | 3 | 2 | 4 | 3 | hypothetical protein |
| SO3594 | 2 | 5 | 7 | 6 | 6 | 2 | 6 | 7 | transcriptional regulatory protein RstA, putative |
| SO3595 | 2 | 3 | 5 | 3 | 4 | 5 | 3 | 4 | sensor protein RstB, putative |
| SO3597 | 4 | 10 | 8 | 7 | 5 | 6 | 7 | 8 | conserved hypothetical protein |
| SO3598 | 6 | 6 | 4 | 5 | 3 | | 6 | 6 | cysteine synthase B (cysM) |
| SO3599 | 11 | 20 | 28 | 32 | 2 | 5 | 23 | 23 | sulfate ABC transporter, periplasmic sulfate-binding protein (cysP) |
| SO3602 | 11 | 25 | 21 | 14 | 11 | 13 | 17 | 23 | sulfate ABC transporter, ATP-binding protein (cysA-1) |
| SO3611 | | 2 | | | | | | 2 | ATPase, AAA family |
| SO3613 | 2 | 3 | 3 | | 3 | 4 | | | phosphoribosylglycinamide formyltransferase 2 (purT) |
| SO3614 | | | | | | | | 2 | hypothetical protein |
| SO3615 | 16 | 22 | 12 | 14 | 19 | 15 | 12 | 15 | hypothetical protein |
| SO3622 | 3 | 2 | 5 | 8 | | 3 | 2 | 5 | conserved domain protein |
| SO3631 | 3 | 4 | 3 | 4 | 2 | 3 | 6 | | glycerate dehydrogenase (hprA) |
| SO3633 | 7 | 6 | 4 | 4 | 7 | 3 | 6 | 8 | DnaJ domain protein |
| SO3636 | 10 | 16 | 8 | 12 | 9 | 8 | 13 | 15 | organic solvent tolerance protein (imp) |
| SO3637 | 19 | 17 | 17 | 22 | 17 | 13 | 20 | 19 | survival protein surA (surA) |
| SO3638 | | | 2 | 2 | 3 | 2 | | 3 | pyridoxal phosphate biosynthetic protein PdxA (pdxA) |
| SO3639 | 3 | 4 | 2 | 4 | 6 | 2 | 2 | 6 | dimethyladenosine transferase (ksgA) |
| SO3642 | 20 | 36 | 21 | 25 | 27 | 31 | 23 | 25 | methyl-accepting chemotaxis protein |
| SO3646 | 3 | 4 | | 2 | 2 | | 2 | 2 | dihydrofolate reductase (folA) |
| SO3649 | 14 | 13 | 16 | 11 | 16 | 10 | 12 | 17 | GTP-binding protein, GTP1/Obg family |
| SO3651 | 8 | 10 | 11 | 11 | 7 | 9 | 11 | 11 | ribosomal protein L27 (rpmA) |
| SO3652 | 16 | 18 | 13 | 14 | 16 | 17 | 14 | 17 | ribosomal protein L21 (rplU) |
| SO3653 | 3 | 8 | 4 | 8 | 4 | 2 | | 5 | octaprenyl-diphosphate synthase (ispB) |
| SO3654 | | 2 | | | | | | 2 | uracil-DNA glycosylase (ung) |
| SO3655 | | 2 | | 2 | | 2 | | | pilin, putative |
| SO3656 | 3 | 2 | 2 | 3 | 4 | 3 | | | hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3664 | 13 | 13 | 6 | 10 | 7 | 4 | 4 | 6 | long-chain-fatty-acid--CoA ligase (fadD-2) |
| SO3665 | 4 | 3 | 5 | 4 | 2 | 4 | 5 | 6 | ABC transporter, ATP-binding/permease protein, putative |
| SO3667 | 4 | 10 | 30 | 27 | | | 29 | 31 | conserved hypothetical protein |
| SO3668 | | | 5 | 6 | | | | 3 | conserved hypothetical protein |
| SO3669 | 14 | 14 | 53 | 64 | 3 | 5 | 67 | 71 | heme transport protein (hugA) |
| SO3670 | | | 3 | 4 | | | 3 | 4 | TonB1 protein (tonB1) |
| SO3671 | 3 | 4 | 5 | 4 | | | 6 | 7 | TonB system transport protein ExbB1 (exbB1) |
| SO3673 | 2 | 6 | 22 | 22 | | | 16 | 17 | hemin ABC transporter, periplasmic hemin-binding protein (hmuT) |
| SO3675 | 4 | 3 | 13 | 9 | | | 12 | 9 | hemin ABC transporter, ATP-binding protein (hmuV) |
| SO3676 | | 3 | 2 | 4 | 3 | | 2 | 4 | hypothetical protein |
| SO3678 | 2 | 3 | | | 2 | 3 | | | conserved hypothetical protein |
| SO3680 | 2 | 2 | 2 | | | 2 | 3 | | universal stress protein family |
| SO3681 | 3 | 5 | 5 | 6 | 6 | 5 | 8 | 7 | universal stress protein family |
| SO3683 | 13 | 15 | 17 | 14 | 13 | 13 | 19 | 10 | coniferyl aldehyde dehydrogenase |
| SO3689 | | | 2 | | | | | 2 | sigma-54 dependent nitrogen response regulator |
| SO3692 | | | 2 | 2 | | | 3 | | ABC transporter, ATP-binding protein |
| SO3695 | 7 | 12 | 7 | 6 | 8 | 8 | 9 | 6 | dihydroorotase, homodimeric type (pyrC) |
| SO3696 | | 2 | 2 | | | | | | hypothetical protein |
| SO3698 | | | | | | | | 3 | hypothetical protein |
| SO3705 | 3 | 2 | 4 | 4 | 3 | | 7 | 4 | 5-methylthioadenosine nucleosidase/S-adenosylhomocysteine nucleosidase, putative |
| SO3711 | | | | | 3 | | | | conserved hypothetical protein |
| SO3714 | | | | 2 | | | | | sugar-binding protein, putative |
| SO3715 | 8 | 8 | 14 | 12 | 9 | 5 | 11 | 13 | oxygen-insensitive NAD(P)H nitroreductase |
| SO3718 | 7 | 9 | 8 | 7 | 9 | 8 | 5 | 10 | thiol:disulfide interchange protein, DsbA family |
| SO3720 | 6 | 5 | 6 | 4 | 12 | 7 | 6 | 4 | conserved hypothetical protein |
| SO3722 | 3 | 5 | | | 5 | 4 | | 2 | conserved hypothetical protein |
| SO3723 | | | | 3 | | | 5 | 4 | adenylylsulfate kinase (cysC) |
| SO3726 | 5 | 11 | 17 | 19 | 7 | 2 | 17 | 23 | sulfate adenylyltransferase, subunit 1 (cysN) |
| SO3727 | 14 | 15 | 24 | 23 | 10 | 5 | 28 | 32 | sulfate adenylyltransferase, subunit 2 (cysD) |
| SO3728 | | | | 2 | | | 4 | 4 | uroporphyrin-III C-methyltransferase (cobA) |
| SO3731 | | | | | | 2 | | | hypothetical protein |
| SO3733 | 28 | 34 | 30 | 30 | 38 | 24 | 37 | 30 | hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3735 | 2 | | | 2 | | | | 2 | monofunctional biosynthetic peptidoglycan transglycosylase (mtgA) |
| SO3736 | 4 | 3 | 4 | 4 | 2 | | 5 | 5 | phosphoadenosine phosphosulfate reductase (cysH) |
| SO3737 | 19 | 17 | 35 | 39 | 13 | 8 | 41 | 57 | sulfite reductase (NADPH) hemoprotein beta-component (cysI) (cysI) |
| SO3738 | 2 | 5 | 11 | 9 | 2 | | 13 | 15 | sulfite reductase (NADPH) flavoprotein alpha-component (cysJ) |
| SO3740 | 35 | 35 | 33 | 36 | 39 | 34 | 43 | 43 | NAD(P) transhydrogenase, alpha subunit (pntA) |
| SO3741 | 13 | 15 | 14 | 14 | 16 | 8 | 16 | 15 | NAD(P) transhydrogenase, beta subunit (pntB) |
| SO3743 | 4 | 8 | 7 | 6 | 5 | 3 | 4 | 7 | transcriptional regulator, TetR family |
| SO3745 | 7 | 9 | 4 | 6 | 7 | 4 | 4 | 8 | ADP-heptose synthase (rfaE) |
| SO3746 | 9 | 15 | 12 | 11 | 12 | 9 | 13 | 17 | lipid A biosynthesis lauroyl acyltransferase (htrB) |
| SO3747 | 2 | 3 | | 3 | | | 2 | 3 | sodium/hydrogen exchanger family/TrkA domain protein |
| SO3748 | 22 | 27 | 21 | 19 | 19 | 23 | 21 | 21 | LysM domain protein |
| SO3749 | | 2 | 3 | 3 | 2 | | 2 | 2 | hypothetical protein |
| SO3757 | | | | | | | | 2 | rubisco operon transcriptional regulator (rbcR) |
| SO3760 | 3 | 3 | 2 | | | | | | glutamate-ammonia-ligase adenylyltransferase (glnE) |
| SO3761 | 2 | 14 | 9 | 8 | 9 | 11 | 8 | 12 | hypothetical protein |
| SO3763 | 6 | 4 | 3 | 2 | 4 | 3 | 3 | 5 | spermine/spermidine synthase family protein |
| SO3765 | 32 | 40 | 37 | 45 | 37 | 41 | 43 | 49 | conserved hypothetical protein |
| SO3767 | | | 2 | | 2 | | | | hypothetical protein |
| SO3768 | | | 2 | 2 | | | 2 | | ion transporter |
| SO3769 | | | | | | | | 2 | conserved hypothetical protein |
| SO3770 | 17 | 15 | 14 | 15 | 11 | 8 | 10 | 16 | conserved hypothetical protein TIGR00153 |
| SO3772 | 16 | 26 | 19 | 23 | 18 | 20 | 21 | 21 | conserved hypothetical protein |
| SO3774 | 71 | 75 | 40 | 57 | 65 | 62 | 47 | 46 | proline dehydrogenase/delta-1-pyrroline-5-carboxylate dehydrogenase, putative |
| SO3775 | 5 | | 3 | 3 | 5 | 5 | 5 | 9 | hypothetical protein |
| SO3776 | | 2 | | | | 2 | | 2 | conserved hypothetical protein |
| SO3777 | | | | | | | | 2 | hypothetical protein |
| SO3778 | | | | | | | 2 | | adenylate cyclase CyaB, putative |
| SO3779 | 3 | 2 | 2 | 4 | 8 | 3 | 6 | 6 | ABC transporter, ATP-binding protein CydC (cydC) |
| SO3780 | 4 | 8 | 4 | | 7 | 7 | | 3 | ABC transporter, ATP-binding protein CydD (cydD) |
| SO3781 | 4 | 5 | 7 | 5 | 5 | 5 | 5 | 4 | hypothetical protein |
| SO3783 | 9 | 11 | 2 | 5 | 5 | 6 | | 2 | ATP-dependent RNA helicase, DEAD box family |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3786 | 2 | 3 | 4 | 5 | 3 | 2 | 4 | 7 | hypothetical protein |
| SO3787 | 3 | 3 | 4 | 5 | | 3 | 4 | 4 | hypothetical protein |
| SO3789 | | 2 | 2 | | | | 2 | | aminotransferase, class V |
| SO3790 | | | | 2 | 3 | | 4 | | conserved hypothetical protein |
| SO3791 | | | 2 | | | | 2 | 3 | renal dipeptidase family protein |
| SO3796 | | | | | 3 | | | | conserved hypothetical protein |
| SO3797 | 9 | 8 | 3 | 8 | 4 | 6 | 5 | 4 | peptidase, U32 family |
| SO3798 | | 3 | 2 | 2 | 2 | | 2 | | ribosomal large subunit pseudouridine synthase A (rluA-2) |
| SO3799 | 4 | 4 | 4 | 3 | 5 | | 2 | 5 | regulatory protein AsnC (asnC) |
| SO3800 | 7 | 5 | 3 | | 6 | 5 | | | serine protease, subtilase family |
| SO3802 | 8 | 9 | 13 | 18 | 10 | 7 | 13 | 12 | ABC transporter, ATP-binding protein |
| SO3803 | 5 | 9 | 5 | 10 | 9 | 3 | 8 | 10 | hypoxanthine phosphoribosyltransferase (hpt-2) |
| SO3804 | | | 3 | 3 | 2 | | 4 | 2 | phenylacrylic acid decarboxylase, 3-octaprenyl-4-hydroxybenzoate carboxy-lyase, putative |
| SO3805 | 9 | 12 | 7 | 10 | 6 | 5 | 6 | 10 | UDP-N-acetylmuramate:L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase (mpl) |
| SO3810 | 2 | | | | | | 2 | 3 | OmpA-like transmembrane domain protein |
| SO3811 | 16 | 16 | 14 | 12 | 17 | 10 | 17 | 19 | lipoprotein, putative |
| SO3812 | 4 | 7 | 5 | 5 | 4 | 3 | 3 | 4 | peptidyl-prolyl cis-trans isomerase A (ppiA) |
| SO3815 | 9 | 8 | 11 | 11 | 10 | 8 | 5 | 11 | conserved hypothetical protein |
| SO3817 | | | 2 | | | | | | 2-dehydropantoate 2-reductase (panE) |
| SO3821 | 2 | 2 | | | | | | | rtn protein |
| SO3827 | 9 | 12 | 10 | 11 | 7 | 9 | 10 | 9 | 2-dehydro-3-deoxyphosphooctonate aldolase (kdsA) |
| SO3828 | | | | | 2 | | | | conserved hypothetical protein |
| SO3832 | 3 | | 2 | 3 | 3 | | 2 | 3 | hemK family protein |
| SO3833 | 8 | 15 | 10 | 12 | 7 | 7 | 11 | 11 | peptide chain release factor 1 (prfA) |
| SO3834 | 12 | 15 | 8 | 6 | 16 | 14 | 10 | 5 | glutamyl-tRNA reductase (hemA) |
| SO3835 | 6 | 4 | 6 | 5 | 3 | 4 | 5 | 4 | outer membrane lipoprotein LolB (lolB) |
| SO3836 | 3 | 7 | 7 | 8 | 5 | | | 4 | 4-diphosphocytidyl-2C-methyl-D-erythritol kinase (ispE) |
| SO3837 | 29 | 31 | 26 | 26 | 27 | 29 | 17 | 25 | ribose-phosphate pyrophosphokinase (prsA) |
| SO3838 | 16 | 15 | 14 | 21 | 17 | 13 | 14 | 18 | methyl-accepting chemotaxis protein |
| SO3841 | | | 2 | | | | | | hypothetical protein |
| SO3842 | 12 | 12 | 11 | 17 | 8 | 8 | 11 | 9 | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3843 | | 4 | 3 | 4 | | | | 2 | ribosomal small subunit pseudouridine synthase A (rsuA-2) |
| SO3844 | 89 | 112 | 90 | 81 | 91 | 97 | 80 | 93 | peptidase, M13 family |
| SO3848 | 2 | 4 | 2 | | | 2 | | 2 | hypothetical protein |
| SO3854 | | | | | 2 | | 2 | 2 | ISSo12, transposase |
| SO3855 | 21 | 30 | 21 | 23 | 25 | 16 | 24 | 25 | malate oxidoreductase (sfcA) |
| SO3856 | 8 | 12 | 6 | 5 | 6 | 9 | 4 | 6 | hypothetical protein |
| SO3862 | | | | | | | | 2 | molybdenum transport regulatory protein ModE (modE) |
| SO3863 | 12 | 14 | 12 | 15 | 10 | 19 | 6 | 9 | molybdenum ABC transporter, periplasmic molybdenum-binding protein (modA) |
| SO3865 | 17 | 16 | 9 | 13 | 25 | 17 | 13 | 10 | molybdenum ABC transporter, ATP-binding protein (modC) |
| SO3866 | 2 | | | 3 | | 2 | 2 | | site-specific recombinase, phage integrase family |
| SO3867 | 2 | | | | | | | 2 | transcriptional regulator, Cro/CI family |
| SO3872 | | | | | | | 2 | | arylsulfate sulfotransferase |
| SO3878 | 4 | 3 | 3 | | 2 | 4 | 3 | 2 | ISSo5, transposase |
| SO3880 | | | | | | | | 2 | ISSo13, transposase |
| SO3888 | 2 | | | | | | | | conserved hypothetical protein |
| SO3890 | 4 | 3 | 5 | 5 | 6 | 3 | 5 | 8 | methyl-accepting chemotaxis protein |
| SO3892 | | 2 | 3 | | 2 | 2 | 4 | | hypothetical protein |
| SO3895 | 4 | 4 | | 4 | 3 | 2 | | 3 | HesA/MoeB/ThiF family protein |
| SO3896 | 50 | 57 | 49 | 49 | 56 | 64 | 55 | 51 | outer membrane porin, putative |
| SO3897 | 13 | 14 | 7 | 8 | 11 | 10 | 7 | 14 | DNA topoisomerase IV, A subunit (parC) |
| SO3899 | 4 | 6 | 4 | 8 | 5 | 2 | 8 | 6 | DNA topoisomerase IV, B subunit (parE) |
| SO3901 | | | | | 3 | | | | lacZ expression regulator (icc) |
| SO3904 | 56 | 72 | 63 | 68 | 63 | 63 | 66 | 65 | outer membrane protein TolC (tolC) |
| SO3905 | 2 | 3 | | | 2 | | | 3 | conserved hypothetical protein |
| SO3906 | 9 | 13 | 18 | 17 | 13 | 14 | 18 | 17 | conserved hypothetical protein |
| SO3907 | 2 | 2 | 5 | 5 | | | 3 | 6 | conserved hypothetical protein |
| SO3908 | 2 | | 4 | | 3 | | 2 | 3 | enoyl-CoA hydratase/isomerase family protein |
| SO3912 | 2 | 2 | 5 | 7 | 2 | | | 6 | TIM-barrel protein, yjbN family |
| SO3913 | | | 5 | 4 | | | 4 | 4 | conserved hypothetical protein |
| SO3914 | 33 | 41 | 59 | 60 | 9 | 9 | 67 | 67 | TonB-dependent receptor, putative |
| SO3917 | 8 | 7 | 4 | 4 | 5 | 5 | 2 | 5 | replicative DNA helicase (dnaB) |
| SO3918 | 26 | 28 | 28 | 25 | 26 | 19 | 23 | 27 | conserved hypothetical protein |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO3920 | | | | | 8 | 7 | | | periplasmic Fe hydrogenase, large subunit (hydA) |
| SO3921 | | | | | 3 | 4 | | | periplasmic Fe hydrogenase, small subunit (hydB) |
| SO3927 | 30 | 33 | 32 | 29 | 28 | 44 | 25 | 26 | ribosomal protein L9 (rplI) |
| SO3928 | 16 | 19 | 16 | 17 | 18 | 17 | 13 | 17 | ribosomal protein S18 (rpsR) |
| SO3930 | 22 | 24 | 20 | 23 | 19 | 26 | 23 | 22 | ribosomal protein S6 (rpsF) |
| SO3931 | | 2 | 2 | | 3 | | 6 | 6 | conserved hypothetical protein |
| SO3934 | 14 | 15 | 8 | 11 | 10 | 11 | 12 | 11 | RNA methyltransferase, TrmH family, group 3 |
| SO3935 | 8 | 15 | 10 | 11 | 13 | 9 | 14 | 21 | ribonuclease R (vacB) |
| SO3936 | 3 | | 3 | 4 | 2 | | | | sodium-type flagellar protein MotX (motX) |
| SO3937 | 19 | 17 | 17 | 20 | 21 | 6 | 16 | 14 | adenylosuccinate synthetase (purA) |
| SO3939 | 10 | 10 | 12 | 13 | 10 | 6 | 9 | 11 | ribosomal protein S9 (rpsI) |
| SO3940 | 21 | 30 | 25 | 27 | 23 | 26 | 26 | 25 | ribosomal protein L13 (rplM) |
| SO3941 | | | | | | 2 | 2 | | conserved hypothetical protein |
| SO3942 | 34 | 37 | 43 | 51 | 43 | 40 | 44 | 47 | serine protease, HtrA/DegQ/DegS family |
| SO3943 | 5 | 13 | 6 | 6 | 9 | 13 | 8 | 10 | protease DegS (degS) |
| SO3948 | 7 | 5 | 8 | 11 | 5 | 7 | 8 | 10 | UDP-N-acetylglucosamine 1-carboxyvinyltransferase (murA) |
| SO3949 | 2 | 8 | 4 | 4 | 7 | | 3 | 4 | BolA/YrbA family protein |
| SO3950 | 7 | 3 | 4 | 4 | 4 | 5 | 6 | 3 | SpoIIAA family protein |
| SO3951 | | | | 2 | 3 | | 2 | 3 | conserved hypothetical protein |
| SO3952 | 4 | 6 | 4 | 4 | 2 | 2 | 4 | 5 | mce-related protein |
| SO3954 | 17 | 18 | 14 | 14 | 15 | 15 | 15 | 16 | ABC transporter, ATP-binding protein, putative |
| SO3956 | 10 | 9 | 8 | 14 | 6 | 9 | 15 | 7 | carbohydrate isomerase, KpsF/GutQ family |
| SO3957 | 2 | 4 | 2 | 5 | 2 | | 2 | 5 | conserved hypothetical protein |
| SO3958 | | | 3 | 2 | | | 2 | 2 | conserved hypothetical protein |
| SO3959 | 7 | 12 | 9 | 10 | 6 | 11 | 5 | 7 | conserved hypothetical protein |
| SO3960 | 14 | 23 | 13 | 16 | 19 | 20 | 15 | 19 | ABC transporter, ATP-binding protein |
| SO3961 | 11 | 8 | 7 | 11 | 8 | 8 | 7 | 14 | RNA polymerase sigma-54 factor (rpoN) |
| SO3962 | | | | | | | 2 | | ribosomal subunit interface protein (yfiA-3) |
| SO3963 | 2 | 5 | 3 | 3 | 2 | 4 | 2 | 3 | nitrogen regulatory IIA protein (ptsN) |
| SO3964 | | 3 | | | 2 | | | 2 | conserved hypothetical protein |
| SO3965 | | | | | | | 2 | | phosphocarrier protein NPR (ptsO) |
| SO3966 | 7 | 8 | 8 | 9 | 5 | 4 | 7 | 11 | magnesium transporter (mgtE-2) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
|---|---|---|---|---|---|---|---|---|---|
| SO3967 | 8 | 10 | 7 | 6 | 12 | 14 | 4 | 3 | molybdenum ABC transporter, periplasmic molybdenum-binding protein, putative |
| SO3969 | 11 | 18 | 17 | 19 | 16 | 11 | 15 | 19 | OmpA family protein |
| SO3973 | | 4 | 6 | 4 | 3 | 4 | 5 | 4 | RIO1/ZK632.3/MJ0444 family, putative |
| SO3974 | 4 | | | | | 2 | | 2 | conserved domain protein |
| SO3980 | 18 | 15 | 22 | 24 | 13 | 15 | 8 | 12 | cytochrome c552 nitrite reductase |
| SO3981 | | | | | 2 | | | | nitrate/nitrite sensor protein NarQ (narQ) |
| SO3982 | 3 | | 3 | | 4 | 2 | | | DNA-binding nitrate/nitrite response regulator |
| SO3984 | 6 | 5 | 3 | 2 | 4 | | 2 | 8 | magnesium transporter, putative |
| SO3988 | 23 | 26 | 27 | 28 | 28 | 27 | 30 | 26 | aerobic respiration control protein ArcA (arcA) |
| SO3990 | 5 | 6 | 6 | 2 | 8 | | 4 | 3 | dipeptidyl peptidase IV |
| SO3991 | 10 | 14 | 11 | 13 | 11 | 7 | 9 | 11 | fructose-1,6-bisphosphatase (fbp) |
| SO3994 | 8 | 7 | 11 | 12 | 7 | 4 | 12 | 12 | hypothetical protein |
| SO4002 | 2 | 6 | 3 | 5 | 7 | 6 | 7 | 6 | sensory transduction histidine kinase |
| SO4003 | 2 | 3 | 4 | 3 | 5 | 2 | 7 | 7 | response regulator |
| SO4006 | 3 | 3 | 3 | | 3 | 3 | 4 | 4 | hypothetical protein |
| SO4007 | 2 | 4 | 3 | 2 | | 3 | 2 | 2 | hypothetical protein |
| SO4008 | 39 | 58 | 34 | 51 | 43 | 49 | 37 | 45 | hypothetical protein |
| SO4011 | 2 | | 2 | | 3 | | | | conserved hypothetical protein |
| SO4012 | 2 | 2 | 3 | | 3 | 4 | 3 | 2 | hypothetical protein |
| SO4013 | 8 | 8 | 9 | 9 | 8 | 6 | 8 | 10 | hypothetical protein |
| SO4014 | 19 | 28 | 17 | 17 | 24 | 22 | 24 | 21 | AcrB/AcrD/AcrF family protein |
| SO4015 | 7 | 6 | 9 | 11 | 4 | 8 | 3 | 5 | conserved hypothetical protein |
| SO4016 | 2 | 3 | 2 | 2 | 3 | | | 2 | conserved hypothetical protein |
| SO4017 | 11 | 15 | 16 | 14 | 12 | 15 | 13 | 17 | transglycosylase, Slt family |
| SO4018 | 2 | | 2 | | | | | | hypothetical protein |
| SO4019 | | | 2 | | 5 | | | | hypothetical protein |
| SO4022 | 3 | | | 2 | 3 | 3 | | 2 | peptidase, M16 family |
| SO4028 | 9 | 12 | 5 | 9 | 8 | 9 | 4 | 4 | single-strand binding protein (ssb) |
| SO4029 | | | | | | | | 2 | transporter, putative |
| SO4030 | 15 | 19 | 22 | 19 | 23 | 7 | 10 | 20 | excinuclease ABC, A subunit (uvrA) |
| SO4034 | 58 | 55 | 39 | 42 | 48 | 42 | 52 | 48 | ATP-dependent RNA helicase DeaD (deaD) |
| SO4036 | 2 | 2 | 2 | | 5 | | 2 | 4 | hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO4038 | 2 | | 2 | | 5 | | 2 | 3 | hypothetical protein |
| SO4043 | 2 | 2 | | 3 | 3 | | | | TonB domain protein |
| SO4047 | 4 | 6 | 6 | 5 | 6 | 4 | | 2 | cytochrome c family protein |
| SO4048 | 2 | 3 | 3 | 4 | 4 | 3 | | | cytochrome c family protein |
| SO4052 | | | | | | 2 | | | transcriptional regulator, MarR family |
| SO4053 | 13 | 12 | 4 | 7 | 19 | 11 | 6 | 9 | methyl-accepting chemotaxis protein |
| SO4054 | | | 2 | | | | 2 | 4 | 5,10-methylenetetrahydrofolate reductase (metF) |
| SO4055 | | | 4 | 4 | 3 | | 3 | 2 | aspartokinase II/homoserine dehydrogenase, methionine-sensitive (metL) |
| SO4057 | 2 | 2 | 3 | 3 | 2 | | 4 | 4 | met repressor (metJ) |
| SO4062 | | | 2 | | | | | | polysulfide reductase, subunit A (psrA) |
| SO4066 | 21 | 22 | 16 | 15 | 15 | 11 | 10 | 12 | phosphoribosylaminoimidazole-succinocarboxamide synthase, putative |
| SO4068 | 9 | 14 | 18 | 16 | 17 | 15 | 14 | 18 | hypothetical protein |
| SO4070 | 4 | 3 | 6 | 7 | 4 | 5 | 6 | 7 | conserved hypothetical protein |
| SO4072 | 6 | 11 | 7 | 12 | 8 | 3 | 6 | 6 | MiaB-like putative RNA modifying enzyme YliG (yliG) |
| SO4077 | 8 | 7 | 15 | 14 | 5 | 5 | 19 | 21 | TonB-dependent receptor, putative |
| SO4078 | 6 | 4 | 5 | 4 | 9 | | | 5 | pmba protein (pmbA) |
| SO4079 | 4 | 10 | 7 | 15 | 4 | 6 | 6 | 16 | conserved hypothetical protein |
| SO4080 | 21 | 24 | 19 | 18 | 22 | 21 | 20 | 14 | conserved hypothetical protein |
| SO4085 | | 2 | | 2 | | | 3 | | chitinase A (chiA) |
| SO4089 | 10 | 14 | 14 | 13 | 12 | 11 | 15 | 18 | HlyD family secretion protein |
| SO4090 | 4 | 2 | 6 | 4 | 2 | 4 | 6 | 5 | outer membrane efflux protein |
| SO4091 | 6 | 9 | 9 | 6 | 3 | 7 | 6 | 5 | tldD protein (tldD) |
| SO4093 | 3 | 2 | 2 | 2 | 3 | 5 | 2 | 3 | conserved hypothetical protein |
| SO4094 | 7 | 4 | 6 | 6 | 5 | 5 | 7 | 5 | ribonuclease G (cafA) |
| SO4095 | | | | | 3 | | | | maf protein (maf) |
| SO4097 | 10 | 12 | 13 | 12 | 8 | 9 | 9 | 6 | rod shape-determining protein MreC (mreC) |
| SO4098 | 18 | 25 | 25 | 28 | 23 | 19 | 27 | 28 | rod shape-determining protein MreB (mreB) |
| SO4100 | | 2 | 3 | 3 | | 2 | 2 | | MSHA biogenesis protein MshQ (mshQ) |
| SO4102 | | | 2 | | | | | 2 | MSHA biogenesis protein MshO (mshO) |
| SO4105 | 21 | 18 | 25 | 22 | 19 | 20 | 23 | 25 | MSHA pilin protein MshA (mshA) |
| SO4106 | 6 | 8 | 11 | 11 | 7 | 7 | 9 | 10 | MSHA pilin protein MshB (mshB) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO4107 | | | | | 2 | | | | hypothetical protein |
| SO4108 | 4 | 4 | 6 | 7 | 7 | 3 | 6 | 5 | MSHA biogenesis protein MshG (mshG) |
| SO4109 | 9 | 15 | 8 | 12 | 13 | 7 | 6 | 10 | MSHA biogenesis protein MshE (mshE) |
| SO4110 | 5 | 11 | 8 | 8 | 11 | 11 | 7 | 9 | MSHA biogenesis protein MshN, putative |
| SO4111 | 6 | 3 | 3 | 4 | 3 | 7 | | 2 | MSHA biogenesis protein MshM (mshM) |
| SO4112 | 4 | 4 | 7 | 12 | 9 | 3 | 8 | 6 | MSHA biogenesis protein MshL (mshL) |
| SO4113 | 3 | | 2 | | 4 | 2 | 2 | | MSHA biogenesis protein MshK (mshK) |
| SO4114 | 6 | 5 | 5 | 8 | 4 | 4 | 6 | 4 | MSHA biogenesis protein MshJ (mshJ) |
| SO4116 | 2 | 2 | | | 3 | | | 2 | MSHA biogenesis protein MshH (mshH) |
| SO4118 | 13 | 14 | 13 | 12 | 9 | 8 | 9 | 13 | malate oxidoreductase, putative |
| SO4120 | 6 | 6 | 7 | 8 | 4 | 3 | 8 | 11 | ribosomal protein L31 (rpmE) |
| SO4121 | | 2 | | 2 | 4 | | | 3 | hypothetical protein |
| SO4122 | 3 | 2 | | 3 | 2 | | | 2 | primosomal protein N' (priA) |
| SO4123 | 12 | 13 | 7 | 6 | 9 | 6 | 12 | 4 | arginyl-tRNA synthetase (argS) |
| SO4124 | 8 | 3 | 4 | 6 | 11 | 8 | 6 | 4 | cell division protein FtsN, putative |
| SO4128 | 18 | 19 | 21 | 13 | 23 | 14 | 20 | 24 | SPFH domain/Band 7 family protein |
| SO4129 | 24 | 33 | 27 | 29 | 26 | 17 | 28 | 30 | SPFH domain/Band 7 family protein |
| SO4130 | | 2 | | | | | | | conserved hypothetical protein |
| SO4131 | 3 | | | 2 | 2 | | | | conserved hypothetical protein |
| SO4133 | 10 | 11 | 13 | 14 | 16 | 12 | 16 | 12 | uridine phosphorylase (udp) |
| SO4134 | | 2 | | 2 | 4 | 5 | 9 | 7 | conserved hypothetical protein |
| SO4135 | | | 2 | 3 | | | | 3 | ISSo7, transposase |
| SO4138 | | 2 | | 2 | 2 | | | | conserved hypothetical protein |
| SO4139 | 15 | 24 | 16 | 19 | 20 | 20 | 20 | 27 | conserved domain protein |
| SO4140 | | 2 | 3 | | 3 | | 2 | | transcriptional regulator, LysR family |
| SO4141 | 2 | 5 | 3 | 3 | 5 | | 5 | 6 | oxidoreductase, short-chain dehydrogenase/reductase family |
| SO4149 | 2 | | | | | | | | RTX toxin, putative |
| SO4151 | | | 2 | | | | 3 | 3 | polysaccharide deacetylase family protein |
| SO4154 | | 3 | | | | | | | transcriptional regulator, LysR family |
| SO4155 | 4 | | | | 4 | 4 | | | sensor histidine kinase |
| SO4157 | 2 | | | | 3 | | | | DNA-binding response regulator |
| SO4160 | | | | 2 | | | | | flagellar biosynthesis protein, putative |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO4162 | 5 | 7 | 10 | 9 | 7 | 6 | 9 | 8 | ATP-dependent protease HslV (hslV) |
| SO4163 | 42 | 48 | 51 | 59 | 38 | 43 | 54 | 57 | heat shock protein HslVU, ATPase subunit HslU (hslU) |
| SO4164 | 2 | 4 | 3 | 4 | 2 | | 2 | 4 | conserved hypothetical protein |
| SO4173 | 4 | 5 | 4 | 5 | 5 | 5 | 3 | 3 | sensor histidine kinase |
| SO4177 | 2 | | | 2 | | | 2 | 2 | hypothetical protein |
| SO4178 | 2 | 4 | 2 | 6 | | 5 | 3 | 2 | conserved hypothetical protein |
| SO4181 | 2 | 4 | | | | | | | RNA pseudouridylate synthase family protein |
| SO4189 | | 2 | | 3 | | | 3 | | conserved hypothetical protein |
| SO4190 | 16 | 15 | 10 | 13 | 11 | 6 | 12 | 10 | inorganic pyrophosphatase, manganese-dependent (ppaC) |
| SO4193 | 5 | 5 | 3 | 5 | 4 | 4 | 6 | 6 | conserved hypothetical protein |
| SO4199 | 17 | 20 | 15 | 22 | 21 | 20 | 15 | 23 | ubiquinone/menaquinone biosynthesis methlytransferase UbiE (ubiE) |
| SO4200 | 3 | 2 | 2 | | 4 | 3 | | | conserved hypothetical protein |
| SO4201 | 11 | 9 | 6 | 11 | 6 | 10 | 10 | 12 | ubiquinone biosynthesis protein AarF (aarF) |
| SO4202 | 6 | 9 | 7 | 6 | 6 | 7 | 7 | 11 | Sec-independent protein translocase protein TatA (tatA) |
| SO4203 | 5 | 12 | 5 | 7 | 10 | 7 | 9 | 10 | Sec-independent protein translocase protein TatB (tatB) |
| SO4205 | | | | | | 2 | | | hypothetical protein |
| SO4206 | 2 | | | | | | | | hydrolase, TatD family |
| SO4207 | 4 | 11 | 3 | 4 | 3 | 8 | 3 | | GGDEF domain protein |
| SO4208 | 10 | 11 | 5 | 7 | 9 | 8 | 5 | 7 | delta-aminolevulinic acid dehydratase (hemB-2) |
| SO4211 | 97 | 110 | 101 | 118 | 117 | 95 | 111 | 128 | preprotein translocase, SecA subunit (secA) |
| SO4212 | | | 2 | | | | | | peptidase, M23/M37 family |
| SO4214 | 2 | | | | | | | | UDP-3-0-acyl N-acetylglucosamine deacetylase (lpxC) |
| SO4215 | 16 | 23 | 27 | 25 | 16 | 17 | 29 | 23 | cell division protein FtsZ (ftsZ) |
| SO4216 | 14 | 16 | 20 | 17 | 17 | 10 | 15 | 16 | cell division protein FtsA (ftsA) |
| SO4217 | 2 | 4 | 3 | 2 | 3 | 3 | 4 | 3 | cell division protein FtsQ (ftsQ) |
| SO4218 | 5 | 6 | 9 | 5 | 8 | 2 | 4 | 8 | UDP-N-acetylmuramate--alanine ligase (murC) |
| SO4219 | 21 | 21 | 21 | 16 | 17 | 17 | 25 | 28 | UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase (murG) |
| SO4221 | 4 | 2 | 7 | 6 | 7 | | 6 | 7 | UDP-N-acetylmuramoylalanine--D-glutamate ligase (murD) |
| SO4223 | 6 | 14 | 10 | 14 | 8 | 6 | 10 | 16 | UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanyl ligase (murF) |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| **Locus** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Run 1** | **Run 2** | **Description** |
| SO4224 | | 2 | 2 | 4 | 3 | | 2 | 2 | UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase (murE) |
| SO4225 | 11 | 12 | 8 | 11 | 8 | 7 | 8 | 10 | peptidoglycan synthetase FtsI (ftsI) |
| SO4226 | | 2 | | | | 3 | | | cell division protein FtsL (ftsL) |
| SO4227 | 7 | 13 | 8 | 6 | 9 | 8 | 4 | 7 | conserved hypothetical protein TIGR00006 |
| SO4228 | | 2 | 2 | 3 | 2 | | 2 | | conserved hypothetical protein TIGR00242 |
| SO4230 | 2 | | 2 | 2 | | | | 7 | glycerol kinase (glpK) |
| SO4232 | | | 2 | 3 | | | | 4 | long-chain fatty acid transport protein |
| SO4233 | | 3 | | | | | 2 | | 3-isopropylmalate dehydratase, small subunit (leuD) |
| SO4234 | | | | | 2 | | 2 | | 3-isopropylmalate dehydratase, large subunit (leuC) |
| SO4235 | | | 2 | 2 | | | | | 3-isopropylmalate dehydrogenase (leuB) |
| SO4236 | | 2 | | | | | | | 2-isopropylmalate synthase (leuA) |
| SO4241 | | | | | 3 | | | | ATP-dependent DNA helicase RecQ (recQ) |
| SO4244 | | 4 | 3 | 2 | | 5 | | | hypothetical protein |
| SO4246 | 6 | 8 | 9 | 12 | 7 | 9 | 7 | 8 | ribosomal protein L33 (rpmG) |
| SO4247 | 11 | 11 | 12 | 15 | 12 | 10 | 12 | 14 | ribosomal protein L28 (rpmB) |
| SO4249 | 10 | 10 | 8 | 7 | 7 | 8 | | 6 | DNA/pantothenate metabolism flavoprotein (dfp) |
| SO4250 | 4 | 7 | 2 | 5 | 5 | 2 | 2 | 5 | deoxyuridine 5-triphosphate nucleotidohydrolase (dut) |
| SO4251 | | 2 | 3 | | 2 | 2 | 2 | | transcriptional regulator, TetR family |
| SO4252 | | 6 | 2 | | 2 | | 3 | | prolyl oligopeptidase family protein |
| SO4254 | 8 | 7 | 7 | 6 | 8 | 6 | 5 | 5 | GTP cyclohydrolase I (folE) |
| SO4255 | 3 | 6 | 8 | 6 | 7 | 4 | 8 | 10 | orotate phosphoribosyltransferase (pyrE) |
| SO4256 | 8 | 10 | 4 | 6 | 5 | 9 | 7 | 6 | ribonuclease PH (rph) |
| SO4257 | 12 | 9 | 8 | 10 | 8 | 9 | 9 | 7 | conserved hypothetical protein TIGR00255 |
| SO4258 | 5 | 7 | 12 | 13 | 5 | 2 | 11 | 10 | site-specific recombinase, phage integrase family |
| SO4261 | | 2 | | | | | | | hypothetical protein |
| SO4262 | | | | | 2 | | 2 | 2 | hypothetical protein |
| SO4263 | | | | 2 | 2 | | 2 | | conserved hypothetical protein |
| SO4264 | 7 | 7 | 6 | 8 | 13 | 4 | 3 | 8 | type I restriction-modification system, S subunit (hsdS-2) |
| SO4265 | 12 | 12 | 10 | 10 | 14 | 9 | 12 | 18 | type I restriction-modification system, M subunit (hsdM-2) |
| SO4266 | | 2 | | | | | | | conserved hypothetical protein |
| SO4267 | | | | 2 | | | | 3 | type I restriction-modification system, R subunit (hsdR-2) |
| SO4270 | | | 2 | | | | | | hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO4280 | | | | 2 | | | | | hypothetical protein |
| SO4281 | | 2 | | | 2 | | | | potassium uptake protein KtrA, putative |
| SO4283 | | 3 | 2 | | 3 | 4 | 4 | | apbE family protein |
| SO4284 | 4 | 3 | 3 | | 2 | 4 | 3 | 2 | ISSo5, transposase |
| SO4286 | 2 | 3 | | 5 | | 3 | 2 | 3 | chemotaxis motB protein (motB) |
| SO4289 | | 2 | 3 | 2 | 2 | | | | phosphate ABC transporter, ATP-binding protein (pstB-2) |
| SO4299 | | | | | 2 | | | 2 | chloramphenicol acetyltransferase (cat) |
| SO4308 | 2 | 2 | 4 | 2 | 3 | | 2 | | diaminopimelate epimerase (dapF) |
| SO4309 | 4 | 2 | 2 | 4 | 4 | 4 | 3 | 5 | diaminopimelate decarboxylase (lysA) |
| SO4310 | | | | | | | | 3 | hypothetical protein |
| SO4311 | 4 | 4 | 5 | 4 | 5 | 3 | 4 | 5 | cyay protein (cyaY) |
| SO4313 | 12 | 15 | 11 | 17 | 15 | 16 | 10 | 14 | porphobilinogen deaminase (hemC) |
| SO4315 | 19 | 33 | 23 | 31 | 24 | 30 | 29 | 28 | hemX protein (hemX) |
| SO4316 | 12 | 18 | 16 | 21 | 18 | 19 | 18 | 18 | hemY protein, putative |
| SO4317 | 3 | 7 | 3 | 5 | 7 | 2 | 6 | 6 | RTX toxin, putative |
| SO4318 | 9 | 8 | 7 | 7 | 10 | 13 | 8 | 6 | toxin secretion ATP-binding protein (rtxB) |
| SO4319 | 17 | 18 | 17 | 17 | 26 | 15 | 11 | 12 | HlyD family secretion protein |
| SO4320 | 40 | 46 | 21 | 32 | 43 | 34 | 28 | 34 | agglutination protein (aggA) |
| SO4321 | 15 | 16 | 22 | 17 | 11 | 15 | 17 | 16 | OmpA family protein |
| SO4323 | 9 | 14 | 10 | 8 | 19 | 15 | 8 | 10 | GGDEF domain protein |
| SO4325 | | 4 | | 2 | | | 4 | 6 | ATP-dependent DNA helicase Rep (rep) |
| SO4326 | 2 | | | 3 | | | | | transcriptional regulator, TetR family |
| SO4327 | 21 | 25 | 24 | 27 | 21 | 21 | 20 | 30 | HlyD family secretion domain protein |
| SO4329 | 10 | 13 | 13 | 7 | 8 | 14 | 8 | 8 | conserved hypothetical protein |
| SO4334 | 5 | 2 | 3 | 6 | 6 | 5 | 6 | 4 | inner membrane protein, putative |
| SO4340 | 5 | 9 | 9 | 5 | 8 | 7 | 8 | 5 | conserved hypothetical protein |
| SO4343 | 4 | 7 | 6 | 6 | 4 | 2 | 9 | 9 | aminotransferase, class V |
| SO4344 | | 4 | 2 | 2 | 2 | 2 | | | threonine dehydratase (ilvA) |
| SO4345 | | 2 | 5 | | 3 | | 2 | | dihydroxy-acid dehydratase (ilvD) |
| SO4349 | 4 | 6 | 2 | 8 | 5 | | 9 | 11 | ketol-acid reductoisomerase (ilvC) |
| SO4350 | 2 | | 2 | | 2 | | | 2 | transcriptional regulator ilvY (ilvY) |
| SO4351 | 3 | 5 | 2 | 3 | | 2 | | 3 | CBS domain protein |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO4356 | 7 | 13 | 9 | 5 | 12 | 7 | 11 | 10 | conserved domain protein |
| SO4358 | 3 | | 2 | 4 | | 2 | | 4 | anaerobic dimethyl sulfoxide reductase, A subunit (dmaA-2) |
| SO4364 | | 5 | 6 | 2 | 4 | | 3 | 5 | ATP-dependent DNA helicase RecG (recG) |
| SO4365 | 19 | 18 | 22 | 22 | 18 | 19 | 17 | 18 | hypothetical protein |
| SO4367 | | | | 2 | 2 | | | | acyltransferase family protein |
| SO4371 | 2 | | | | | 2 | | 2 | conserved hypothetical protein |
| SO4372 | 2 | | | | | | | | thioester dehydrase family protein |
| SO4373 | 8 | 12 | 4 | 4 | 8 | 10 | 7 | 10 | glycosyl transferase, group 2 family protein |
| SO4374 | | 6 | 2 | 3 | | 4 | | 4 | histidine ammonia-lyase, putative |
| SO4377 | 2 | 4 | 4 | 4 | 3 | 2 | 3 | 5 | membrane protein, putative |
| SO4378 | 2 | 3 | 3 | 4 | 3 | 2 | 2 | 4 | FAD-binding protein |
| SO4380 | | | | | | | 2 | | 3-oxoacyl-(acyl-carrier-protein) synthase II, putative |
| SO4381 | | | | 2 | | | | | thioester dehydrase family protein |
| SO4382 | 5 | 3 | 2 | 7 | 4 | 4 | 6 | 5 | 3-oxoacyl-(acyl-carrier-protein) reductase (fabG-2) |
| SO4383 | | 2 | | 2 | 3 | | | 2 | 3-oxoacyl-(acyl-carrier-protein) synthase II (fabF-2) |
| SO4384 | 28 | 30 | 22 | 23 | 22 | 27 | 12 | 22 | hypothetical protein |
| SO4385 | | | | 2 | | | | | von Willebrand factor type A domain protein |
| SO4391 | 5 | 3 | 3 | 5 | 6 | 5 | 8 | 9 | hypothetical protein |
| SO4393 | | | | 2 | | | | | acetyltransferase, GNAT family |
| SO4394 | | 2 | | | | | | | phage shock protein E (pspE-2) |
| SO4396 | | | 2 | 3 | 2 | | 3 | 3 | acyl carrier protein phosphodiesterase (acpD) |
| SO4397 | | | | | 2 | | | | conserved hypothetical protein |
| SO4398 | | | | 2 | | | | | conserved hypothetical protein TIGR00256 |
| SO4399 | 6 | 11 | 10 | 5 | 8 | 5 | 9 | 11 | hypothetical protein |
| SO4403 | 5 | 14 | 11 | 8 | 5 | 13 | 9 | 10 | hypothetical protein |
| SO4405 | | | | | 2 | 2 | | | catalase/peroxidase HPI (katG-2) |
| SO4408 | 22 | 31 | 24 | 22 | 29 | 20 | 15 | 27 | virulence regulator BipA (bipA) |
| SO4410 | 22 | 30 | 15 | 24 | 25 | 21 | 33 | 28 | glutamine synthetase, type I (glnA) |
| SO4414 | | | | 3 | | | 3 | | conserved domain protein |
| SO4418 | | 2 | 2 | 3 | | 2 | 2 | 3 | trypanothione synthetase domain protein |
| SO4420 | 2 | 2 | | 3 | 3 | 3 | 3 | 2 | peptidase, M23/M37 family |
| SO4426 | 2 | 2 | | | | | 2 | | RNA pseudouridylate synthase family protein |

| | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | |
|---|---|---|---|---|---|---|---|---|---|
| Locus | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Description |
| SO4427 | | | 2 | 2 | 3 | | | 3 | sensor histidine kinase |
| SO4428 | 5 | 8 | 7 | 8 | 9 | | 5 | 8 | DNA-binding response regulator |
| SO4438 | | | 3 | 2 | 3 | | 2 | 2 | hypothetical protein |
| SO4439 | | | | | | | | 2 | conserved hypothetical protein |
| SO4443 | | | | | | 2 | | | hypothetical protein |
| SO4445 | 2 | | | | | | | | response regulator/sensor histidine kinase |
| SO4446 | 5 | 3 | 2 | 3 | 3 | | 3 | 2 | molybdenum ABC transporter, ATP-binding protein |
| SO4448 | | 2 | 2 | 5 | 2 | | | | molybdenum ABC transporter, periplasmic molybdenum-binding protein |
| SO4449 | 4 | 4 | 6 | 5 | 5 | | 8 | 9 | molybdenum cofactor biosynthesis protein E (moaE) |
| SO4450 | 2 | 2 | | 2 | | | | | molybdenum cofactor biosynthesis protein D (moaD) |
| SO4451 | | 6 | 6 | 2 | | 3 | 3 | 4 | molybdenum cofactor biosynthesis protein C (moaC) |
| SO4452 | | 3 | 2 | 3 | 2 | | 3 | 3 | molybdenum cofactor biosynthesis protein A (moaA) |
| SO4453 | 24 | 21 | 17 | 17 | 24 | 16 | 16 | 20 | electron transfer flavoprotein-ubiquinone oxidoreductase, putative |
| SO4454 | 10 | 12 | 5 | 16 | 15 | 11 | 10 | 13 | methyl-accepting chemotaxis protein |
| SO4456 | | | 2 | | 3 | | 3 | 2 | conserved hypothetical protein |
| SO4457 | | 3 | | | | | 3 | 3 | GGDEF domain protein |
| SO4465 | | | 2 | | | | | | conserved domain protein |
| SO4466 | 3 | 2 | | | 2 | | 2 | | methyl-accepting chemotaxis protein |
| SO4469 | 2 | 4 | 7 | 5 | 5 | | 5 | | alcohol dehydrogenase, iron-containing |
| SO4470 | | | | 3 | 2 | | 2 | | conserved hypothetical protein |
| SO4471 | | | | | | 2 | | | nitrogen regulation protein (ntrB) |
| SO4472 | | | 3 | 3 | | | 5 | 3 | nitrogen regulation protein NR(I) (ntrC) |
| SO4473 | | 3 | | 2 | | | 3 | 2 | outer membrane protein, putative |
| SO4475 | | | | | | | | 2 | cation efflux family protein |
| SO4476 | | 3 | | | | | | | spheroplast protein y precursor, putative |
| SO4477 | 2 | | | | | | 3 | 2 | transcriptional regulatory protein CpxR (cpxR) |
| SO4478 | | | 4 | | 2 | | 4 | 3 | sensor protein CpxA (cpxA) |
| SO4479 | 3 | | | | 2 | | 2 | 3 | sigma-54 dependent transcriptional regulator |
| SO4480 | | | 4 | 2 | 3 | | 6 | 11 | aldehyde dehydrogenase (aldA) |
| SO4488 | 2 | | | | | | | 2 | sensor histidine kinase |
| SO4490 | | | | | | | 2 | | hypothetical protein |
| SO4492 | 8 | 7 | 6 | 9 | 10 | 5 | 6 | 9 | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO4497 | 4 | 3 | 3 | | 2 | 4 | 3 | 2 | ISSo5, transposase |
| SO4503 | 2 | | 2 | | | | | | formate dehydrogenase accessory protein FdhD, putative |
| SO4505 | 5 | 7 | 7 | 8 | 5 | 6 | 6 | 6 | conserved hypothetical protein |
| SO4506 | 10 | 7 | 3 | 6 | 11 | 3 | 3 | 3 | iron-sulfur cluster-binding protein |
| SO4509 | 38 | 50 | 44 | 51 | 38 | 32 | 38 | 38 | formate dehydrogenase, alpha subunit |
| SO4510 | 8 | 5 | 5 | 8 | 6 | 2 | 6 | 8 | formate dehydrogenase, iron-sulfur subunit (fdhB-1) |
| SO4511 | 2 | | | 2 | 2 | | | | formate dehydrogenase, C subunit, putative |
| SO4513 | 16 | 18 | 5 | 3 | 48 | 36 | 4 | 3 | formate dehydrogenase, alpha subunit |
| SO4514 | 4 | 2 | 4 | 3 | 5 | | 3 | 4 | formate dehydrogenase, iron-sulfur subunit (fdhB-2) |
| SO4516 | 6 | 10 | 15 | 10 | | 3 | 15 | 12 | ferric vibriobactin receptor (viuA) |
| SO4520 | 24 | 24 | 20 | 20 | 36 | 15 | 20 | 20 | oxygen-independent coproporphyrinogen III oxidase, putative |
| SO4523 | 28 | 41 | 40 | 44 | 20 | 26 | 63 | 63 | iron-regulated outer membrane virulence protein (irgA) |
| SO4525 | | | | | 2 | 2 | | | conserved hypothetical protein |
| SO4529 | 5 | 7 | | 6 | 3 | | 5 | 5 | RNA methyltransferase, TrmH family, group 2 |
| SO4537 | 12 | 13 | 10 | 11 | 14 | 14 | 5 | 11 | peptidase, putative |
| SO4557 | 13 | 23 | 17 | 24 | 17 | 19 | 20 | 23 | methyl-accepting chemotaxis protein |
| SO4558 | | | | 2 | | | | 3 | hypothetical protein |
| SO4559 | | | | | 3 | | | | conserved domain protein |
| SO4561 | 8 | 5 | 4 | 7 | 5 | 7 | | | conserved hypothetical protein |
| SO4564 | | | | | | 2 | | | TonB2 protein, putative |
| SO4567 | | | | | | | | 2 | transcriptional regulator, AsnC family |
| SO4573 | 4 | 4 | | 3 | 2 | | 3 | 4 | 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylic acid synthase/2-oxoglutarate decarboxylase (menD) |
| SO4575 | 3 | 7 | 4 | 4 | 6 | | 3 | 8 | O-succinylbenzoate-CoA synthase (menC) |
| SO4583 | 2 | | | | | | | 2 | RNA polymerase sigma-32 factor (rpoH) |
| SO4584 | 2 | 7 | 2 | 3 | 4 | 5 | 3 | 4 | cell division permease protein FtsX (ftsX) |
| SO4585 | 10 | 10 | 13 | 15 | 8 | 7 | 9 | 12 | cell division ATP-binding protein FtsE (ftsE) |
| SO4586 | 24 | 31 | 41 | 41 | 25 | 29 | 38 | 41 | cell division protein FtsY (ftsY) |
| SO4587 | | 2 | 2 | 2 | 2 | | 2 | 2 | conserved hypothetical protein TIGR00095 |
| SO4591 | 3 | 3 | 3 | 2 | 6 | 7 | | 4 | tetraheme cytochrome c (cymA) |
| SO4593 | | | | | 3 | 4 | 2 | 4 | hypothetical protein |
| SO4597 | | 2 | 3 | 3 | 5 | 9 | 2 | 4 | heavy metal efflux system protein, putative |
| SO4598 | | 4 | 2 | 3 | 8 | 8 | 3 | 5 | heavy metal efflux pump, CzcA family |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO4599 | 5 | 4 | 6 | 4 | 4 | | 3 | 5 | ribonuclease, T2 family |
| SO4602 | 35 | 53 | 37 | 39 | 40 | 44 | 43 | 40 | glycerol-3-phosphate acyltransferase (plsB) |
| SO4603 | 4 | | 4 | 4 | 5 | 2 | | 2 | LexA repressor (lexA) |
| SO4615 | | | | 2 | | | | | SCO1/SenC family protein |
| SO4616 | | 2 | | | | | 2 | | polysaccharide deacetylase family protein |
| SO4618 | | | | | 2 | | | | prolyl oligopeptidase family protein |
| SO4619 | 5 | 9 | 6 | 10 | 7 | 7 | 6 | 10 | yhgI protein (yhgI) |
| SO4620 | 4 | 2 | 5 | 3 | 9 | 10 | 3 | 7 | fumarate reductase, flavoprotein subunit precursor (ifcA-2) |
| SO4626 | | | | | 2 | | | | bioH protein (bioH) |
| SO4628 | 13 | 11 | 10 | 16 | 11 | 11 | 9 | 10 | sulfatase |
| SO4629 | 21 | 22 | 22 | 20 | 19 | 14 | 26 | 27 | conserved hypothetical protein |
| SO4633 | 4 | 5 | 7 | 4 | 3 | 3 | 2 | 5 | transcriptional regulatory protein OmpR (ompR) |
| SO4634 | 3 | 5 | 5 | 3 | 4 | 3 | 3 | 5 | osmolarity sensor protein EnvZ (envZ) |
| SO4635 | | 2 | | | | 2 | 2 | 2 | methyl-accepting chemotaxis protein |
| SO4637 | | | | 2 | | | | 3 | DNA-binding response regulator |
| SO4638 | 2 | | | 2 | 2 | | 2 | 2 | sensor histidine kinase |
| SO4640 | 11 | 15 | 10 | 9 | 10 | 9 | 11 | 13 | antioxidant, AhpC/Tsa family |
| SO4642 | | | | 2 | | 2 | | 2 | conserved hypothetical protein |
| SO4643 | | 3 | 3 | | 2 | 3 | 3 | 3 | hypothetical protein |
| SO4645 | | 2 | | | 2 | | | | hypothetical protein |
| SO4647 | 2 | 2 | | | | 3 | 3 | 2 | DNA-binding response regulator |
| SO4648 | | | 2 | 2 | | 2 | | 2 | sensor histidine kinase |
| SO4650 | | | | | | | | 4 | conserved hypothetical protein |
| SO4651 | | | 3 | 5 | | | 4 | 6 | conserved hypothetical protein |
| SO4652 | | | 10 | 10 | | | 13 | 24 | sulfate ABC transporter, periplasmic sulfate-binding protein (sbp) |
| SO4653 | | | | | | | | 2 | sulfate ABC transporter, permease protein (cysT-2) |
| SO4655 | 3 | 6 | 19 | 11 | | | 26 | 26 | sulfate ABC transporter, ATP-binding protein (cysA-2) |
| SO4656 | | 2 | | 2 | | | | | hypothetical protein |
| SO4658 | 2 | 2 | | | | 2 | 2 | | conserved hypothetical protein |
| SO4659 | 6 | 6 | 4 | 6 | 2 | 5 | 7 | 5 | conserved hypothetical protein |
| SO4661 | 2 | 4 | | 2 | 2 | 6 | 2 | | hypothetical protein |
| SO4662 | 6 | 10 | 7 | 12 | 10 | 7 | 10 | 11 | lemA protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO4666 | 22 | 27 | 23 | 22 | 23 | 18 | 18 | 19 | cytochrome c (cytcB) |
| SO4667 | 8 | 8 | 3 | 6 | 5 | 9 | 7 | 7 | GTP-binding protein EngB (engB) |
| SO4669 | 11 | 17 | 15 | 12 | 17 | 8 | 14 | 12 | DNA polymerase I (polA) |
| SO4670 | 3 | 4 | 4 | 5 | 3 | | 4 | 3 | enhancing lycopene biosynthesis protein (elbB) |
| SO4672 | | | 3 | 3 | 2 | | | 2 | glpE protein (glpE) |
| SO4673 | 9 | 7 | 12 | 11 | 14 | 4 | 11 | 7 | threonine 3-dehydrogenase (tdh) |
| SO4674 | 13 | 13 | 16 | 12 | 15 | 10 | 13 | 17 | 2-amino-3-ketobutyrate coenzyme A ligase (kbl) |
| SO4675 | | | 2 | | | | | | transcriptional regulator, TetR family |
| SO4676 | 7 | 6 | 6 | 6 | 6 | 7 | 5 | 9 | 3-deoxy-D-manno-octulosonic-acid (KDO) transferase (kdtA) |
| SO4677 | 4 | 8 | 3 | 6 | 5 | 7 | 4 | 4 | conserved hypothetical protein |
| SO4678 | 11 | 8 | 4 | 10 | 9 | 12 | 13 | 10 | heptosyl transferase, glycosyltransferase family 9 protein |
| SO4679 | 19 | 19 | 16 | 16 | 15 | 21 | 16 | 21 | glycosyl transferase, group 1 family protein |
| SO4680 | 13 | 15 | 17 | 14 | 15 | 13 | 17 | 16 | conserved hypothetical protein |
| SO4681 | 21 | 23 | 20 | 21 | 25 | 19 | 19 | 19 | glycosyl transferase, group 1 family protein |
| SO4682 | 3 | 3 | 2 | 3 | 5 | | 6 | 4 | glycosyl transferase, group 1 family protein |
| SO4684 | 5 | 6 | 5 | 2 | 2 | 3 | 2 | 5 | phosphopantetheine adenylyltransferase (coaD) |
| SO4685 | 15 | 24 | 16 | 14 | 13 | 16 | 19 | 21 | conserved hypothetical protein |
| SO4686 | | | 2 | | 2 | | | | NAD dependent epimerase/dehydratase family protein |
| SO4687 | 2 | | | | | | | | UDP-glucose 6-dehydrogenase (ugd) |
| SO4688 | | | | | 2 | 2 | | 4 | glycosyl transferase, group 2 family protein |
| SO4690 | | | | | 2 | 2 | | | conserved hypothetical protein |
| SO4692 | 35 | 40 | 33 | 39 | 25 | 28 | 31 | 36 | AcrB/AcrD/AcrF family protein |
| SO4693 | 40 | 47 | 42 | 44 | 39 | 44 | 44 | 46 | multidrug resistance protein, AcrA/AcrE family |
| SO4696 | | | 2 | | 2 | | 2 | | conserved hypothetical protein |
| SO4697 | | | | | | | 2 | | glutathione S-transferase (gst) |
| SO4699 | 10 | 15 | 15 | 14 | 13 | 5 | 14 | 16 | oligopeptidase A (prlC) |
| SO4702 | 13 | 15 | 13 | 11 | 14 | 4 | 9 | 15 | glutathione reductase (gor) |
| SO4704 | | | | | | 2 | | | hypothetical protein |
| SO4711 | | 3 | | | | | | | HD domain protein |
| SO4712 | | | | 2 | | | | | ABC transporter, ATP-binding protein, putative |
| SO4713 | | 2 | | | 2 | | 2 | | menaquinone-specific isochorismate synthase, putative |
| SO4716 | | | | | | | 2 | | acetyltransferase, GNAT family |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO4717 | 10 | 8 | 4 | 5 | 3 | 5 | 6 | 12 | sensor histidine kinase |
| SO4718 | 8 | 10 | 6 | 8 | 8 | 6 | 5 | 12 | sigma-54 dependent response regulator |
| SO4719 | 31 | 36 | 31 | 30 | 24 | 35 | 18 | 19 | conserved hypothetical protein |
| SO4721 | 3 | 2 | | | 2 | 2 | 4 | 2 | ABC transporter, ATP-binding protein |
| SO4722 | 2 | | | 3 | 2 | 2 | | 2 | molybdopterin-guanine dinucleotide biosynthesis protein (mobA) |
| SO4723 | 16 | 20 | 11 | 13 | 13 | 11 | 12 | 12 | molybdopterin biosynthesis MoeA protein, putative |
| SO4724 | | 2 | | 2 | | 2 | | | molybdenum cofactor biosynthesis protein A, putative |
| SO4725 | 3 | 5 | 3 | 4 | 4 | 6 | 3 | 3 | conserved hypothetical protein |
| SO4726 | 5 | 4 | 4 | 3 | 7 | 2 | 2 | 3 | formamidopyrimidine-DNA glycosylase (mutM) |
| SO4728 | 5 | 6 | 9 | 8 | 9 | 6 | 6 | 10 | conserved hypothetical protein |
| SO4729 | 9 | 7 | 8 | 12 | 6 | 4 | 7 | 11 | conserved hypothetical protein |
| SO4730 | 4 | 10 | 9 | 8 | 13 | 3 | 13 | 11 | oxygen-independent coproporphyrinogen III oxidase (hemN) |
| SO4731 | | 4 | 5 | 5 | 5 | | 5 | 8 | adenosine deaminase (add) |
| SO4732 | 7 | 12 | 6 | 9 | 7 | 6 | 7 | 8 | conserved hypothetical protein |
| SO4733 | 7 | 4 | 4 | 6 | 5 | 5 | 5 | 5 | lysophospholipase L2 (lypA) |
| SO4734 | | | 3 | | | | 2 | 2 | sensory box protein |
| SO4737 | 3 | 5 | 2 | 3 | 7 | 2 | 3 | 5 | iron-sulfur cluster-binding protein |
| SO4738 | 4 | 3 | 2 | 4 | 3 | 4 | 4 | 4 | hypothetical protein |
| SO4739 | 8 | 4 | 2 | 10 | 7 | 6 | 12 | 9 | naphthoate synthase (menB) |
| SO4741 | 12 | 19 | 12 | 18 | 16 | 9 | 14 | 19 | glucosamine--fructose-6-phosphate aminotransferase (isomerizing) (glmS) |
| SO4742 | 4 | 3 | 5 | 6 | 3 | 3 | 2 | | transcriptional regulator, DeoR family |
| SO4743 | 46 | 64 | 55 | 58 | 29 | 32 | 63 | 63 | TonB-dependent receptor, putative |
| SO4745 | 11 | 11 | 7 | 6 | 4 | 7 | 8 | 12 | UDP-N-acetylglucosamine pyrophosphorylase (glmU) |
| SO4746 | 12 | 12 | 16 | 15 | 12 | 12 | 17 | 15 | ATP synthase F1, epsilon subunit (atpC) |
| SO4747 | 72 | 82 | 78 | 75 | 81 | 64 | 86 | 79 | ATP synthase F1, beta subunit (atpD) |
| SO4748 | 49 | 51 | 46 | 45 | 49 | 51 | 50 | 43 | ATP synthase F1, gamma subunit (atpG) |
| SO4749 | 94 | 106 | 94 | 95 | 102 | 103 | 98 | 105 | ATP synthase F1, alpha subunit (atpA) |
| SO4750 | 16 | 17 | 16 | 15 | 14 | 18 | 15 | 12 | ATP synthase F1, delta subunit (atpH) |
| SO4751 | 40 | 47 | 41 | 42 | 39 | 43 | 41 | 45 | ATP synthase F0, B subunit (atpF) |
| SO4755 | 11 | 16 | 13 | 10 | 12 | 9 | 14 | 19 | ParB family protein |
| SO4756 | | | 3 | 3 | | | 3 | 2 | ParA family protein |
| SO4757 | 3 | 7 | | 4 | 6 | | 2 | 3 | glucose-inhibited division protein B (gidB) |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SO4758 | 29 | 35 | 28 | 34 | 26 | 30 | 27 | 32 | glucose-inhibited division protein A (gidA) |
| SOA0003 | 10 | 4 | 8 | 6 | 9 | 3 | 7 | 5 | type II restriction endonuclease, putative |
| SOA0004 | 12 | 9 | 6 | 16 | 8 | 4 | 9 | 14 | type II DNA modification methyltransferase |
| SOA0005 | | 2 | | 3 | 4 | | 2 | 3 | hypothetical protein |
| SOA0006 | | | | | 2 | | 2 | 3 | ParA family protein |
| SOA0008 | | | 2 | | | | | | hypothetical protein |
| SOA0011 | 5 | 5 | 6 | 4 | 4 | 3 | | 5 | conserved hypothetical protein |
| SOA0018 | 2 | | | | 2 | 2 | | 4 | TnSon1, conserved hypothetical protein |
| SOA0019 | 5 | 5 | 2 | 6 | 4 | 4 | 3 | 3 | TnSon1, resolvase |
| SOA0022 | | | 2 | | | | 2 | 2 | proteic killer active protein (higB) |
| SOA0023 | 3 | | 2 | 3 | 2 | 3 | | 2 | proteic killer suppressor protein (higA) |
| SOA0031 | | | 3 | | | | 3 | 6 | partition protein, ParB family, putative |
| SOA0032 | | 2 | | | | 2 | | | conserved hypothetical protein |
| SOA0033 | | 2 | 2 | 2 | | | 2 | 3 | hypothetical protein |
| SOA0040 | | | | | | | 3 | 2 | hypothetical protein |
| SOA0041 | 5 | 7 | 5 | 5 | 9 | 6 | 7 | 5 | transcriptional regulator, PemK family |
| SOA0042 | 3 | | 5 | 6 | | | 4 | 4 | hypothetical protein |
| SOA0045 | | | | | | 2 | | | site-specific recombinase, phage integrase family |
| SOA0048 | 18 | 22 | 26 | 23 | 24 | 17 | 23 | 24 | prolyl oligopeptidase family protein |
| SOA0049 | 4 | 5 | 3 | | 7 | 5 | 4 | 6 | toxin secretion ABC transporter, ATP-binding subunit/permease protein, putative |
| SOA0051 | 9 | 10 | 8 | 10 | 6 | 8 | 6 | 7 | hypothetical protein |
| SOA0056 | | 2 | 2 | 2 | | | | 2 | hypothetical protein |
| SOA0059 | 2 | 2 | 5 | 4 | 3 | 6 | | 2 | conserved hypothetical protein |
| SOA0060 | 3 | 5 | 8 | 7 | 4 | 3 | 6 | 8 | acetyltransferase, GNAT family |
| SOA0061 | 4 | 4 | 7 | 5 | 3 | 7 | 6 | 6 | parA protein, putative |
| SOA0062 | 7 | | 4 | 6 | 4 | 4 | 4 | 3 | hypothetical protein |
| SOA0067 | | | | | | | 2 | | hypothetical protein |
| SOA0069 | | | | | | | | 3 | conserved hypothetical protein |
| SOA0070 | 3 | | 7 | | 3 | | 5 | | hypothetical protein |
| SOA0075 | | | | | | 2 | | | hypothetical protein |
| SOA0077 | 3 | 4 | 3 | 3 | 2 | | 2 | 2 | site-specific recombinase, resolvase family |
| SOA0079 | | | 2 | 2 | | | 2 | | conserved hypothetical protein |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SOA0080 | | | 6 | 3 | | | | 4 | hypothetical protein |
| SOA0086 | | 2 | 2 | | | 4 | | 2 | site-specific recombinase, resolvase family |
| SOA0088 | | | 3 | | 2 | | | | plasmid stabilization protein ParE, putative |
| SOA0095 | 3 | 4 | 3 | 6 | 6 | 2 | 5 | 10 | partitioning protein A |
| SOA0096 | 11 | 14 | 14 | 13 | 13 | 14 | 15 | 15 | partitioning protein B |
| SOA0099 | 6 | 4 | 12 | 7 | 8 | 4 | 6 | 11 | conserved hypothetical protein |
| SOA0100 | 7 | 17 | 21 | 26 | 17 | 9 | 21 | 19 | conserved hypothetical protein |
| SOA0106 | 26 | 28 | 22 | 28 | 18 | 22 | 28 | 23 | methyl-accepting chemotaxis protein |
| SOA0108 | | | | | | | | 2 | hypothetical protein |
| SOA0110 | 14 | 19 | 18 | 24 | 24 | 22 | 18 | 24 | lipoprotein, putative |
| SOA0112 | 18 | 21 | 24 | 32 | 29 | 23 | 26 | 31 | lipoprotein, putative |
| SOA0114 | 15 | 20 | 17 | 18 | 16 | 23 | 23 | 20 | outer membrane protein A (ompA) |
| SOA0115 | 18 | 21 | 24 | 32 | 29 | 23 | 26 | 31 | lipoprotein, putative |
| SOA0119 | | | | | | | | 2 | ISSo13, transposase |
| SOA0131 | | | | 2 | | | 3 | 3 | hypothetical protein |
| SOA0132 | 3 | 2 | 3 | 3 | 5 | 2 | 5 | 2 | conserved hypothetical protein |
| SOA0135 | 6 | 5 | 6 | 4 | 6 | 5 | 7 | 6 | hypothetical protein |
| SOA0138 | 12 | 11 | 12 | 11 | 11 | 6 | 12 | 10 | hypothetical protein |
| SOA0139 | | 2 | 2 | 3 | 2 | | 2 | | hypothetical protein |
| SOA0140 | 33 | 35 | 30 | 19 | 31 | 19 | 30 | 30 | hypothetical protein |
| SOA0141 | 15 | 16 | 4 | 2 | 11 | 6 | | 3 | hypothetical protein |
| SOA0142 | | | | | | | | 2 | ISSo13, transposase |
| SOA0149 | | | | 2 | | | | 3 | conserved hypothetical protein |
| SOA0150 | 2 | 3 | 2 | 2 | 3 | 3 | 3 | 5 | hypothetical protein |
| SOA0151 | | | | | | | | 2 | ISSo13, transposase |
| SOA0153 | | 3 | 2 | 4 | 12 | 11 | 7 | 6 | heavy metal efflux pump, CzcA family |
| SOA0154 | | | 3 | 3 | 13 | 13 | 8 | 10 | heavy metal efflux protein, putative |
| SOA0157 | 2 | | | | 2 | 2 | | | hypothetical protein |
| SOA0160 | 3 | 5 | 8 | 5 | 11 | | 5 | 4 | esterase, putative |
| SOA0161 | 6 | 6 | 8 | 8 | 7 | 3 | 9 | 13 | zinc-binding dehydrogenase |
| SOA0164 | | | | 3 | | | | 3 | iron-containing alcohol dehydrogenase |
| SOA0165 | 11 | 11 | 12 | 8 | 14 | 9 | 7 | 13 | transcriptional regulator, LysR family |

| Locus | 45 min Control | | 45 min Cr | | 90 min Control | | 90 min Cr | | Description |
|---|---|---|---|---|---|---|---|---|---|
| | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | Run 1 | Run 2 | |
| SOA0169 | 2 | | | | 2 | 2 | | 4 | TnSon1, conserved hypothetical protein |
| SOA0170 | 5 | 5 | 2 | 6 | 4 | 4 | 3 | 3 | TnSon1, resolvase |
| SOA0171 | 3 | 8 | 3 | 2 | 2 | 3 | 2 | 3 | hypothetical protein |
| SOA0172 | | 2 | | 4 | 4 | | | | site-specific recombinase, resolvase family |
| SOA0173 | | | 3 | 6 | 2 | | 4 | 3 | hypothetical protein |

# VITA

Melissa Renee Thompson was born Melissa Renee Rugen on February 23, 1979 in Columbia, Mo. She spent her early childhood in the town of Green Ridge, Mo. Melissa started elementary school in Mexico, Mo attending Hawthorne Elementary School. She then transferred to Butler, Mo. attending Butler Elementary followed by Albany, Mo. for grades second thru fourth attending Virginia E. George Elementary. She finished her elementary and secondary education in California, Mo. graduating Salutatorian of her class from California High School in May of 1997. In December of 2001, she completed her Bachelor's of Science degree in Biological Sciences with Minors in Chemistry and Spanish from the University of Missouri-Columbia. After finishing her degree, she worked as a Senior Laboratory Technician for Dr. Kathleen Newton at the University of Missouri-Columbia. Her work with Dr. Newton included assisting in annotating the wild-type *Zea mays* mitochondrial genome, which was published in 2004.

In August of 2003, she enrolled in the Graduate School of Genome Science and Technology, a joint program between the University of Tennessee-Knoxville and Oak Ridge National Laboratory where she pursued a Doctoral of Philosophy degree in Life Sciences completing the degree in 2007. After graduation she moved to St. Louis, Mo to pursue a Post-Doctoral Fellow with Pfizer, Inc.