Model based prediction of physiology of *G. sulfurreducens* by flux balance and thermodynamics based metabolic flux analysis approaches

by

Srinath Garg Govindarajan

A thesis submitted in conformity with the requirements for the degree of

Master of Applied Science

Graduate Department of Chemical Engineering and Applied Chemistry

University of Toronto

©Copyright by Srinath Garg Govindarajan, 2009
ABSTRACT

Model based prediction of physiology of *G. sulfurreducens* by flux balance and thermodynamics based metabolic flux analysis approaches

Master of Applied Science, 2009

Srinath Garg Govindarajan

Graduate Department of Chemical Engineering and Applied Chemistry

University of Toronto

The development of genome scale metabolic models have been aided by the increasing availability of genome sequences of microorganisms such as *Geobacter sulfurreducens*, involved in environmentally relevant processes such as the *in-situ* bioremediation of U(VI). Since microbial activities are the major driving forces for geochemical changes in the sub-surface, understanding of microbial behavior under a given set of conditions can help predict the likely outcome of potential subsurface bioremediation strategies. Hence, a model based lookup table was created to capture the variation in physiology of *G. sulfurreducens* in response to environmental perturbations. Thermodynamically feasible flux distributions were generated by incorporating thermodynamic constraints in the model. These constraints together with the mass balance constraints formed the thermodynamics based metabolic flux analysis model (TMFA). Metabolomics experiments were performed to determine the concentration of intracellular metabolites. These concentrations were posed as constraints in the TMFA model to improve the model accuracy.
ACKNOWLEDGEMENT

At the outset, I would like to thank Prof. Radhakrishnan Mahadevan for having given me
an opportunity to be a part of a large and friendly research group. I would also like to
acknowledge his constant support, motivation and guidance throughout the course of my work.

My heartfelt thanks and deep appreciation goes to my parents for their continued support-
both moral and financial throughout the years of my study and for their unshakeable trust in all
my endeavours, despite them being overseas. In particular, I would like to thank my late mother
Smt. Hemamalini Govindarajan for all her love and moral support throughout the years. I would
not find myself where I am today if not for her words of wisdom.

I would like to thank all members of the Laboratory for Metabolic Systems Engineering
Group for their continued support and constant encouragement. I would particularly like to thank
Dr. Helen Vrionis for helping me out in the experimental portion of my work and Chenomx Inc,
Alberta for having helped out with the NMR analysis of my biological samples.

I would like to acknowledge funding support throughout the term of my program from
the Office of Science (US Department of Energy); Environmental Remediation Sciences
Program.

[iii]
# TABLE OF CONTENTS

ABSTRACT ................................................................................................................................. ii

ACKNOWLEDGEMENT ............................................................................................................. iii

TABLE OF CONTENTS .............................................................................................................. iv

LIST OF TABLES .................................................................................................................... vii

LIST OF FIGURES ................................................................................................................... ix

NOMENCLATURE .................................................................................................................... xii

MATHEMATICAL SYMBOLS ..................................................................................................... xiv

LIST OF EQUATIONS ................................................................................................................. xv

THESES OUTLINE ..................................................................................................................... 1

CHAPTER 1: INTRODUCTION ..................................................................................................... 3

  1.1: Literature Review: .............................................................................................................. 3

  1.2: Field Dynamics: ............................................................................................................... 4

  1.3: The Challenges in implementing *in situ* bioremediation: .................................................. 7

  1.4: The Need for Constraint Based Modeling: ...................................................................... 8

  1.5: Limitations of FBA-Need for TMFA: .............................................................................. 15

  1.6: Thermodynamics –Based Metabolic Flux Analysis (TMFA): ........................................... 18

  1.7: Metabolomics of *G. sulfurreducens*: ............................................................................. 21

STATEMENT OF OBJECTIVES ................................................................................................. 23

  2.1: Generating the Pre Calculated Lookup table: ................................................................. 26

  2.2: *In Silico* Model Simulations: ......................................................................................... 28

  2.3: Identifying the Infeasibility Thresholds: ......................................................................... 29
2.4: Simulating the Metabolic Flux Distributions: .................................................................30

2.5: Thermodynamics based Metabolic Flux Analysis: .........................................................30

2.6: Group contribution theory and estimation of $\Delta_r G^\circ$ of reactions in *G. sulfurreducens* model: ..................................................................................................................................................31

2.6.a: Lumping Reactions to eliminate metabolites with unknown $\Delta_f G^\circ$: ..................32

2.7: Model Based Deletion Analysis: .....................................................................................33

2.8: Thermodynamic Variability Analysis : .........................................................................34

2.9: Strain and Culture conditions: .......................................................................................35

2.9a: Quenching Metabolism: ...............................................................................................36

2.9b: Flash freezing the samples: .........................................................................................36

CHAPTER 3: RESULTS AND DISCUSSION .............................................................................37

3.1: Prediction of Fluxes under Acceptor and Donor limiting physiological modes: 37

3.2: Influence of Nutrient availability on growth yields: ....................................................39

3.3: Predicting the Threshold Flux and Threshold Concentration for the donor and acceptor at different uptake rates: ..............................................................................45

3.4: Variation in Energy yield, hydrogen secretion with biomass yield: ..........................47

3.4.a: Energy Yield as a function of biomass yield: .............................................................48

3.4.b: Hydrogen Secretion rate as a function of biomass yield: ..........................................49

3.4.c: Fraction of electrons to biomass formation as a function of biomass yield: ........52

3.5. Using TVA to predict candidate reactions for regulation and possible bottleneck reactions : ......................................................................................................................................................54

3.6: Comparing gene expression data with model predicted candidates for regulation: ..................................................................................................................................................59

3.7: Analysis of the MDH activity using the results of TVA: ..............................................63
3.8: *in silico* phenotype predictions: .................................................................64

3.9: Detecting the intracellular concentration of metabolites in *G. sulfurreducens*: .67

3.9.a: Constraining the solution space using absolute metabolite concentrations: ...69

4. CONCLUSIONS AND RECOMMENDATIONS: ..................................................71

5. REFERENCES: ....................................................................................................75

6. APPENDIX: ........................................................................................................79
LIST OF TABLES

Table: 1. A List of the model predicted bottleneck reactions in the metabolic network of *G. sulfurreducens* for optimal growth on acetate as the energy source and fumarate as the sole electron acceptor (p.g:55).

Table: 2. A List of the model predicted bottleneck reactions in the metabolic network of *G. sulfurreducens*, net flux and the exchange factors predicted from C\textsuperscript{13} labelled metabolic flux analysis experiments (p.g:58).

Table.3: *in silico* model predictions of phenotypes on twelve different growth conditions. Incorrect phenotype predictions are indicated as (+). The symbol + indicates presence of growth while the symbol - indicates absence of growth. WT indicates wild type strain (p.g:66).

Table: 4. List of major intracellular metabolites and their detected concentrations in the cell pellet of *G. sulfurreducens* cultured on acetate as the electron donor and fumarate as the sole electron acceptor. A complete set of metabolites detected and their associated concentrations is provided in the appendix section (p.g:68).

Table: 5. List of major intracellular metabolites and their detected concentrations in the cell supernatant of *G. sulfurreducens* cultured on acetate as the electron donor and fumarate as the sole electron acceptor. A complete set of metabolites detected and their associated concentrations is provided in the appendix section (p.g:69).

Table: 6. Minimum and maximum free energy ranges (kcal/mol) for the reaction ALAD\textsubscript{L} computed by the model in the presence and in the absence of concentration constraints (p.g:70).
Table.7: List of model predicted bottleneck reactions, their associated genes and the ranges in free energy under optimal growth using acetate as the electron donor and fumarate as the electron acceptor. Analysis assuming no uncertainty in standard free energy change (p.g:80).

Table.8: List of metabolites and their minimum and maximum activities (mM) at optimal growth with fumarate as the sole acceptor. Analysis assuming no uncertainty in standard free energy change (p.g:81).

Table.9: List of predicted strong candidate’s reactions subject to regulation, their associated genes and the minimum and maximum model predicted free energy changes (kcal/mol) (p.g:82).

Table.10: List of predicted weak candidate’s reactions subject to regulation, their associated genes and the minimum and maximum model predicted free energy changes (kcal/mol) (p.g:83).

Table.11: List of all the metabolites identified using NMR spectroscopy in the cell pellet and their intracellular concentrations (p.g:84).

Table.12: List of all the metabolites identified using NMR spectroscopy in the cell supernatant and their intracellular concentrations (p.g:86).

Table.13: List of simulation parameters used in the lookup table and their associated values (p.g:88).
LIST OF FIGURES

Fig. 1.1: Layout of the *in situ* test plot installed at the Old Rifle UMTRA site at Rifle, CO. (Adapted from Anderson et al. 2003) (p.g:6).

Fig. 1.2: General paradigm for FBA a) A model system consisting of three metabolites A, B and C and three reactions b) Mass balance equations accounting for all reactions and transport mechanisms c) Fluxes of this system constrained based on allowable constraints d) Optimization of the system with different objectives (Adapted from Kauffman et al., 2003). In the figure b₁, b₂ and b₃ are exchange fluxes that can enter or leave the system boundary, while v₁, v₂, v₃ and v₄ are internal fluxes (p.g:12).

Fig. 1.3: A conceptual interpretation for integrating an *in silico* model of microbial metabolism with a subsurface flow and reactive transport model (Adapted from Scheibe et.al., 2009) (p.g:14).

Fig1.4: Redundant pathway for the conversion of malate to oxaloacetate (p.g:16).

Fig.2.1: Flow diagram used to generate the one thousand element look-up table (p.g:27).

Fig.3.1: Predicted fluxes (mmol/gDW hr) through central metabolism of *G. sulfurreducens* during acetate limited and Fe(III)ₙ limited physiological modes (p.g:38).
Fig. 3.2: Frequencies of the biomass yield levels within the one thousand element look-up table generated for **linear** distribution in concentrations of acetate, Fe(III) and NH₄ using the *in silico* model of *G. sulfurreducens* for various environmental conditions. Legend: Fe lim- Fe(III) limited growth, NH₄ lim- NH₄ limited growth, Fe and NH₄ lim- Fe(III) and NH₄ limited growth, Ac lim- acetate limited growth (p.g:41).

Fig. 3.3: Frequencies of biomass yield levels within the one thousand element look-up table generated by the *in silico* model of *G. sulfurreducens* for various environmental conditions using **exponential** distribution of concentrations for acetate, Fe(III) and NH₄. Legend: Fe lim- Fe(III) limited growth, Ac lim-acetate limited growth, Fe lim and N₂ fix- Fe(III) limited and N₂ fixation dependent growth, N₂ fix- N₂ fixation dependent growth (p.g:44).

Fig. 3.4: Prediction from genome scale model of growth rates at different uptake rates of acetate and Fe(III) and for different acetate and Fe(III) concentrations when acetate or Fe(III) availability is the factor limiting growth. a) acetate uptake as a function of growth rate when acetate is limiting nutrient, b) acetate concentration as a function of growth rate when acetate is limiting nutrient, c) Fe(III) consumption rate as a function of growth rate when Fe(III) availability is the limiting nutrient, d) Fe(III) concentration as a function of growth rate when Fe(III) is the limiting nutrient (p.g:46).

Fig. 3.5: Fraction of electrons used for ATP synthesis as a function of biomass yield generated using an exponential distribution in the concentrations for acetate, Fe(III), and NH₄ (p.g:49).

Fig. 3.6: Fraction of electrons used for hydrogen production as a function of biomass yield generated using an exponential distribution in the concentrations for acetate, Fe(III), and NH₄ (p.g:51).
Fig. 3.7: Fraction of electrons used for biomass synthesis as a function of biomass yield generated using an exponential distribution in the concentrations for acetate, Fe(III), and NH₄ (p.g:53).

Fig. 3.8: List of model predicted thermodynamically constrained reactions (bottlenecks) in the metabolic network of *G. sulfurreducens* for optimal growth on acetate as the energy source and fumarate as the sole electron acceptor. Legends are shown as insets in the above figure. Different color schemes are used to show the bounds in the absence and presence of uncertainty (p.g:56).

Fig. 3.9: Thermodynamically feasible activity ranges for all metabolites with a feasible range that is less than the bounds placed on the metabolites activities (0.01mM-60mM) (p.g:59).

Fig. 3.10: Comparison of gene expression fold changes for 21 different conditions against model predicted **strong candidates** for regulation using a threshold covariance of 0.20 (p.g:61).

Fig. 3.11: Comparison of gene expression fold changes for 21 different conditions against model predicted **weak candidates** for regulation using a threshold covariance of 0.20 (p.g:62).

Fig. 3.12: A surface plot showing the variation in the free energy of the MDH reaction corresponding to changes in the malate and oxaloacetate concentrations (mM) (p.g:64).

Fig. 4.1: Gene expression fold changes for the predicted bottleneck reactions (p.g:79).
NOMENCLATURE

Ac: Acetate

Fe(III)_S: Soluble Fe(III) (Ferric Iron)

Fum: Fumarate

Pyr: Pyruvate

NH_4: Ammonium

U(VI): Uranium(VI)

U(IV): Uranium(IV)

Fe(III): Ferric Iron

Fe(II): Ferrous Iron

Uptake Rate: The rate at which a metabolite is consumed by a cell. Expressed in mmol/gDW hr

Consumption Rate: The overall rate of consumption of a metabolite. Expressed in mmol/gDW hr

DOE: Department of Energy

Exchange Flux: the reaction flux that can either enter or leave the system. The metabolite can either move to the inside of the cell from outside or vice versa. Expressed in mmol/gDW hr

FBA: Flux balance analysis

TMFA: Thermodynamics based metabolic flux analysis

mRNA: Messenger RNA (Ribo Nucleic acid)
NMR: Nuclear magnetic resonance spectroscopy

PPCK: Phosphoenol pyruvate carboxykinase reaction

PPDK: Phosphoenol pyruvate di kinase reaction

PPS: Phosphoenol pyruvate synthase reaction

WT: Wild type strain of *G. sulfurreducens*

MDH: Malate dehydrogenase reaction

FUM: Fumarase reaction

Visible cluster: Number of clusters that are visible above the noise level of the NMR instrument.

Confidence level: A combination of sensitivity of the NMR, uniqueness of identifying a particular metabolite and the visible cluster.

UMTRA: Uranium mill tailings remedial action site located in Colorado.

Stoichiometric matrix: A mathematical representation of the metabolic network of an organism. Rows correspond to compounds (metabolites) while columns correspond to the reactions. The entries in the matrix reflect the stoichiometric coefficients. In case of *G. sulfurreducens* this matrix has 611 reactions and 542 metabolites.

Acetate (donor) limiting mode: A condition where the flux of acetate is in excess for a given flux of Fe(III)\textsubscript{s}.

Fe(III)\textsubscript{s} (acceptor) limiting mode: A condition where the flux of the acceptor, Fe(III)\textsubscript{s} is in excess for a given flux of acetate.
MATHEMATICAL SYMBOLS

S: The stoichiometric matrix

m: The total number of metabolites in the metabolic network

n: The total number of reactions in the metabolic network.

c^T v: The objective function for the FBA formulation

q_{Ac}: Uptake rate of acetate (mmol/gDW hr)

q_{Fe}: Fe(III) consumption rate (mmol/gDW hr)

v_i: Flux through a reaction i in the metabolic network of \textit{G. sulfurreducens} (mmol/gDW hr)

v_{max}: The maximum specific uptake rate of a substrate (mmol/gDW hr)

\Delta r_{G_i}^o: Standard Free energy change (kcal/mol)

\Delta r_{G_i}^{\circ o}: Standard free energy change of formation of a metabolite (kcal/mol)

\Delta r_{G_i}^\dagger: Free energy change computed by the TMFA model taking into account the activities of the metabolites (kcal/mol)

R: The universal gas constant (kcal/mol K)

x_j: The activity of compound j.

z_i: A decision variable that can assume values of either 0 or 1 depending on the magnitude of flux through a reaction i in the metabolic network.
LIST OF EQUATIONS

Equation (i). The classical FBA formulation

\[
\textit{Maximize} \ (C^T \ v) \\
\text{s.t.} \ S. \ v = 0 \\
0 \leq v_i \leq v_\infty \\
i = 1, 2\ldots, n
\]

Equation (ii). The constraints used in the TMFA formulation

\[
S. \ v = 0 \\
0 \leq v_i \leq Z_i v_{\text{max}}, \{i = 1, \ldots, n\}, \\
\Delta_r G_i' - K + KZ_i < 0, \{i = 1, \ldots, n| \Delta_r G_i^o \text{ is known}\}, \\
\Delta_r G_i^o + RT \sum_{j=1}^{m} n_{ij} \ln(x_j) = \Delta_r G_i', \{i = 1, \ldots, n + L| \Delta_r G_i^o \text{ is known}\}, \\
\Delta_r G_i' - Ky_i < 0, \{i = n, \ldots, n + L\}, \\
y_i + \sum_{j=1}^{m} \alpha_{ij} z_j \leq \sum_{j=1}^{m} \alpha_{ij} \{i = n, ..., n + L\},
\]

Equation (iii). Estimating the \(\Delta G^o\) of the entire compound using group contribution method

\[
\Delta_f G_{\text{est}}^o = \Delta_{gr} G_{\text{o}}^o + \sum n_i \Delta_{gr} G_i^o
\]
Equation (iv). Calculating the energy yield from uptake and consumption rates of substrates

\[ \text{Energy Yield} = \frac{q_{Fe}}{8q_{Ac}} \]

Equation (v). Calculating the fraction of electrons used for hydrogen production

\[ \text{Fraction of electrons to Hydrogen} = \frac{2q_{H2}}{8q_{Ac}} \]

Equation (vi). Calculating the fraction of electrons used for biomass synthesis

\[ \text{Fraction of electrons to biomass} = 1 - (\text{Energy yield} + \text{Fraction of electrons to Hydrogen}) \]

Equation (vii). Uptake rate of acetate in the model-based lookup table based on Michaelis-Menten kinetics

\[ R_{Ac} = \frac{V_{max1}[Ac]}{K_1 + [Ac]} + \frac{V_{max2}[Ac]}{K_2 + [Ac]} + \frac{V_{max3}[Ac]}{K_3 + [Ac]} \]

Equation (viii). Uptake rate of Fe(III) in the model-based lookup table based on Michaelis-Menten kinetics

\[ R_{Fe} = \frac{V_{maxFe}[Fe]}{K_{Fe} + [Fe]} \]

Equation (ix). Uptake rate of NH\(_4\) in the model-based lookup table based on Michaelis-Menten kinetics

\[ R_{NH4} = \frac{V_{maxNH4}[NH4]}{K_{NH4} + [NH4]} \]
THESIS OUTLINE

This is a Masters of Applied Science thesis that conforms to the guidelines of the Department of Chemical Engineering and Applied Chemistry, University of Toronto. The first part of this thesis focuses on the development of a model based lookup table to investigate the dependence of cell physiology of *G. sulfurreducens* on environmental perturbations. The lookup table was developed by using a previously published model of *G. sulfurreducens* (Mahadevan et al., 2006). A portion of this work is adapted from the journal paper “**Coupling a genome scale metabolic model with a reactive transport model to describe in situ uranium bioremediation**” which was published in the Microbial Biotechnology journal in March, 2009. Srinath Garg is a co-author of this paper and the results described herein are his contribution to the manuscript. The second part of the thesis focuses on developing a thermodynamics based metabolic model for *G. sulfurreducens* by adapting a previously published framework for *E. coli* (Henry et al., 2007) and using this updated model to predict potential bottlenecks in the metabolic network and reactions subject to regulatory control. This part of the thesis has been drafted as a manuscript titled “**Analysis of regulation in the central metabolism of G. sulfurreducens using constraint based modeling**” and is intended to be submitted to the BMC Research Notes journal shortly. Srinath Garg is the primary author of this manuscript.

The main body of the thesis has four chapters. Chapter 1 gives a detailed introduction to uranium contamination in the subsurface and suitable management alternatives as well as the motivation to develop a suitable computational framework in order to investigate the growth and activity of *G. sulfurreducens*.
The importance of integrating thermodynamic constraints with an existing constraint-based model and the need to incorporate absolute metabolite concentrations in our model-based analysis is detailed in this chapter. The statement of objectives is listed at the conclusion of this chapter. Chapter 2 details the various methods (computational and experimental) used in this work. Chapter 3 describes and discusses the simulation and experimental results, while Chapter 4 summarizes the results of this work and provides a guide for future directions.
CHAPTER 1: INTRODUCTION

1.1: Literature Review:

1.1. a: Nature of Bioremediation—Overview of the Problem:

Metals (e.g., chromium, mercury) and radionuclides (e.g., uranium, strontium) are major subsurface contaminants in many US Department of Energy (DOE) facilities. These elements are long lived due to their long half-life periods and possess high mobility in subsurface environments, posing a potential risk to humans and the natural environment. In-situ bioremediation is being evaluated as a potential technology for remediation of uranium at DOE sites. This method focuses on the microbial mediated reduction of uranium from the completely soluble U(VI) valence state to the relatively insoluble U(IV) valence state. Since elemental uranium cannot be degraded into daughter products, current research focuses on making use of this radionuclide as an electron acceptor in microbial respiratory processes (Scheibe et al. 2006). Previous studies have suggested that dissimilatory metal reducing bacteria are capable of reduction of soluble U(VI) to insoluble U(IV) and this has subsequently been shown to be an enzymatically driven process (Lovley et al. 1991). Subsequent laboratory experiments have demonstrated the direct enzymatic reduction of U(VI) (Suzuki et al. 2002, Lovley et al. 1995, Holmes et al. 2002). These studies also showed that an electron donor such as acetate was effective in stimulating metal reducers especially those belonging to Geobacter species.

These results suggested that biostimulation (the modification of the environment, including nutrient addition, to stimulate existing bacteria capable of bioremediation) may be a viable method for treatment of uranium contaminated sites. Thus, in-situ bioremediation can be implemented by introducing a soluble electron donor (e.g., acetate), thereby stimulating the in-
situ activity of dissimilatory metal reducing organisms such as *G. sulfurreducens*. Bio stimulation experiments conducted at the Old Rifle UMTRA site in Western Colorado showed significant removal of hexavalent uranium [U(VI)] in the shallow alluvial aquifers (Anderson et al. 2003, S.B. Yabusaki et al. 2007). Anderson et al. (2003) showed that the U(VI) concentrations in the subsurface could be lowered below prescribed standards by stimulating the activity of dissimilatory metal breathing bacteria with acetate as the electron donor. The reduction of U(VI) to U(IV) was shown to be coupled to Fe(III) reduction (Anderson et al. 2003), leading to an accumulation of Fe(II) in the subsurface and enrichment of *Geobacteraceae* (S.B. Yabusaki et al. 2007). Furthermore, when bio-available Fe(III) was depleted near the point of acetate amendment, sulphate reducers were seen to out-compete Fe(III) reducers such as *G. sulfurreducens* as evidenced by the increase in the sulphide concentrations after 45 days of acetate injection. This increase in sulphide concentration was coupled to an increase in U(VI) concentration in the subsurface (S.B. Yabusaki et al. 2007). It was further shown that the acetate oxidizing sulphate reducers were not as effective as Fe(III) reducers at U(VI) immobilization (Anderson et al. 2003). Hence it becomes imperative to keep Fe(III) reduction active and not inject excess of acetate into the subsurface. This in turn calls for a thorough understanding of the field dynamics and the factors controlling U(VI) remediation at the site.

1.2: **Field Dynamics:**

In particular, it is well established that microbial community dynamics are largely controlled by the availability of electron donor, electron acceptor and nutrients, and that long term bioremediation performance is insensitive to initial microbial population (the presence of
Dehalacoccoides needed for anaerobic biodecay of chlorinated solvents is an exception. Thus biostimulation of indigenous organisms is generally preferred over bioaugmentation. (Emily N Devillier, Master of Science, 2006). Fig:1.1 shows the layout of a Uranium contaminated aquifer located at the Old Rifle site in Colorado for the 2003 study by Anderson et al. This aquifer was selected to evaluate the potential ability of G. sulfurreducens in removing hexavalent uranium from groundwater. An electron donor such as acetate (1-3 mM) was injected into the subsurface through an injection gallery with 20 injection wells over a time period of 3 months. The injection wells were installed up gradient of 15 monitoring wells. It was shown that in 9 days, the U(VI) concentration had dropped below the prescribed treatment level indicating efficient reduction of U(VI) to U(IV) coupled with enrichment of Geobacter species. Throughout the acetate amendment, the composition of the microbial community changed several times with a period of enrichment of dissimilatory iron reducing bacteria (DIRB) being followed by a phase of sulphate reducers dominating the subsurface. These results indicated that in situ remediation of U(VI) is feasible but the strategy needs better optimization to maintain the long term activity of G. sulfurreducens (Anderson et al., 2003). This calls for developing rigorous models that can accurately predict the response of Geobacter species in response to dynamic changes in their growth environment.
Fig. 1.1: Layout of the *in situ* test plot installed at the Old Rifle UMTRA site at Rifle, CO. showing the locations of the injection galleries and the observation wells (Adapted from Anderson et al. 2003).
1.3: The Challenges in implementing *in situ* bioremediation:

While the processes governing U(VI) behaviour during acetate biostimulation are understood in a general mechanistic sense on the basis of earlier laboratory studies and the 2002 and 2003 field scale experiments, (S.B.Yabusaki et al. 2007), control of Uranium bioremediation at the site and engineering long term strategies is still an open issue. The inability to predictively model the growth and activity of microorganisms in the subsurface, limits predicting the fate and transport of contaminants in the subsurface (Scheibe et al. 2006). The challenge is therefore being able to identify and quantify the effect of key growth parameters (Ac, Fe (III), and NH₄) on the growth yield and the cell physiology of *G. sulfurreducens*, as the concentration of these three growth factors vary the most during *in situ* bioremediation of U(VI). This would help in designing efficient bioremediation strategies prior to field trials. The design of successful management alternatives and field strategies has been greatly aided by the development of several codes that simulate groundwater flow, solute transport and geochemical reactions such as the reactive transport modeling code-HYDROGEOCHEM (Scheibe et al., 2006, S.B.Yabusaki et al., 2007, Steefel et al., 2005, Yeh et al., 2004). Over the last 20 years, advances in computer hardware and software have made field-scale simulations of hydrologic processes coupled with complex chemical and biological reaction networks computationally feasible.

Computer codes are now available that are capable of modeling three dimensional transient multiphase fluid flow under non-isothermal conditions and transport of multiple chemical species subject to kinetically coupled or equilibrium controlled reactions associated with various physical, chemical and biological processes (Yeh et al., 2004, Fang et al., 2003). In addition, detailed models of microbial metabolism can be applied to bioremediation due to the
availability of completely sequenced genomes of microorganisms (Lovley 2003, Lovley et al. 2008). Bioremediation of metals and radionuclides is a less mature technology as it involves reactions that reduce the contaminant mobility by potentially reversible precipitation, sorption or complexation reactions that involve complex interactions between geochemical and transport processes. In addition, the kinetics of remobilization is another concern. Models for microbial growth and reaction kinetics have been described for Fe, Mn and other trace metals in pure and mixed cultures (Truex et al., 1997; Spear et al., 1999; Liu et al., 2002; Lim et al., 2007). Several field scale applications of coupled flow transport and biogeochemical reaction models that describe microbial growth in terms of Monod kinetics have been recently reported (Scheibe et al., 2006; Yabusaki et al., 2007; Luo et al., 2008). However the following factors are seen to limit the ability of current models to predict the activity of microbial community accurately.

1. Influence on the microbial metabolism due to a diversity of environmental conditions.

2. Optimization of growth and activity in a given environment by adaptation of microorganisms metabolism and the complexity of the subsurface environment further magnified by hydrologic flow and mass transport and biogeochemical reactions creating a spatially variable environment.

1.4: The Need for Constraint Based Modeling:

The above factors demand that a model be able to accurately predict the microbial growth and activity when dealing with dynamic conditions that are commonplace when microbial activity is stimulated artificially (addition of a suitable amendment such as acetate). Knowledge
of genes, their function and expression make it theoretically possible to construct a reaction based model for the biochemical mechanism involved in cellular metabolism. However, experimental validation of kinetic parameters for a large number of metabolites is infeasible. An alternative approach is the constraint-based modeling which involves finding a steady state flux distribution through all biochemical reactions (Price et al., 2004, Reed et al., 2006, Feist and Palsson 2008). The constraints include the following:

1. Mass Balance: the rate of production and consumption of intracellular metabolites are equal at steady state.

2. Thermodynamic: Irreversibility of reactions.

3. Enzymatic capacity: Include bounds on the enzyme rates.

4. Availability of Nutrients.

In this method, the potential candidate reactions are mapped based on the genome sequence and other omics data such as fluxomics (deals with the physiological state of a cell, time dependent fluxes), proteomics etc. Then the cellular objective such as maximization of growth rate is assumed to be the objective as this objective gives a growth advantage to G. sulfurreducens.

In most cases maximization of biomass (i.e., growth) is assumed to be the objective, although other optimization choices such as maximization of the ATP production rate and minimization of resource utilization could also be used (Schuster et al., 2007) Thus, we can formulate a linear optimization program as follows:
Maximize $(c^T v)$

\[ s.t. \ S.v = 0 \]

\[ 0 \leq v_i \leq v_\infty \]

\[ i = 1, 2\ldots, n \]

Where $S$ is the stoichiometric matrix- a network topology with the stoichiometry of all the biochemical reactions in the model. The dimension of the $S$ matrix is $m \times n$, where $m$ is the total number of metabolites in $S$ and $n$ is the total number of reactions in $S$. $V_i$ is the flux through reaction $i$. Typically $m$ is less than $n$, leading to an underdetermined system (possibility of plurality of solutions due to more fluxes than metabolites). The above equation forms the basis of the classic Flux Balance Analysis (FBA) (Varma and Palsson 1994). The FBA is useful as it facilitates a quantitative interpretation of metabolic physiology of a micro organism, for practical applications. The FBA is also useful for the stoichiometric analysis of metabolic flux distributions and requires information on relatively fewer parameters such as the stoichiometry of the metabolic reactions, metabolic requirements for growth and certain strain specific parameters such as the maintenance energy requirements. A conceptual diagram of the flux balance approach is given in Fig.1.2. Flux balance models are based on the assumption that the metabolic transients in a cell are rapid when compared to the cellular growth rate and the dynamic nature of an organism’s environment.
The metabolic transients are typically shorter than a few minutes and hence the metabolic fluxes are assumed to be in a quasi steady state with respect to the growth and process transients. This assumption leads to the result that all metabolic fluxes leading to the formation and degradation of any metabolite are balanced (Eq.1). In the FBA, the number of fluxes exceeds the number of metabolites, making the system underdetermined. Due to the underdetermined nature of the model, pluralities of solutions are possible and the cell has infinitely many choices to distribute its metabolic fluxes. The choices are however constrained by the stoichiometric matrix. This domain defines the so called ‘metabolic solution space’. Within this domain, a particular flux distribution can be chosen based on assumed behaviour such as maximal growth rate or ATP production rate which are distinct user-defined objective functions. Then, the optimal flux distribution so calculated using linear programming methods represent the strains ‘metabolic phenotype’.
Fig 1.2: General Paradigm for FBA a) A model system consisting of three metabolites A, B and C and three reactions b) Mass balance equations accounting for all reactions and transport mechanisms c) Fluxes of this system are constrained based on allowable constraints d) Optimization of the system with different objectives (Adapted from Kauffman et al., 2003). In the figure $b_1$, $b_2$ and $b_3$ are exchange fluxes that can enter or leave the system boundary, while $v_1$, $v_2$, $v_3$ and $v_4$ are internal fluxes.
Constraint-based models have shown to accurately predict the microbial yield under different environmental conditions (Reed et al. 2006). Since this approach does not assume constant yield coefficients, it is well suited for modelling microorganisms in heterogeneous environments, and accounts for changes in metabolic network in response to nutrient limitation (Varma et al., 1993). This approach has also been extended to model the diauxic growth (the diphasic response of a culture based on phenotypic adaptation to the addition of a second substrate) of *E.coli* (Mahadevan et al., 2002). Previous studies have shown that a consortia of organisms can be effectively modeled using the constraint based modelling approach. In this work, an example of a subsurface remediation scenario essentially catalyzed by a single microbial genus (*G. sulfurreducens*) is considered. The initial step is including a constraint-based microbial model in a comprehensive hybrid model to adequately describe *in situ* uranium bioremediation. To accomplish this step, a previously published model of *G. sulfurreducens* was used (Mahadevan et al., 2006). This model represented the growth and activity of *G. sulfurreducens* under various conditions (Mahadevan et al., 2006, Segura et.al., 2008) and was a valuable tool for developing strategies for increasing the respiration rate of *G. sulfurreducens* for fuel cell applications (Izallalen et al., 2008). Furthermore, since the subsurface is a spatially variable environment, accurate description of the subsurface contaminant transport calls for the integration of the constraint based *in silico* model with the subsurface flow and reactive transport model that include equilibrium, complexation and biogeochemical reactions happening in the subsurface. This integration can be implemented in two distinct ways as illustrated in Fig.1.3.
Fig.1.3: A conceptual interpretation for integrating an *in silico* model of microbial metabolism with a subsurface flow and reactive transport model. (Adapted from Scheibe et.al., 2009)

The above schematic describes the two methods to integrate an *in silico* model of microbial metabolism with a flow and reactive transport model. Either direct calls can be made to the *in silico* model during the execution of the flow and reactive transport model, or, alternatively, the *in silico* model itself can be solved off-line for a number of possible scenarios to create a so called lookup table which can then be referenced at each time step and grid cell during the execution of the flow and the reactive transport model. The direct integration makes the *in silico* model a callable routine from the reactive transport model to replace functions that perform rate calculations for important microbial mediated reactions. Direct integration is computationally intractable, as a separate linear programming problem needs to be implemented for each cell grid at each time step. In a typical field scale simulation involving spatial and temporal variations in bio-geochemical parameters, the number of such grid cells can be very high.
Hence, in the present work, an alternative approach based on the development of a look-up table is pursued. This model is also shown to adequately describe the change in the metabolic state of *G. sulfurreducens* in response to change in concentration of a combination of growth factors.

### 1.5: Limitations of FBA-Need for TMFA:

The Flux balance approach is useful in identifying the metabolic flux distribution for a given cellular objective. However, as metabolite concentrations are not incorporated in the FBA approach, it is incapable of representing the changes in the metabolome (Varma et al., 1994). Previous *in vivo* and *in silico* studies on *G. sulfurreducens* revealed the presence of several redundant pathways (Segura et al., 2008). This redundancy in the central metabolism of *G. sulfurreducens* was attributed to either the presence of alternate biochemical pathways (Fig:1.4) or due to gene duplicates which act as isozymes for the same enzymatic function. Alternate biochemical pathways provide compensation through another pathway that is stoichiometrically or energetically equivalent.
Previous model based comparison of the *in silico* and *in vivo* growth phenotypes using the FBA and gene deletion approach resulted in 16 cases of incorrect prediction (Segura et al., 2008). However, incorporation of further constraints on the activity of these redundant pathways resulted in only 8 false positive cases. Of these, five false positives were identified for the PPCK (phosphoenol pyruvate carboxykinase) mutant grown on Fe (III) citrate as the electron acceptor and acetate as the electron donor. This prediction was primarily attributed to the thermodynamical infeasibility of the MDH (Malate dehydrogenase reaction). The MDH reaction has a standard free energy change of +30 kJ/mol. Thus, a high malate to oxaloacetate ratio is required for the free energy change to be negative and for the reaction to be feasible.

These results indicate that the FBA approach might not be best suited to predict growth phenotypes and also show the need to include thermodynamic data in the analysis of the metabolism of *G. sulfurreducens*. 

Fig. 1.4: Redundant pathway for the conversion of malate to oxaloacetate
Furthermore, Flux balance analysis solutions have been shown to generate internal flux cycles (Beard et al., 2002). These are defined as reactions of the form $A \rightarrow B \rightarrow C \rightarrow A$.

According to the first law of thermodynamics, the net flux through these cycles must be zero (Henry et al., 2007). These limitations of the FBA approach calls for providing tighter constraints on the system. One such type of constraint is the thermodynamic constraint of feasibility of a biochemical reaction.

Although, thermodynamics have been used in a variety of applications in the analysis of biological systems, they are yet to be applied for examining entire metabolic networks due to scarcity in thermodynamic data and the existence of a database to integrate the thermodynamic data with the metabolic chemistry.

However, the advent of the group contribution method has made it possible to accurately predict thermodynamic properties such as the free energy (Mavrovouniotis et al., 1990, 1991). Furthermore, the availability of thermodynamic data has grown over the time (Benson et al., 1968). Recently, several genome scale models of organisms have been developed (Feist et al., 2008, Palsson et al., 2008.).

Previous studies of the lactic acid fermentation pathway by Mavrovouniotis found that pathways which were feasible could be falsely labelled as infeasible pathways in the absence of thermodynamic data, ionic strength of a solution and cell pH. Thus it becomes imperative to incorporate thermodynamic data in the analysis of metabolism of $G. sulfurreducens$. 

[17]
Here, a new methodology called the Thermodynamics based metabolic flux analysis (TMFA) is proposed for integrating thermodynamic data and constraints into a constraint based metabolic model of \textit{G. sulfurreducens} to ensure that flux distributions generated by the model are thermodynamically feasible.

We hypothesize that the inclusion of thermodynamic constraints would increase the predictive ability of the model to predict the lethality of deletion mutants, specifically the PPCK mutant. This knowledge would in turn, be helpful to study the physiology of these mutants and would also be helpful in understanding the role of key central metabolic pathways (acetate assimilation, Gluconeogenic and Anapleurotic pathways).

\textbf{1.6: Thermodynamics –Based Metabolic Flux Analysis (TMFA):}

The TMFA uses the FBA (Flux Balance Analysis) as its core. It also uses a set of linear thermodynamic constraints in addition to the mass balance constraints used in the FBA. The TMFA provides information on the range of metabolite activity profiles and the free energy change of a reaction along with the reaction fluxes (Henry et al., 2007) and hence can be used to generate flux distributions that are free of any thermodynamic infeasibility.

Linear thermodynamic constraints have previously been used in conjunction with FBA to eliminate thermodynamically infeasible flux loops (Beard et al., 2002, Qian et al., 2003, Beard et al., 2004). Non linear constraints have been utilized to eliminate flux distributions that utilize reactions that cannot be thermodynamically feasible under physiological conditions (Beard et al., 2005). The application of these constraints to large scale systems is computationally intractable.
Thus, in the TMFA, a set of mixed integer linear constraints is used to produce flux distributions that are thermodynamically feasible and in addition provide information on the feasible metabolite activity profiles and $\Delta_r G_i^\prime$. (The Gibbs free energy change of a reaction in the metabolic network). The constraints used in the TMFA are as below:

\begin{align*}
S \cdot v &= 0 \\
0 &\leq v_i \leq Z_i v_{\text{max}}, \{i = 1, \ldots, n\},
\end{align*}

\begin{align*}
\Delta_r G_i^\prime - K + KZ_i &< 0, \quad \{i = 1, \ldots, n| \Delta_r G_i^o \text{ is known}\},
\end{align*}

\begin{align*}
\Delta_r G_i^o + RT \sum_{j=1}^{m} n_{ij} \ln(x_j) &= \Delta_r G_i^\prime, \quad \{i = 1, \ldots, n + L| \Delta_r G_i^o \text{ is known}\},
\end{align*}

\begin{align*}
\Delta_r G_i^\prime - Ky_i &< 0, \quad \{i = n, \ldots, n + L\},
\end{align*}

\begin{align*}
y_i + \sum_{j=1}^{m} \alpha_{ij} z_j &\leq \sum_{j=1}^{m} \alpha_{ij}, \quad \{i = n, \ldots, n + L\},
\end{align*}

In the above Eq.1, $S$ is the stoichiometric matrix having the dimension $m \times n$; $v$ is the $n \times 1$ flux vector. $m$ equals the total number of metabolites in $S$ and $n$ equals the total number of reaction in $S$. The value of $n$ is typically greater than the number of reactions in the model as each of the reversible reactions needs to be split up into separate irreversible forward and backward component reactions during the construction of $S$. The result of this separation has the effect of constraining the flux through each reaction $i$ to be greater than or equal to zero. $Z_i$ is a decision variable associated with every reaction $i$ in the stoichiometric matrix $S$. The value of $Z_i$
is zero when the flux through the reaction $i$, $V_i$ is zero and one if the flux through reaction $i$ is positive. This is reflected in Eq.2. $V_{\text{max}}$ in Eq.2 is the upper bound on the flux through any reaction $i$. This is typically set to a physiologically reasonable value such as 100 mmol/gDW hr. Eqs 1 and 2 are the mass balance constraints encountered in the classical FBA setup with the only change being the separation of a reversible component into irreversible forward and backward reactions.

Eq.3 is reflective of the second law of thermodynamics, according to which the Gibbs free energy change of a reaction needs to be negative for a reaction to be thermodynamically feasible. This constraint ensures that the flux distribution and the activity profiles generated are free from any thermodynamic infeasibility and adhere to the second law of thermodynamics. $K$ in Eq.3 is a constant whose value is large enough so that Eq.3 is satisfied if $Z_i$ and $V_i$ are zero.

This constraint is applied to only those reactions with a non zero flux and ensures that the $\Delta_r G < 0$ for such reactions. Eq.4 is the Gibbs free energy equation used to set the value of $\Delta_r G^o$ of a reaction $i$, provided the metabolite activities are known. $\ln(x_j)$ represents the activity of compound $j$ and $n_{ij}$ is the stoichiometric coefficient of compound $j$ in reaction $i$. This constraint also accounts for any transport of ions across the cell membrane. Since $\Delta_r G^o$ must be known to formulate the thermodynamic constraints for a reaction, Eqs. 3 and 4 cannot be applied to reactions whose standard free energy change is unknown. Instead, these reactions are lumped with reactions whose free energy change is known. Special constraints are applied to reactions whose $\Delta_r G^o$ is unknown. These are given by Eqs 4-6. $L$ is the number of lumped reactions and $\alpha_{ij}$ is a coefficient that equals one if reaction $j$ is one of the original reactions with unknown $\Delta_r G^o$ that makes up the lumped reaction $i$. Eq.5 is the thermodynamic feasibility constraint for the
lumped reaction except that the decision variable $y_i$, is equal to one if the reaction is infeasible and zero if the reaction is feasible.

Eq.6 eliminates those flux distributions that involve flux through a set of reactions that comprise an infeasible lumped reaction. The continuous independent variables in this optimization are $\ln(x_j)$, $v_i$ and $\Delta G_i^-$. The binary independent variables are $y_i$ and $Z_i$. In this work, the TMFA was used to predict the lethality of deletion mutants under different environments and compare the results with FBA predictions. These constraints help in quantifying the fluxes in the cell metabolism simultaneously with the assessment of thermodynamic feasibility. This eliminates the possibility of getting flux distributions that need to be corrected later due to infeasibilities.

### 1.7: Metabolomics of *G. sulfurreducens*:

The term metabolomics refers to both the qualitative and quantitative analysis of low molecular weight metabolites present in and around growing cells at a given time during their growth or production cycle (Mashego et al., 2007). Metabolomics is seen to complement emerging research areas such as new drug discovery and metabolic engineering. The quantification of metabolites is mainly carried out by analytical methods such as chromatographic techniques such as LC-MS, GC-MS and even NMR. It is important to quantify the intracellular metabolite concentration of metabolites. Previous studies have shown that the metabolite concentration plays a direct regulatory role. For example, the feedback inhibition mechanism of enzymes by metabolite concentration is a common control mechanism in amino acid biosynthesis pathways (Oliver et al., 1998). Furthermore, successful engineering of
microbial cells for enhanced product production requires a thorough knowledge of primary metabolism and its regulation \textit{in vivo} (Bailey, 1991). Metabolomics constitutes reactants and products of \textit{in vivo} reactions directly related to phenotypes rather than transcriptomics (mRNA) and proteome (proteins). It has also been shown that identification of metabolite levels in a cell can lead to identification of bottlenecks in a metabolic network (Theobald et al., 1993).

Hence, quantitative understanding of metabolism and its regulation demands the quantification of intracellular and extracellular metabolites. These results make it feasible to consider determining the intracellular concentration of metabolites in \textit{G. sulfurreducens}. In this work, in an attempt to improve the TMFA model predictions and pose experimentally determined concentrations as constraints to the model, thereby shrinking the solution space even further, intracellular concentration of metabolites in WT \textit{G. sulfurreducens} cells cultured on acetate as the electron donor and fumarate as the sole acceptor were determined using 1D NMR spectroscopy. It is hypothesized that this TMFA model would enable better understanding of the \textit{in vivo} regulation of metabolism in \textit{G. sulfurreducens}.
STATEMENT OF OBJECTIVES

In summary, *Geobacter sulfurreducens* is a well studied representative of the *Geobacteraceae* family of micro-organisms that have a unique mode of metabolism and are capable of utilizing metals such as Fe(III) as their electron acceptor. Stimulating the dissimilatory activity of *G. sulfurreducens* by introducing a soluble electron donor such as sodium acetate has been shown to result in the removal of radioactive contaminants such as U(VI) from the subsurface environment. Thus, the efficiency of subsurface remediation can be enhanced through an improved understanding of the physiology of *G. sulfurreducens*.

- Consequently there is a need to understand how the physiology of *G. sulfurreducens* changes in response to environmental perturbations during *in situ* bioremediation of U(VI). Thus, it is hypothesized that the constraint based modeling approach can predict important changes in the fluxes of substrates that can result from changes in environmental conditions and the associated difference in growth yield can give an insight on the impact and the extent of cell growth in the subsurface.

- It is hypothesized that the inclusion of thermodynamic constraints in a constraint based model of metabolism would increase the predictive ability of the model to predict the lethality of deletion mutants (mutants that lack the activity of a particular enzyme such as PEP carboxy kinase), specifically the PPCK mutant. This knowledge would in turn, be helpful to study the physiology of these mutants and would also be helpful in
understanding the role of key central metabolic pathways (acetate assimilation, Gluconeogenic and Anapleurotic pathways).

- It is also hypothesized that thermodynamic constraints would permit the detection of potential thermodynamic bottlenecks in the metabolic network and would give insight into the ability of a reaction in the metabolic network to be subject to regulation.

Finally, it is hypothesized that inclusion of measured intracellular metabolite concentrations in the TMFA model would increase the predictive ability of the model by further shrinking the solution space, thereby enabling better understanding of regulation in vivo in G. sulfurreducens.
CHAPTER 2: MATERIALS AND METHODS

In this section, the method used to develop a novel algorithm to generate the pre-calculated lookup table based on key reactants (acetate, Fe(III) and ammonium) is illustrated. This is followed by the methods used to identify the threshold fluxes and concentrations of the donor and acceptor under donor and acceptor limiting conditions and for simulating the in silico metabolic flux distribution under these two different physiological modes. This analysis helps in identifying the threshold flux and the concentration below for the donor and acceptor below which growth becomes infeasible.

We also report on the group contribution method used to determine the standard free energy change of reactions in the metabolic network of G. sulfurreducens. Details of in silico gene deletion studies used to predict mutant phenotypes on different combination of substrates (electron donor and electron acceptor) by incorporating thermodynamic and mass balance constraints are explained. This is followed by details on the thermodynamic variability analysis used to identify potential bottlenecks in the metabolic network and reactions that can serve as a regulatory control point. Next, these model based predictions are compared against experimental gene expression data obtained under 21 different conditions. Finally, details of the strain and culture conditions used for culturing G. sulfurreducens on acetate/fumarate is explained along with details of the protocols used for quenching metabolism.
2.1: Generating the Pre Calculated Lookup table:

The first step in integrating the constraints based in silico model with the reactive transport model is the generation of a pre calculated lookup table. The steps involved in generating the lookup table are as follows:

1. Uptake rates were calculated as a function of concentration for selected key reactants (Ac, Fe (III), NH₄). Although other reactants such as PO₄⁻, SO₄²⁻ can also influence the growth yield and activity of G. sulfurreducens, these reactants were selected as their concentrations varies the most during in situ U(VI) remediation.

2. For each combination of the selected key reactants, uptake rates were posed as constraints and the metabolic model was executed to define the yield and other relevant fluxes.

3. The fluxes and the resulting concentration of the key reactants were stored as a lookup table (a matrix composed of the fluxes and concentrations of substrates for the one thousand possible combination of growth factors) for reference at each time step and grid cell.

The generation of the lookup table involved choosing ten different concentrations of acetate and Fe(III), and NH₄, respectively, resulting in one thousand different possible combinations. The maximum possible concentration for these reactants was chosen based on earlier field scale experimental data. For each of these cases, the uptake rates were calculated and imposed as constraints and maximized the growth to obtain the optimal metabolic flux distribution. A novel algorithm was developed to solve the lookup table problem (Fig:2.1.)
Fig. 2.1: Flow diagram used to generate the one thousand elements Lookup Table.

The algorithm starts with defining the FBA problem for *G. sulfurreducens*. Then, we proceed to fix the substrate (acetate) uptake rate. This is primarily done to simulate the acceptor (Fe (III)) limiting scenario, as when the acetate uptake rate is fixed, for a given flux of Fe(III), the flux of acetate is in excess. This condition is hence reflective of the scenario when the acceptor Fe(III) limits metabolism. The linear program was then solved.

The resulting solution can either be feasible or infeasible. Infeasibilities primarily arise due to excess flux of acetate forced into the cell, than that can be used by *G. sulfurreducens*. We practically fix the acetate uptake rate in the model to force excess acetate flux into the cell. The
other reason contributing to an infeasible solution is the inability of *G. sulfurreducens* to meet its maintenance energy requirements. If the solution is found to be feasible, we proceed to compute the nitrogen and ammonium dependent growth rate. The ammonium dependent growth rate was calculated by making ammonium exchange reversible and forcing nitrogen exchange to zero. The nitrogen dependent growth, likewise, was calculated by forcing ammonium exchange to zero and setting nitrogen exchange to be reversible. It is well known that below a certain threshold ammonium concentration, nitrogen fixation is possible and this involves the expression of ammonium uptake protein. The two growth rates were then compared and the higher growth rate was identified consistent with the cellular objective of growth maximization. The biomass yield and other fluxes were then defined based on the higher growth rate. However, if the solution is found to be infeasible after solving the linear program for the first time, we proceed to free the upper bound on the acetate uptake and then resolve the linear program. The solution thus obtained can again be feasible or infeasible. If the solution is found to be feasible, then the nitrogen and ammonium dependent growth rates are computed. The growth rates are then compared and the highest growth rate is outputted. Infeasibilities can also arise when resolving the linear program. These infeasibilities are attributed to the insufficient acetate uptake to meet the Fe(III)<sub>s</sub> consumption rate and the inability of *G. sulfurreducens* to meet the maintenance energy requirements.

**2.2: In Silico Model Simulations:**

Since the FBA model requires the uptake rate of the limiting nutrient as an input, the uptake rates of the key nutrients (acetate, Fe (III)<sub>s</sub>, and NH<sub>4</sub>) were specified using the Michaelis-Menten kinetics. The maximum specific uptake rate parameters (*V*<sub>max</sub>) in the Michaelis Menten
uptake kinetics were set to be 100 mmol/gDW hr for Fe(III) and 0.1 mmol/gDW hr for NH₄. The half saturation constants used were 1mM and 0.1 mM for Fe(III) and NH₄ respectively. These parameters were chosen so that the growth rates were consistent with microbial physiology (Esteve-Nunez et al., 2004; Esteve-Nunez et al., 2005). Since *G. sulfurreducens* has been shown to have three acetate transporters, the uptake rate of Acetate was specified as a summation of three separate equations with different maximum specific uptake rates and half saturation constants. The values used were \( V_{\text{max}1} = 3 \text{ mmol/gDW hr} \), \( V_{\text{max}2} = 0.6 \text{ mmol/gDW hr} \) and \( V_{\text{max}3} = 1.8 \text{ mmol/gDW hr} \) respectively. The corresponding half saturation constants used were \( K_1 = 0.161 \text{ mM} \), \( K_2 = 0.008 \text{ mM} \) and \( K_3 = 0.034 \text{ mM} \) (Richter et al., submitted).

**2.3: Identifying the Infeasibility Thresholds:**

Identifying threshold concentrations of the substrates (electron donor, electron acceptor) is important to identify the minimum concentration of these substrates that would provide sufficient ATP to meet the maintenance energy demand.

In order to identify the threshold acetate flux and threshold acetate concentration during acetate limiting conditions, the flux of acetate was varied from 0 mmol/gDW hr to 10 mmol/gDW hr, as the maximum specific uptake rate of acetate was set to 10 mmol/gDW hr in the lookup table, while allowing an excess of Fe(III) so that the donor limits metabolism in this case. The Fe(III) consumption rate was fixed at 1000 mmol/gDW hr.

Similarly, the threshold Fe(III) flux was identified under acetate excess conditions by allowing unlimited acetate uptake while constraining the Fe(III) consumption rate between 0 and 100 mmol/gDW hr and allowing excess influx of acetate so that Fe(III) limits metabolism in this case.
2.4: Simulating the Metabolic Flux Distributions:

The metabolic flux distribution for Fe (III) and acetate limiting conditions were simulated using a previously published metabolic model of *G. sulfurreducens*. For acceptor limiting scenario, the Fe (III) consumption rate was constrained to be lower than 50 mmol/gDW hr. For acetate limiting scenario, the acetate uptake was fixed to 10 mmol/gDW hr so as to simulate acetate limitation.

2.5: Thermodynamics based Metabolic Flux Analysis:

TMFA uses the mass balance constraints of the metabolic flux analysis. FBA is helpful in studying the metabolic capabilities of cellular systems by constraining the net production rate of every metabolite in the system to zero as

\[ S \cdot v = 0, \]

Where \( S \) is an \( m \times n \) matrix of the stoichiometric coefficients for the \( n \) reactions and \( m \) metabolites in the model, \( v \) is the \( n \times 1 \) vector of the steady state fluxes through the \( n \) reactions in the model. The introduction of these constraints will exclude thermodynamically infeasible fluxes, an example being flux through thermodynamically infeasible flux loops. Furthermore, these constraints also allow quantification of the ranges in the metabolite activity profiles needed to drive reactions in the direction of flux reported in all calculated flux distributions. This is essential for developing kinetic models of metabolism and metabolic control analysis (Famili et al., 2005, Wang et al., 2004, Heinrich and Rapoport 1974, Hatzimanikatis et al., 1996).
2.6: Group contribution theory and estimation of $\Delta_r G^\circ$ of reactions in *G. sulfurreducens* model:

Knowledge of $\Delta_r G^\circ$ of a reaction is essential for the formulation of the thermodynamic constraints and these must be determined either experimentally or from previously reported data. However, experimental values are only available only for a small fraction of the reactions in the *G. sulfurreducens* model. The group contribution method provides a means of estimating the $\Delta_r G^\circ$ of nearly every reaction (Mavrovouniotis et al., 1990, 1991). Since the *G. sulfurreducens* model has metabolites in common with the iJR904 model, the $\Delta_r G^\circ$ of those reactions in the *G. sulfurreducens* model that has metabolites in common with the iJR904 model were estimated utilizing the $\Delta_r G^\circ$ values of metabolites in the iJR904 model. The $\Delta_r G^\circ$ of all other reactions were estimated using the group contribution method through a website based implementation provided by the authors (M.D. Jankowski et al., 2008).

All other reactions for which the $\Delta_r G^\circ$ value could not be estimated were lumped together to produce a set of net reactions for which the value $\Delta_r G^\circ$ could be estimated. Since an experimental measure of $\Delta_r G^\circ$ is unavailable for most compounds in *G. sulfurreducens* metabolism, the group contribution method developed by Mavrovouniotis provides a means of estimating the $\Delta_r G^\circ$ for most metabolites by providing the estimated $\Delta_r G^\circ$. The group contribution methodology considers a single compound to be made up of smaller structural sub groups. The Gibbs free energy change associated with a sub group, $\Delta_{gr} G^\circ$ can be found in literature along with corrections for biochemical cofactors such as CoA, $\text{NAD}^+$/NADH (Mavrovouniotis et al., 1990, 1991).
The following expression may be used to estimate the $\Delta G^\circ$ of the entire compound,

$$
\Delta_f G_{est}^\circ = \Delta_{gr} G_0^\circ + \sum n_i \Delta_{gr} G_i^\circ
$$

(7)

In the above expression, the contributions of each of the subgroups are summed along with an origin term, $\Delta_{gr} G_0^\circ$ and $n_i$ is the number of instances of subgroup $i$ in the compound. $\Delta_{gr} G_i^\circ$ is the contribution of subgroup $i$ to $\Delta_f G_{est}^\circ$. The values calculated using the group contribution methodology are based on the standard condition of a solution with pH 7 and zero ionic strength. However, when a reaction takes place in a solution, the reactants are bound to dissociate into several ionic forms (Alberty, R.A 1998). For the dissociation of ATP, many intermediate ionic forms are possible: $\text{ATP}^4^+$, $\text{HATP}^3^-$ and $\text{H}_2\text{ATP}^2^-$.

In the group contribution method, the total amount of ATP is given by the single most common charged form found in a pH7 solution ($\text{ATP}^4^+$). Thus the reaction for the hydrolysis of ATP to ADP is given by:

$$
\text{ATP}^{4^-} + \text{H}_2\text{O} \leftrightarrow \text{ADP}^{3^-} + \text{Pi}^{2^-} + \text{H}^+
$$

(8)

The form of the reactants used in this work is the most common ionic form of a species at pH 7 is observed in the cytosol of an $E.coli$ cell.

**2.6.a: Lumping Reactions to eliminate metabolites with unknown $\Delta_f G^\circ$:**

Those compounds in the $G.\ sulfurreducens$ model for which $\Delta_f G_{est}^\circ$ could not be estimated were lumped with other reactions in the model to yield a single reaction and these compounds were eliminated. Consider the following sequence of reactions,

$$
X \to Y + Z
$$

(9)
$$Y + A \rightarrow B \quad (10)$$

Let’s say $\Delta_f G_{est}^o$ of compound Y is unknown and the $\Delta_f G_{est}^o$ of all other compounds X, A, Z and B are known. Now we add the above two reactions in order to eliminate B. Addition yields,

$$X + A \rightarrow Z + B \quad (11)$$

After this lumping, the two reactions (9) and (10) are removed from the model and are replaced by a single reaction (11). Furthermore, metabolite B is not explicitly accounted for in the network. Thus, the effect of lumping reactions effectively reduces the total number of reactions in the model.

2.7: Model Based Deletion Analysis:

For the model based deletion analysis, a previously published model of *G. sulfurreducens* was used. Deletion mutant growth rates were predicted by calculating the maximum possible growth rate of the mutant in the presence of a specific environmental condition. Constraints on the exchange fluxes were incorporated to test the optimal growth under different conditions. The following growth medium constituents were used in the model based analysis.

Electron Donors: Acetate, Pyruvate and Hydrogen.

Electron Acceptors: Ferric Citrate, Fumarate.

The presence of hydrogen was modeled by allowing an influx of 10 mmol/gDW hr of hydrogen. *In vivo* predictions have indicated the inability of *G. sulfurreducens* to grow when pyruvate is the sole electron donor (Segura et al., 2008). Hence, the constraint on the pyruvate
uptake was set to 0.015 mmol/gDW hr, so as to meet the non growth associated ATP requirement of 0.45 mmol/gDW hr.

This was done to allow the use of pyruvate as the carbon source and not the sole electron donor in silico. The pyruvate uptake constraint was chosen since it is active only during growth on pyruvate and hence would not impact growth rate predictions in other environments. Growth with acetate and fumarate was modeled by allowing an influx of 5 mmol/gDW hr of acetate and constraining the uptake rate of fumarate to 25 mmol/gDW hr. Similarly, Fe(III) citrate reduction was simulated by allowing the Fe(III) citrate consumption rate to be 150 mmol/gDW hr and constraining the acetate uptake rate to 10 mmol/gDW hr. These uptake rates were chosen, as they are consistent with previously reported values (Mahadevan et al., 2006, Esteve-Nunez et al., 2005, Segura et al., 2008). Lower growth rates were observed under Fe(III) citrate reduction rather than fumarate reduction.

2.8 Thermodynamic Variability Analysis:

The thermodynamic variability analysis is similar to the flux variability analysis (which involves maximizing and minimizing of every flux in the metabolic network at optimal growth conditions), except that in this case, the standard free energy change of every reaction in the model is maximized and minimized to find the range of free energies for every reaction in the model under optimal growth conditions. This analysis is performed subject to both mass balance and thermodynamic constraints. It is evident that the solution of this series of optimization problems results in thermodynamically allowable ranges for the free energies for every reaction in the network.
For this analysis, simulations were done for growth on acetate as the donor and fumarate as the electron acceptor. For growth on acetate as the energy source and fumarate as the sole electron acceptor, the uptake rate for acetate was set to 5 mmol/gDW hr and the uptake rate of fumarate was set to 25 mmol/gDW hr so that these uptake rates were consistent with experimentally determined values (Esteve-Nunez et al., 2005). The thermodynamic variability analysis is done with both the mass balance and the thermodynamic constraints and an additional constraint that forces the growth rate to be constrained. For growth conditions involving acetate as the energy source and fumarate as the sole acceptor, the optimal growth rate was forced to 0.05 1/hr (Segura et al., 2008). The system was also studied when Fe(III) citrate was the sole acceptor. In this case, the acetate uptake and the Fe(III) uptake rates were set to 10 mmol/gDW hr and 150 mmol/gDW hr respectively (Segura et al., 2008). The optimal growth rate was set to 0.04 1/hr.

2.9: Strain and Culture conditions:

Wild type G. sulfurreducens (DL1) strain was cultured in NBAF medium, containing 15mM acetate as the electron donor and 40mM fumarate as the sole electron acceptor under strict anaerobic conditions to 30°C. Cell growth was monitored at various time points by measuring the optical density using a spectrophotometer.

The composition of the growth medium per liter of de ionized water was 0.42g of KH₂PO₄, 0.22g of K₂H PO₄, 0.2g of NH₄Cl, 0.38g KCl, 0.36g NaCl, 0.04g CaCl₂·2H₂O, 0.1g MgSO₄·7H₂O, 1.8g NaHCO₃, 0.5g of Na₂CO₃, 2.04 g of NaC₂H₃O₂·3H₂O, 6.4g of Na₂C₄H₄O₄, 0.5ml of 0.1% resazurin, 1 ml of 100mM Na₂SeO₄, 10 ml of vitamin solution and 10ml of NB trace mineral solution.
2.9a: Quenching Metabolism:

Quenching essentially refers to quickly arresting the metabolic activity prior to harvesting the cells by minimal perturbations. For this, 2L of cells cultured on NBAF medium (batch culture) were harvested at mid log growth phase (O.D measured at 600nm). The metabolism was quenched using 60% cold methanol (-80°C) in a glove box. The medium containing the cells was transferred to separate 250mL centrifuge bottles and centrifuged at 5500rpm for 30 min at 4°C. After centrifuging, cell pellets were separated from the supernatant in a glove box.

2.9b: Flash freezing the samples:

Conical tubes containing the cell pellets and the supernatant were immediately flash frozen in liquid nitrogen and were stored at -80 °C until further analysis. The samples were sent to Chenomx Inc, Alberta for further analysis using NMR spectroscopy.
3.1: Prediction of Fluxes under Acceptor and Donor limiting physiological modes:

In order to examine the changes in the metabolic flux distribution, we compared the fluxes throughout central metabolism under donor and acceptor limiting modes. The figure shown below shows the metabolic flux distribution for acetate limited and Fe(III) limited physiological modes. Fluxes (mmol/gDW hr) shown to the left are acetate limited fluxes while fluxes shown to the right are Fe(III) limited fluxes. Fluxes (mmol/gDW hr) that increase under acceptor limiting condition are shown in red while fluxes that decrease under donor limiting condition are shown in blue. On inspection, it can be seen that the fluxes through some of the pathways are significantly different under the donor and the acceptor limiting physiological modes. This is attributed to favoring of catabolic reactions under acceptor limiting mode and anabolic reactions under donor limiting mode. Increased TCA cycle flux and hydrogen production are observed during acceptor limiting conditions. However, the flux through the biomass synthesis pathway, pyruvate oxidoreductase and ammonium transport decrease as some of the excess acetate is directed towards catabolic reactions. Thus, under acceptor limiting mode, catabolism of acetate is favored. The constraint based model predicted a 25% drop in the growth yield under acceptor limiting conditions than when donor limits metabolism. During acceptor limiting condition, some of the electrons from acetate oxidation are released as hydrogen (a metabolite), causing the yield to drop under this condition. Furthermore, the constraint-based model successfully captured hydrogen secretion under acceptor limitation (Fig.3.1). The hydrogen so secreted acts as an electron sink. Another interesting finding pertains to the
significant difference in the ammonium uptake rates for the acceptor and donor limiting physiological modes. The model computed a lower ammonium uptake when the acceptor limits growth, than when the donor limits growth (Fig.3.1).

![Diagram of central metabolism of G. sulfurreducens](image)

Fig.3.1: Predicted fluxes (mmol/gDW hr) through central metabolism of *G. sulfurreducens* during acetate limited and Fe(III) limited physiological modes.
This is attributed to the fact that amino acid synthesis requires NH₄, and during acceptor limiting case, since catabolism of acetate is favored, we have a lower amount of acetate entering the biomass synthesis pathway and hence a lower yield. This lowered biomass synthesis results in a lower uptake of NH₄. These results are seen to be in accordance with decreased growth yields observed when \( G. \text{sulfurreducens} \) was cultured in Fe (III) limited chemostats. Hence, the constraint-based model of \( G. \text{sulfurreducens} \) is shown to account for changes in the metabolic network in response to changes in its growth environment.

### 3.2: Influence of Nutrient availability on growth yields:

The typical microbial component of subsurface reactive transport models treats growth yield for a given form of metabolism as a constant (Schafer et.al., 1998, Scheibe et.al., 2006; Reginer et.al., 2005). However, growth yields may be influenced by a range of environmental conditions in reality. The influence of environmental parameters on growth yields can be captured by the constraints-based models. It is known that during in situ bioremediation of U (VI), solid phase Fe(III) oxides are the primary electron acceptor for the growth of \( G. \text{sulfurreducens} \) (Finneran et al., 2002). There is only a finite concentration of Fe(III) at any particular location in the subsurface. Over a period of time, the Fe(III) closest to the site of acetate injection gets depleted and can limit the activity of \( G. \text{sulfurreducens} \) (Vrionis et al., 2005; Anderson et al., 2003). The \( G. \text{sulfurreducens} \) constraint-based model predicts that fluxes under electron acceptor and electron donor limiting conditions are significantly different. For example, under acceptor limiting conditions, the predicted yield is 25% of that under electron donor limiting conditions. Furthermore, under acceptor limiting conditions, hydrogen is
predicted to be produced as some of electrons derived from acetate oxidation are released as hydrogen. These results are in accordance with decreased growth yields observed in Fe(III) limited chemostats and hydrogen production in such cultures (Esteve-Nunez et al., 2005, Esteve-Nunez in preparation).

These results emphasize the predictive ability of the *G. sulfurreducens* constraint based model to represent changes in physiology in response to nutrient limitation and also suggest that the metabolism of *G. sulfurreducens* can change dramatically with changing conditions during in situ bioremediation. Nutrient availability can also influence growth yields to a great extent during bioremediation. For example, at the Rifle site in Colorado, ammonium concentrations varied from 300μM to undetectable within a small field plot (Mouser et al., submitted). Previous studies have shown that the biomass yield of *G. sulfurreducens* per mole of acetate oxidized is significantly higher in the presence of ammonium rather when microbial nitrogen fixation is required and this primarily attributed to the higher energetic cost of nitrogen fixation (Bazylinski et al., 2000; Methe et al., 2005). This yield difference was captured successfully by the constraint based model. It may however be noted that availability of other nutrients such as dissolved iron (O’Neil et al., 2008), oxygen (Mouser et al., submitted) and phosphate are also important growth factors during in situ uranium remediation (N’Guessan et al., 2008; N’Guessan et al., submitted).

In order to simplify the overall modeling effort, the present work considers the influence of only three growth factors (acetate, Fe (III) and NH₄), whose concentrations vary significantly during in situ bioremediation. The physiology of *G. sulfurreducens* was modeled for one thousand possible combinations of environmental conditions over linear spaced concentration ranges of acetate (0-5mM), iron (0-10mM) and ammonium (0-1mM). These concentrations were selected
based on actual field-scale data at the Rifle, Colorado site. The result of this modeling is shown as a plot of the biomass yield (gDW/mol Ac) against the frequency of each of the biomass yields in the lookup table (Fig.3.2). The biomass histogram depicted below was plotted only for those conditions in the lookup table where growth was feasible. Substrate uptake rates were modeled with Michaelis-Menten kinetics as uptake of acetate, Fe(III) and ammonium were assumed to follow Michaelis-Menten kinetics (Esteve-Nunez et al., 2005).

Fig.3.2: Frequencies of the biomass yield levels within the one thousand element lookup table generated for linear distribution in concentrations of Acetate, Fe(III) and NH$_4$ using the in silico model of *G. sulfurreducens* for various environmental conditions. Legend: Fe lim- Fe(III) limited growth, NH$_4$ lim- NH$_4$ limited growth, Fe and NH$_4$ lim- Fe(III) and NH$_4$ limited growth, Ac lim- acetate Limited growth.
The coupling of a constraint based *in silico* model with a Michaelis–Menten kinetic model is called hybrid constraint based model. The figure shown above depicts significant variation in the growth yields levels within the one thousand element lookup table, with changes in the concentration of growth factors (Ac, Fe(III) and NH₄). The variation in the biomass yield levels was attributed to conditions of nutrient limitation. This result shows that the constraint-based model of *G. sulfurreducens* can predict the condition specific differences in microbial metabolism, and are hence more effective than empirical Monod type models that assume a constant yield.

The biomass histogram depicted above shows the distribution in growth yields for one thousand possible combinations of growth factors. The constraint based model predicted feasible growth for 450 out of the 1000 possible combinations of nutrient availability. The biomass yield was found to vary between 2.57 and 3.5 gdw/mol acetate consumed. Four different physiological states (Ac lim, Fe(III) lim, Fe(III) and NH₄ lim, Ac lim) were identified. Similar biomass yields were possible for different physiological states. It was also found that depending on nutrient availability under a range on environmental conditions, the biomass yield could vary up to 36%.

An interesting aspect of the above plot is that marginal yields were not observed as soon as the maintenance energy exceeded its threshold value. However it must be noted, that the relatively narrow range of yields reported above is an artifact of the bin resolution used to generate the histogram plot. Ten different concentrations were used for acetate, Fe(III) and ammonium to generate a thousand combinations with corresponding simulated yields. Since the above histogram is based on ten discrete concentrations, between a minimum and a maximum, leading to a lack of simulated yields (for example) for acetate concentrations between 0 and
1mM. Therefore, in the simulations reported, the lookup table was regenerated based on an exponential distribution of concentration values that provided closer spacing (higher resolution) at the low end of the concentration range. An exponential distribution in the substrate concentrations was selected to increase the resolution of the lookup table at lower concentrations where fluxes are expected to change rapidly and metabolic limitations are expected. Accordingly, the range of simulated yields in the model runs was larger than that indicated by the histogram shown above. (Fig. 3.3) below indicates the revised histogram, which indicates a range of 1.95 gdw/mol Ac to 4.05 gdw/mol Ac, for the bin-centres which is a significant variation in the growth yield. When the simulations were done using exponential concentrations for acetate, Fe(III) and ammonium, the hybrid model was seen to predict no growth of *Geobacter* for 600 of the one thousand different relevant combinations of acetate, Fe(III), and ammonium availability that were modeled, since the concentrations of one or more of these substrates was too low to provide enough ATP for maintenance energy requirements. The simulations were done by setting the maximum specific uptake rate of ammonium to be 0.5 mmol/gDW hr. For the remaining 400 of the 1000 possible combinations of nutrient availability, the predicted growth yields varied over nearly an order of magnitude from 0.5 gDW/mol Ac to 4.2 gDW/mol Ac (Fig.3.3). This reinforces the concept of significant yield variability based on changes in nutrient availability. However, due to lumping of values in the histogram plot, the range of individual values is larger than the values at the bin midpoint. A small number of biomass yield values, typically lower than 1 gDW/mol Ac, were not included in the histogram plot as they constituted only 1% of the total number. This result shows that sampling more number of data points at the lower end provides a much better representation of the distribution of the growth yields as
opposed to uniform sampling of points throughout the solution space. The growth yield was found to increase when the maximum specific uptake rate of ammonium was increased to 0.5 mmol/gDW hr, as more acetate (donor) is used up in making biomass in the biomass synthesis pathway. Also, the biomass yield was found to vary by 52%. This huge variation in the biomass yield level is due to the fact that points were sampled exponentially, and hence at lower concentrations, the fluxes can change rapidly, when the simulation was done based on exponential concentrations of acetate, Fe(III), and NH₄. Thus, it can be seen that the biomass yield variation is significantly higher with a higher specific ammonium uptake rate and with exponential concentrations of the donor and acceptor.
Fig. 3.3: Frequencies of biomass yield levels within the 1000-element look-up table generated by the *in silico* model of *G. sulfurreducens* for various environmental conditions using exponential distribution of concentration for acetate, Fe(III) and NH₄. Legend: Fe lim- Fe(III) limited growth, Ac lim-acetate limited growth, Fe lim and N₂ fix- Fe(III) limited and N₂ fixation dependent growth, N₂ fixation dependent growth.

3.3: Predicting the Threshold Flux and Threshold Concentration for the donor and acceptor at different uptake rates:

To simulate donor limiting physiological mode as shown in Fig. 3.3, the uptake rate of acetate was varied between 0-10 mmol/gDW hr, as the maximum specific uptake rate for acetate in the lookup table was fixed to 10 mmol/gDW hr. The Fe(III) consumption rate was fixed at a physiologically reasonable value of 1000 mmol/gDW hr. In essence, unlimited Fe(III) uptake was allowed in the model. Similarly, for the acceptor limiting case, the Fe(III) consumption rate was varied between 0-100 mmol/gDW hr, as the maximum specific consumption rate of Fe(III) in the model was set to be equal to 100 mmol/gDW hr. Unlimited uptake of acetate was allowed by setting the upper limit of the acetate uptake to be 1000 mmol/gDW hr. Under acetate limiting conditions, a minimum threshold flux of 0.9 mmol/gDW hr and a threshold concentration of 0.001 mM were predicted. Similarly, under acceptor limiting conditions, a minimum flux of 7.2 mmol/gDW hr is needed to supply ATP to meet maintenance energy requirements. The threshold concentration for Fe(III) limited growth was identified to be 0.1 mM. Similar threshold limits have been identified for the substrate butyrate during growth of *S. aciditrophicus* (Jackson, B.E. and M.J. McInerney, 2002). The constraint based model predicted no growth of *Geobacter* in 730 of the 1000 possible combinations of nutrient availability. Under these conditions, the
concentration of the substrates (acetate and Fe(III)) was too low to supply enough ATP for maintenance energy requirements.

Fig. 3.4: Prediction from genome scale model of Growth Rates at different uptake rates of acetate and Fe(III) and for different acetate and Fe(III) concentrations when acetate or Fe(III) availability is the factor limiting growth. a) acetate uptake (qAc) as a function of growth rate when acetate is limiting nutrient, b) acetate Concentration as a function of growth rate when acetate is limiting nutrient, c) Fe(III) consumption rate (qFe(III)) as a function of growth rate when Fe(III) availability is the limiting nutrient, d) Fe(III) Concentration as a function of growth rate when Fe(III) is the limiting nutrient.
3.4: Variation in Energy yield, hydrogen secretion with biomass yield:

The conceptual realization that microorganisms have measurable energy requirements arose from the growth yield studies, first in batch reactors and later on in chemostats. In a chemostat, for example, microorganisms are grown within a flow through-reactor under conditions of constant flow (Q). The cell density in a reactor (X) is typically a function of the substrate concentration in the reactor, S. A relation is observed between the specific growth rate \( \mu = 1/X \frac{dX}{dt} \) and the specific rate of substrate consumption \( q = -1/X \frac{dS}{dt} \). For such conditions, the cell yield can be defined as follows:

\[
Y = \frac{\mu}{q} = \frac{(1/hr)}{(mol \ of \ substrate \ utilized \ per \ hour/gm \ dry \ weight \ of \ cells)}
\]  

\[
Y = gm \ dry \ weight / mol \ of \ substrate \ utilized
\]

A plot of the specific growth rate against the substrate utilization rate is a straight line. However, when the growth rate is extrapolated to zero, the substrate utilization rate has a finite non-zero value. This value is known as the maintenance energy, or the energy required for housekeeping functions. A portion of the electrons derived from the oxidation of the electron donor are used to meet the energy requirements. It is known that when a mole of acetate is oxidized, eight electrons are released. Some portion of these electrons is directed to energy generation, while some others participate in hydrogen secretion and formation of biomass. Hence it becomes important to quantify the fraction of electrons directed to energy generation, hydrogen secretion and biomass formation. The stoichiometry for Fe(III) reduction coupled to U(VI) reduction is as given below:
\[
\text{CH}_3\text{COO}^- + 8 \text{Fe(III)} + 4 \text{H}_2\text{O} \rightarrow 8\text{Fe(II)} + 2 \text{HCO}_3^- + 9\text{H}^+ \quad (9)
\]

\[
\text{CH}_3\text{COO}^- + 4 \text{U(VI)} + 4 \text{H}_2\text{O} \rightarrow 4\text{U(IV)} + 2 \text{HCO}_3^- + 9\text{H}^+ \quad (10)
\]

3.4.a: Energy Yield as a function of biomass yield:

The energy yield is the fraction of acetate used for energy generation. It is calculated as the fraction of electrons recovered during Fe(III) reduction relative to the total electrons available. Thus we can write:

\[
\text{Energy Yield} = \frac{q_{\text{Fe}}}{8q_{\text{Ac}}} \quad (11)
\]

Where \( q_{\text{Fe}} \) is the Fe(III) consumption rate (mmol/gDW hr) and \( q_{\text{Ac}} \) is the acetate uptake rate (mmol/gDW hr). Fig.3.5 shows the energy yield as a function of each biomass yield. Furthermore, the maximum and the minimum energy yield possible for each biomass yield are also shown. It can be seen that there is a slightly wider range (75-95% in some cases) released during acetate oxidation is directed towards meeting energetical requirements. This indicates that the majority of electrons that are derived from the oxidation of acetate are directed towards ATP synthesis.
Fig. 3.5: Fraction of electrons used for ATP synthesis as a function of Biomass Yield generated using a exponential distribution in the concentrations for acetate, Fe(III), and NH₄.

3.4.b: Hydrogen Secretion rate as a function of biomass yield:

The in silico model was seen to successfully predict a higher TCA cycle flux (two fold higher) under acetate excess conditions, suggesting that catabolism of acetate is favoured compared to anabolism, and thus more electrons are readily available for metal reduction. This increase of the TCA cycle flux in *G. sulfurreducens* when acetate is supplied in excess has important environmental implications because it implies a higher production of NADH, so more electrons are accessible for exocellular metal reduction, which may be a factor responsible for the *Geobacter* predominance in subsurface environments in which dissimilatory uranium and
vanadium reductions have been stimulates with acetate additions (Anderson et al., 2003, Ortiz-Bernad et al., 2003, Vrionis et al., 2005).

Previous studies have shown that hydrogen secretion occurs as an overflow mechanism in chemostats to cope with the excess of electrons released during acetate oxidation under fumarate limiting conditions (Esteve-Nunez et al, in press). Under acetate-excess conditions, the acetate transporters became down-regulated. In addition, the oxidation-related acetyl-CoA transferase and the assimilatory-related acetate kinase were up regulated and down regulated respectively, indicating that the transcriptional regulation of acetate activation might be the key point for coping with the excess of acetate and increasing the TCA flux (Esteve-Nunez et al, in press). Although uncoupled metabolism related reactions are difficult to predict, the in silico model suggested that hydrogen would act as an electron sink to eliminate the excess of electrons from the cytoplasm when Fe(III) limits metabolism. Thus, hydrogen secretion might be energetically favorable for the cell while disposing the excess of reducing power in the cytoplasm. The fraction of electrons directed towards hydrogen secretion can be calculated by the knowledge of the relevant stoichiometry for hydrogen production. This is given as

\[
2H^+ + 2e^- \rightarrow H_2 \quad (12)
\]

Thus, for every mole of hydrogen secreted, two electrons are released. Hence, the fraction of electrons to hydrogen is given as

\[
\text{Fraction of electrons to Hydrogen} = \frac{2q_{H_2}}{8q_{Ac}} \quad (13)
\]
Where, $q_{H2}$ is the Hydrogen secretion rate (mmol/gDW hr) and $q_{Ac}$ is the acetate uptake rate (mmol/gDW hr). Fig.3.6 shows the maximum and minimum hydrogen secretion rate as a function of each biomass yield in the lookup table.

![Graph showing hydrogen secretion rate vs biomass yield](image)

**Fig.3.6:** Fraction of electrons used for hydrogen production as a function of Biomass Yield generated using an exponential distribution in the concentrations for acetate, Fe(III)$_3$, and NH$_4$. When compared to the energy yield, it can be seen that a relatively smaller fraction of electrons are involved in hydrogen secretion (up to 8%). Furthermore, hydrogen secretion was noticed at lower yields, as these yields reflect the acceptor limiting physiological mode.

[51]
At higher yields, hydrogen secretion was not noticed, as the physiological mode is donor limited. This shift from acceptor limiting to donor limiting mode was captured by the model.

### 3.4.c: Fraction of electrons to biomass formation as a function of biomass yield:

Fig:3.7 shows the fraction of electrons to biomass as a function of yield. It is seen that a smaller fraction of electrons (up to 10% in some cases) are used up for biomass formation. The fraction of electrons to biomass formation can be calculated from the knowledge of the energy yield and the fraction of electrons to Hydrogen secretion.

\[
\text{Fraction of electrons to biomass} = 1 - (\text{Energy yield} + \text{Fraction of electrons to Hydrogen})
\]

However, this fraction seems to be higher than that used for secretion of a metabolite such as hydrogen (Fig.3.7). Hence, it can be concluded that a significant portion of the electrons derived from the oxidation of the donor (acetate) is used to meet energy requirements and for the formation of biomass. It can be seen that the variation is less pronounced when the biomass yield is low. There seems to be a more perceptible variation corresponding to yield ranges of 3.15 gDW/mol Ac to 3.45 gDW/mol Ac. At lower yield levels, the donor is in excess compared to the acceptor, causing the acceptor to limit growth, leading to the secretion of hydrogen. At higher yields, the physiological mode switches to donor limited condition. Thus, lower yields are indicative of acceptor limiting mode while higher yields are indicative of donor limiting physiological mode.
Fig. 3.7: Fraction of electrons used for biomass synthesis as a function of Biomass yield generated using a exponential distribution in the concentrations for acetate, Fe(III)$_4$ and NH$_4$. 
3.5. Using TVA to predict candidate reactions for regulation and possible bottleneck reactions:

For any given reaction in a metabolic network, the magnitude of its free energy change, $\Delta G'$ dictates if this reaction is subject to regulation. Those reactions with a highly negative $\Delta G'$ have the potential to serve as a potential regulatory control points for the pathways in which they participate, as the flux through these reactions is dependent on enzyme regulation (Henry et al., 2007). It has also been proposed that certain reactions that operate close to equilibrium (bottleneck reactions) and consequently have a $\Delta G'$ close to zero are sensitive to minor changes in concentration of the participating metabolites (reactants). Hence, these reactions have limited potential for regulation (Kummel et al., 2006). Thus the $\Delta G'$ ranges generated by the TVA can be used along with the above criterion to determine all reactions in the G. sulfurreducens model that can serve as regulatory control points. Specifically, reactions in the model were identified for which the maximum possible $\Delta G'$ calculated by TVA were $<-1$ kcal/mol. These reactions cannot reach equilibrium under the concentration range studied. In order to focus the analysis on the range of $\Delta G'$ values attainable due to variations in metabolite concentrations alone, the uncertainty associated with the standard free energy change of a reaction was assumed to be zero. The analysis indicated 10 reactions as possible regulatory control points including a non gene associated reaction and included five reactions in the central metabolism subsystem alone. Four reactions were identified as potential bottleneck reactions in the network.

The candidate reactions were further classified as being a strong candidate or a weak candidate for regulation based on the maximum possible $\Delta G'$ calculated by TVA. If the value is $<-1$ kcal/mol strictly, these reactions were classified as strong candidates. However if the
maximum possible $\Delta rG'$ calculated by TVA is $< 6$ kcal/mol and the minimum possible $\Delta rG' < -1$ kcal/mol, these reactions were classified as weak candidates for regulation. Finally, reactions for which $-5 \leq \Delta rG' \leq 0$ were classified as bottleneck reactions. The following table shows a list of potential bottleneck reactions and a list of candidate reactions for regulation in the metabolic network of \textit{G. sulfurreducens}.

Table:1. A List of the model predicted Bottleneck reactions in the metabolic network of \textit{G. sulfurreducens} for optimal growth on acetate as the energy source and fumarate as the sole electron acceptor.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Metabolic Pathway Classification</th>
<th>Associated Gene</th>
<th>Model Predicted Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDH</td>
<td>Central Metabolism</td>
<td>GSU1466</td>
<td>Bottleneck Reaction</td>
</tr>
<tr>
<td>FUM</td>
<td>Central Metabolism</td>
<td>GSU0994</td>
<td>Bottleneck Reaction</td>
</tr>
<tr>
<td>ADSL1</td>
<td>Nucleotide Metabolism</td>
<td>GSU1632</td>
<td>Bottleneck Reaction</td>
</tr>
<tr>
<td>ADSL2</td>
<td>Nucleotide Metabolism</td>
<td>GSU1632</td>
<td>Bottleneck Reaction</td>
</tr>
</tbody>
</table>
Fig. 3.8: List of model predicted thermodynamically constrained reactions (bottlenecks) in the metabolic network of *G. sulfurreducens* for optimal growth on acetate as the energy source and fumarate as the sole electron acceptor. Legends are shown as insets in the above figure. Different color schemes are used to show the bounds in the absence and presence of uncertainty.

Furthermore, in order to study the system of interest, the most thermodynamically constrained reactions were selected and their minimum and maximum possible free energies were evaluated using TVA, first in the absence of any uncertainty, and then accounting for the uncertainty in the standard free energy change by allowing the error in the group contribution method. The results of this analysis are depicted in the figure above, showing expanded bounds on these reactions when the uncertainty is accounted for. This result clearly shows the utility of the TVA in determining the bounds on a reaction when uncertainty is accounted for.
Furthermore, it may be noted that in the case of the fumarase reaction, inclusion of uncertainty changes the bounds marginally, indicating that the concentration of malate (product of the FUM reaction and reactant of the MDH reaction) in the system may be higher. The figure shown below also indicates that when the uncertainty is accounted for, the MDH, ADSL1, ADSL2 reactions cannot behave as bottlenecks under these conditions since they are free to assume a range of $\Delta_r G'$ values. It was noticed that when the simulations were done using Fe(III) citrate as the sole acceptor, the predicted $\Delta_r G'$ ranges were the same as when fumarate was the acceptor. Since the predicted bottleneck reactions are sensitive to perturbations in concentration, they are highly reversible and therefore have significantly different forward and reverse fluxes. This means that there is an exchange factor associated with these fluxes. This factor is the ratio of the reverse flux to the forward flux (assuming the reverse flux is less than the forward flux); this factor is less than unity.

Hence, the model predicted bottleneck reactions were compared with the exchange factors calculated from C$^{13}$ labeled metabolic flux distribution data from acetate limited chemostats with labeled acetate (C.Risso et al., 2008). It was seen that for the reactions MDH and FUM, there is an exchange factor associated with the net flux indicating that the reverse fluxes are significant, and suggesting the potential for these reactions to be operating close to equilibrium. However, experimental results were not available for reactions ADSL1 and ADSL2 probably as they are not a part of the central carbon chemistry of the cell. The results are shown in the table below.
Table: 2. A List of the model predicted bottleneck reactions in the metabolic network of *G. sulfurreducens*, net flux and the exchange factor predicted from C$^{13}$ labelled metabolic flux analysis experiments

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Net flux prediction (mmol/gDW hr)</th>
<th>Exchange Factor determined experimentally</th>
<th>Minimum model predicted $\Delta r G'$ (kcal/mol)</th>
<th>Maximum model predicted $\Delta r G'$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDH</td>
<td>8.06</td>
<td>0.29</td>
<td>-4.70</td>
<td>-1e-06</td>
</tr>
<tr>
<td>FUM</td>
<td>8.14</td>
<td>0.37</td>
<td>-1e-06</td>
<td>-1e-06</td>
</tr>
</tbody>
</table>

Furthermore, in order to study the permissible ranges of the new system variable of metabolite activity, a TVA was performed by maximizing and minimizing the metabolite activity of each metabolite subject to mass balance and thermodynamic constraints. Fig.3.9 shows the result of this analysis. The system was first studied in the absence of uncertainty and then in the presence of uncertainty. The result shows that in the absence of uncertainty, the metabolites participating in bottleneck reactions (mal-L and 25aics) are constrained. In the presence of uncertainty, the activity ranges of most metabolites are unconstrained. These metabolites correspond to those reactions that are most affected by the thermodynamic constraints. Thus, in the presence and in the absence of uncertainty, only a few metabolites are influenced by thermodynamic constraints, suggesting that the cellular system is energetically buffered from large external perturbations, thereby imparting a degree of thermodynamic flexibility to *G. sulfurreducens*. 

[58]
Fig.3.9: Thermodynamically feasible activity ranges for all metabolites with a feasible range that is less than the bounds placed on the metabolites activities (0.01mM-60mM).

3.6: Comparing gene expression data with model predicted candidates for regulation:

As explained earlier, the thermodynamic variability analysis was used to identify potential reactions that are subject to regulation by calculating the ranges in the free energy. Based on the magnitude of these ranges, the reactions were classified as either strong candidate or weak candidates for regulation. Since one or more genes may be associated with these reactions and since genes can be differently expressed (up-regulated or down-regulated) under different conditions, we decided to compare the model predictions against gene expression data.
set, under 21 different conditions. The experimental data pertained to different microarray experiments done to assess the effect of environmental and genetic perturbations on the gene expression levels of *G. sulfurreducens* (Mahadevan et al., 2008). For each condition, the list of genes and their fold changes in gene expression relative to a reference was reported (unperturbed state). In essence, the model predictions were compared against gene expression fold changes for each gene associated with a reaction in the model. The changes, in the range of gene expression were calculated by identifying maximum and minimum fold changes across the 21 conditions after filtering out those conditions having low signal to noise ratio (covariance threshold greater than 20%). This analysis was done for both the strong and the weak candidates for regulation. The results of this analysis are shown below (Fig.3.10 and Fig.3.11). The results indicate that irreversible reaction are most likely to be strongly regulated at the transcript level as enzyme regulation is the dominant mechanism by which flux through these reactions is controlled. Furthermore, irreversible reactions are largely insensitive to perturbations in the reactant or product concentrations due to their highly negative free energy change. Hence, their corresponding gene expression fold changes assume a wider range. For example, it can be seen that for the CS (citrate synthase) and the PPDK (PEP di kinase) reaction, both the model and the gene expression data indicate that these reactions are strongly subject to regulation. This is expected for the PPDK reaction as it involves the translocation of a proton across the cell membrane and hence is energetically more favourable than the PPCK or PPS reaction.

Also it should be noted that for the reaction named SHCHCS2 there are no corresponding genes as it is a non gene associated reaction. Similarly, the analysis of the gene expression data for candidates for weak regulation, suggest that these genes are not as extensively regulated as
the candidates for strong regulation as these reactions have a limited potential for regulation as indicated by their free energy change bounds.

Fig. 3.10: Comparison of gene expression fold changes for 21 different conditions against model predicted strong candidates for regulation using a threshold covariance of 0.20.
Fig.3.11: Comparison of gene expression fold changes for 21 different conditions against model predicted weak candidates for regulation using a threshold covariance of 0.20.
3.7: Analysis of the MDH activity using the results of TVA:

An interesting finding was that the MDH reaction was identified as a thermodynamic bottleneck reaction by the model. This suggested that any change in the concentration of the reactants or the products can force the flux through the reaction to be zero, and consequently, this reaction would form a part of a thermodynamically infeasible flux loop. This effect of the changes in the concentrations on the directionality of the MDH reaction is validated in the following conceptual surface plot that shows the variation in the free energy of the MDH reaction (kcal/mol) in response to perturbations in the malate and oxaloacetate concentrations (mM). Malate concentrations were randomly varied up to 10 mM while oxaloacetate concentrations were varied up to 5 mM. The redox ratio for the NAD/NADH was set to unity. The pH was set to 7 as this a biochemical standard state for [H\(^+\)] and hence, the activity of [H\(^+\)] is unity. The graph clearly shows that a high malate/oxaloacetate concentration is the single most important factor that forces the MDH reaction to be thermodynamically feasible.
Fig.3.12: A surface plot showing the variation in the free energy of the MDH reaction corresponding to changes in the malate and oxaloacetate concentrations (mM).

3.8: in silico phenotype predictions:

3.8a: Wild Type Predictions:

In silico predictions (Table.3) showed that the wild type strain could grow on all combinations of substrates except when pyruvate was the sole carbon source and the energy source. Growth was predicted on pyruvate and hydrogen suggesting that pyruvate could only be used as a carbon source and not as the sole donor in silico. Thus, a constraint was placed on the pyruvate transporter by setting the pyruvate exchange flux to be 0.015 mmol/gDW hr, so that
this flux is sufficient to meet the non growth associated ATP requirement of 0.45 mmol/gDW hr. This was done so that pyruvate could be used only as a carbon source, and moreover, the pyruvate uptake flux is active only when pyruvate is present and does not influence predictions in other environments.

3.8b: PPCK phenotype predictions:

The PPCK mutant was simulated in silico by constraining the flux through the PPCK reaction to be zero. In silico predictions revealed that the mutant grew when fumarate was the sole acceptor, a finding consistent with experimental results (Segura et al., 2008). This is attributed to the PPDK (PEP di kinase reaction), which produces diphosphate leading to proton translocation across the cell membrane giving an energetic advantage to the cell. In contrast, the experimental data suggest that the PPCK mutant cannot grow with Fe(III) as the acceptor, an observation that the FBA model does not predict correctly. It was initially hypothesized that the TMFA model could correctly predict the PPCK knockout phenotype on Fe(III), as MDH might be inactivated due to the additional thermodynamic constraints, leading to a lethal prediction. However, the TMFA model was not able to predict the PPCK phenotype when Fe(III) citrate was the sole acceptor. This is attributed to the fact that MDH is perhaps not the limiting factor for growth under these conditions. Hence, the TMFA predictions are not any different from the FBA predictions, suggesting that thermodynamic constraints alone are not sufficient for correct predictions, and thereby motivating the measurement of intracellular metabolite concentrations.
Table 3: *in silico* model predictions of phenotypes on twelve different growth conditions. Incorrect phenotype predictions are indicated as (+). The symbol + indicates presence of growth while the symbol (-) indicates absence of growth. WT indicates wild type strain.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>WT</th>
<th>PPCK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac/Fum</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ac/Fe(III)</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>Ac+H2/Fum</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ac+H2/Fe(III)</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>Pyr/Fum</td>
<td>_pt</td>
<td>_pt</td>
</tr>
<tr>
<td>Pyr/Fe(III)</td>
<td>_pt</td>
<td>_pt</td>
</tr>
<tr>
<td>Pyr+H2/Fum</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pyr+H2/Fe(III)</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>Ac+Pyr /Fum</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ac+Pyr /Fe(III)</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>Ac+Pyr+H2 /Fum</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ac+Pyr+H2 /Fe(III)</td>
<td>+</td>
<td>(+)</td>
</tr>
</tbody>
</table>

pt: Pyruvate transport constraint
Ac: Acetate
Fe (III): Ferric citrate
Fum: Fumarate

[66]
3.9: Detecting the intracellular concentration of metabolites in *G. sulfurreducens*:

The concentration of intracellular metabolites can be measured by a variety of techniques such as LC-MS, GC-MS, CE-MS and NMR spectroscopy. Recently, vibrational spectroscopic methods such as Raman spectroscopy and IR spectroscopy are also actively being used to measure intracellular metabolites. In this work, the concentration of intracellular metabolites was determined using NMR spectroscopy using a 1D proton NMR. The NMR spectroscopy is based on the principle that the nucleus of certain compounds when placed in an external magnetic field can either absorb or emit electromagnetic radiations at a particular wavelength and frequency called a resonant frequency. This in turn produces a ‘chemical shift’ which is related to the surrounding chemical environment. This chemical shift enables accurate identification and quantification of intracellular metabolite concentrations. 1D proton NMR works on the principle of measuring the chemical shift of protons. Table 4 and Table 5 give a list of major metabolites detected for the triplicates. The term confidence level used in Table’s 4 and 5 refer to a combination of the sensitivity of the NMR instrument, the visible cluster and the uniqueness of metabolite identification. In all, 35 metabolites were detected (See Appendix). It was noticed that the coefficient of variation among the three samples was in some cases above statistically acceptable standards. One possible reason for this is that water extraction was done prior to NMR analysis and since water contains protons and the proton NMR works on the principle of measuring the chemical shift of protons, this gives rise to a condition called water suppression, leading to spectral overlap in some cases. Another reason is that since the culture was essentially a batch culture, a physiological reference steady state condition cannot be achieved. Furthermore, since methanol was used to quench the metabolism, the relatively high
concentration of methanol added resulted in an elevated methanol peak during NMR analysis, overloading the instrument.

It is proposed to design more effective protocols for culturing cells and then quenching their metabolism with minimum perturbations so that the results of the triplicates are more reproducible. However, metabolomics provided a valuable tool to quantify intracellular metabolite concentration for the first time in *G. sulfurreducens*. It was interesting to note that some central metabolites like malate, succinate and 2-oxoglutarate were detected during NMR analysis.

Table: 4. List of major intracellular metabolites and their detected concentrations in the cell pellet of *G. sulfurreducens* cultured on acetate as the electron donor and fumarate as the sole electron acceptor. A complete set of metabolites detected and their associated concentrations is provided in the appendix section.

<table>
<thead>
<tr>
<th>Ac/Fum</th>
<th>Acetate (mM)</th>
<th>Fumarate (mM)</th>
<th>Succinate (mM)</th>
<th>Butyrate (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Pellet_1</td>
<td>15.59</td>
<td>0</td>
<td>0</td>
<td>0.84</td>
</tr>
<tr>
<td>Cell Pellet_2</td>
<td>11.71</td>
<td>2.85</td>
<td>3.98</td>
<td>0.241</td>
</tr>
<tr>
<td>Cell Pellet_3</td>
<td>21.52</td>
<td>12.3</td>
<td>4.13</td>
<td>0.9</td>
</tr>
<tr>
<td>Confidence Level</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>1D Visible Cluster</td>
<td>1/1</td>
<td>1/1</td>
<td>1/1</td>
<td>3/3</td>
</tr>
</tbody>
</table>
Table: 5. List of major intracellular metabolites and their detected concentrations in the cell supernatant of *G. sulfurreducens* cultured on acetate as the electron donor and fumarate as the sole electron acceptor. A complete set of metabolites detected and their associated concentrations is provided in the appendix section.

<table>
<thead>
<tr>
<th>Ac/Fum</th>
<th>Acetate (mM)</th>
<th>Fumarate (mM)</th>
<th>Succinate (mM)</th>
<th>Butyrate (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supernatant_1</td>
<td>20.6</td>
<td>4.5</td>
<td>12.7</td>
<td>0.8</td>
</tr>
<tr>
<td>Supernatant_2</td>
<td>12.8</td>
<td>9.2</td>
<td>4.8</td>
<td>0.27</td>
</tr>
<tr>
<td>Supernatant_3</td>
<td>22.2</td>
<td>27.3</td>
<td>3.7</td>
<td>0.759</td>
</tr>
<tr>
<td>Confidence Level</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>1D Visible Cluster</td>
<td>1/1</td>
<td>1/1</td>
<td>1/1</td>
<td>3/3</td>
</tr>
</tbody>
</table>

**3.9.a: Constraining the solution space using absolute metabolite concentrations:**

In order to further constrain the solution space of the TMFA model using absolute metabolite concentrations, it was decided to evaluate the number of substitutable reactions at optimal growth conditions on fumarate in *G. sulfurreducens*. (Substitutable reactions carry a minimum zero flux and a maximum positive flux at optimal growth conditions). We applied FVA subject to mass balance and thermodynamic constraints to identify the substitutable reactions at optimal growth. We then constrained the concentration bounds on acetate, malate and succinate to their experimentally determined values and again applied the FVA. It was seen
that the number of substitutable reactions decreased from 12 to 11. In this case, the reaction pyruvate kinase was no longer substitutable when the FVA was applied by constraining the concentration of the metabolites. This shows the utility of using absolute metabolite concentrations to further constrain the solution space and also determine unknown flux directions. Next, we constrained the concentration of alanine and malate to their experimental values and furthermore, constrained the concentration of NAD to 2.6 mM based on experimental measurements in the Gram negative model organism, *E. coli* (Bryson D Bennet at al., 2009). We then, applied TVA for optimal growth of WT *G. sulfurreducens* on fumarate. The results indicate that the reaction ALAD_L is sensitive to perturbations in the concentrations of alanine, malate & NAD and forms a part of thermodynamically infeasible flux loop when bounds are placed on the concentrations of these metabolites.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Reactions</th>
<th>Minimum model predicted $\Delta r G'$ (kcal/mol)</th>
<th>Maximum model predicted $\Delta r G'$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac/Fum*</td>
<td>ALAD_L</td>
<td>0.64</td>
<td>4.3</td>
</tr>
<tr>
<td>Ac/Fum</td>
<td>ALAD_L</td>
<td>-4.3</td>
<td>5.70</td>
</tr>
</tbody>
</table>
4. CONCLUSIONS AND RECOMMENDATIONS:

The results presented in this work suggest that constraint-based modeling offers more incentives over the Monod type modeling to describe in situ Uranium bioremediation. The model developed shows that the growth and activity of environmentally relevant microorganisms influencing bioremediation can be accurately predicted from detailed information on the physiology of these microorganisms. This makes it possible to develop microbial models that can be truly predictive and instantaneously capture the changes in physiology of *G. sulfurreducens* in response to changes in a combination of growth factors. Hence, a genome scale *in silico* model of metabolism is applicable to a range of environmental conditions and locations in comparison to empirically calibrated models. For example, the results shown in this work suggest that the constraint based modeling approach can predict important changes in the fluxes of substrates that can result from changes in environmental conditions and the associated difference in growth yield can impact the extent of cell growth in the subsurface. This contrasts with the assumption of constant growth yields and complete metabolism of substrates in Monod type models. Another interesting result from the constraint based modeling approach is the existence of minimum thresholds necessary to maintain metabolism. This concept of minimum metabolic threshold has known to exist for hydrogen metabolism in methanogenic bacteria (Lovley et.al; 1982, Lovley, 1985), However, metabolic based strategies for estimating these environmental thresholds have not been available.

The constraint-based modeling approach summarized in this work is relatively simple compared to the heterogeneity encountered in the subsurface. It considers only three growth
factors that could impact the metabolism and growth yields of *G. sulfurreducens* and ignores parameters such as the death rate of microorganisms. Although, *in situ* Uranium remediation can be modeled by just considering the growth and activity of the genus *Geobacter*, the microbial communities involved in bioremediation processes such as other iron reducers such as *Rhodoferax ferrireducens* and sulphate reducers such as *Desulfovibrio* are much more complex and would need multiple genome scale models. Attempts at increasing the complexity of the modeling approach are underway and would be incorporated in future efforts.

Furthermore, the acetate amendment strategy evidently leads to a situation of acceptor limiting the growth and activity of *G. sulfurreducens* as Iron oxides closest to the site of acetate amendment get depleted over time. This can in turn influence the efficacy of *in situ* remediation. Hence, alternative strategies such as lactate injection are currently being investigated. Injection of lactate stimulates sulphate reducers such as *D. vulgaris* which incompletely oxidises lactate to acetate. This acetate can be used to sustain *Geobacter* activity in the subsurface.

It was shown that incorporating thermodynamic constraints that dictate the thermodynamic feasibility of a reaction is useful in eliminating thermodynamically infeasible flux loops while still allowing the flux through thermodynamically feasible reactions. The usefulness of the thermodynamic variability analysis in addressing thermodynamic bottleneck reactions and identifying potential reactions that serve as a regulatory control point has been illustrated. It was shown that regardless of uncertainty in estimating the standard free energy change, only a few metabolites are affected by thermodynamic constraints in *G. sulfurreducens*. This leads to the conclusion that the majority of the reactions in *G. sulfurreducens* are favourable...
thermodynamically, allowing the reaction to be active under a wide range of metabolite concentrations and hence giving versatility to the cell.

The TVA analysis proved to be an invaluable tool in evaluating the inability of the model to correctly predict the phenotype of the PPCK mutant on Fe(III) Citrate. The TVA analysis indicated that the MDH reaction is essentially a bottleneck reaction and a high malate to oxaloacetate ratio would not be captured by the model as even small changes in the concentrations of the reactants or products would force the flux through the reaction to be zero. This in turn calls for modeling the steady state concentration of oxaloacetate. In order to tightly constrain the metabolite activities, additional non linear constraints involving kinetics must be added to the TMFA formulation. This in turn requires the principles of the law of mass action kinetics or non equilibrium thermodynamics to be included in the analysis. Inclusions of such kinetic constraints are currently being evaluated. Further, it was shown that the TVA could capture the variation in the bounds on the free energy when the uncertainty specified by the group contribution method is taken into account. TVA was also useful in identifying the feasible metabolite activity ranges given the overall bounds on the metabolite activity placed on the system. It was shown that the activity of only a few metabolites was affected by thermodynamic constraints, namely, the reactants and products of the bottleneck reactions. Furthermore, it was shown that metabolomics is a useful tool in determining the intracellular concentration of metabolites of interest and these experimental concentrations can be incorporated in the model to refine the model predictions.

Since in this work, the experiment was essentially done as a batch culture, an accurate physiological reference steady state was not achieved, as control over the microenvironment was
not possible. Furthermore, quenching the metabolism was another issue as addition of a solvent like methanol to quench metabolism resulted in a large methanol peak in the NMR spectra, thereby overloading the instrument. Hence, in some cases, an accurate estimate of concentration was not possible. Defined biomass is a prerequisite for any metabolomics experiment. Hence, the low biomass yield of *G. sulfurreducens* on fumarate was a bottleneck. Hence, to enable accurate NMR detection, the cells had to be cultured in large volumes (2L) in order to get sufficient biomass. In order to overcome these difficulties, it is proposed to perform the experiments in a chemostat, as accurate temperature control would be possible and other parameters such as volume would also be fixed. Moreover, in a chemostat at steady state, specific growth rate would be equal to the dilution rate, which can be well defined and fixed. By fixing the specific growth rate, the fluxes such as the specific substrate uptake rate are fixed. Furthermore, a specific growth limiting medium component such as a specific carbon source can be imposed. In addition, it would be necessary to optimize the protocol for cell quenching and harvesting by using liquid nitrogen rather than methanol to quench metabolism with minimum perturbation. It is proposed that these steps would help in getting more accurate predictions among the replicates. Furthermore, it would be important to repeat the experiments with Fe(III) citrate as the electron acceptor and acetate as the electron donor as this would give valuable information on the metabolism and *in vivo* regulation in *G. sulfurreducens*. In order to help resolve the activity of the PPCK mutant, future efforts should focus on metabolomics experiments on the PPCK mutant using fumarate as the sole acceptor in order to get an idea of the malate and oxaloacetate concentrations in the cell. It is hypothesized that using these concentrations to constrain the solution space of the TMFA algorithm, would help improve its predictive ability.
5. REFERENCES:


6. APPENDIX:

Fig. 4.1: Gene expression fold changes for the predicted Bottleneck reactions.
Table 7: List of model predicted bottleneck reactions, their associated genes and the ranges in free energy under optimal growth using acetate as the electron donor and fumarate as the electron acceptor. Analysis assuming no uncertainty in standard free energy change.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Metabolic Pathway classification</th>
<th>Associated gene</th>
<th>Minimum model predicted $\Delta_r G'$ (kcal/mol)</th>
<th>Maximum model predicted $\Delta_r G'$ (kcal/mol)</th>
<th>Model Predicted Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>FUM</td>
<td>Central Metabolism</td>
<td>GSU0994</td>
<td>-1e-06</td>
<td>-1e-06</td>
<td>Bottleneck</td>
</tr>
<tr>
<td>ADSL1</td>
<td>Nucleotide Metabolism</td>
<td>GSU1632</td>
<td>-0.179</td>
<td>-1e-06</td>
<td>Bottleneck</td>
</tr>
<tr>
<td>ADSL2</td>
<td>Nucleotide Metabolism</td>
<td>GSU1632</td>
<td>-0.179</td>
<td>-1e-06</td>
<td>Bottleneck</td>
</tr>
<tr>
<td>MDH</td>
<td>Central Metabolism</td>
<td>GSU1466</td>
<td>-4.70</td>
<td>-1e-06</td>
<td>Bottleneck</td>
</tr>
</tbody>
</table>
Table 8: List of metabolites and their minimum and maximum activities (mM) at optimal growth with fumarate as the sole acceptor. Analysis assuming no uncertainty in standard free energy change.

<table>
<thead>
<tr>
<th>Metabolite Abbreviation</th>
<th>Minimum Activity (mM)</th>
<th>Maximum Activity (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25aics</td>
<td>44.32</td>
<td>60</td>
</tr>
<tr>
<td>ahdt</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>amopbut-L</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>cysam</td>
<td>44.32</td>
<td>60</td>
</tr>
<tr>
<td>glutrma</td>
<td>0.01</td>
<td>38.13</td>
</tr>
<tr>
<td>glyc</td>
<td>0.363</td>
<td>60</td>
</tr>
<tr>
<td>ile-L</td>
<td>0.01</td>
<td>2.24</td>
</tr>
<tr>
<td>mal-L</td>
<td>3.52</td>
<td>60</td>
</tr>
<tr>
<td>mmoca-S</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Table 9: List of predicted strong candidates reactions subject to regulation, their associated genes and the minimum and maximum model predicted free energy changes (kcal/mol).

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Metabolic Pathway classification</th>
<th>Associated gene</th>
<th>Minimum model predicted $\Delta_r G'$ (kcal/mol)</th>
<th>Maximum model predicted $\Delta_r G'$ (kcal/mol)</th>
<th>Model Predicted Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACHBS</td>
<td>Amino acid Metabolism</td>
<td>GSU1911,GSU1736, GSU1910</td>
<td>-17.9</td>
<td>-2.25</td>
<td>Strong candidate</td>
</tr>
<tr>
<td>ACLS</td>
<td>Amino acid Metabolism</td>
<td>GSU1911,GSU1736, GSU1910</td>
<td>-17.9</td>
<td>-2.25</td>
<td>Strong candidate</td>
</tr>
<tr>
<td>ADCL</td>
<td>Amino acid Metabolism</td>
<td>GSU0523</td>
<td>-41.2</td>
<td>-31.12</td>
<td>Strong candidate</td>
</tr>
<tr>
<td>CS</td>
<td>Central Metabolism</td>
<td>GSU1106</td>
<td>-10</td>
<td>-9.2</td>
<td>Strong candidate</td>
</tr>
<tr>
<td>DHDPS</td>
<td>Amino acid Metabolism</td>
<td>GSU0159</td>
<td>-37.26</td>
<td>-16</td>
<td>Strong candidate</td>
</tr>
<tr>
<td>LDH_L</td>
<td>Central Metabolism</td>
<td>GSU1466</td>
<td>-33.78</td>
<td>-28.89</td>
<td>Strong candidate</td>
</tr>
<tr>
<td>ME1X</td>
<td>Central Metabolism</td>
<td>GSU2308</td>
<td>-4.63</td>
<td>-4.45</td>
<td>Strong candidate</td>
</tr>
<tr>
<td>PC</td>
<td>Central Metabolism</td>
<td>GSU2428</td>
<td>-19.2</td>
<td>-1e-06</td>
<td>Strong candidate</td>
</tr>
<tr>
<td>PPDK</td>
<td>Central Metabolism</td>
<td>GSU0580</td>
<td>-28.03</td>
<td>-2.53</td>
<td>Strong candidate</td>
</tr>
<tr>
<td>SHCHCS2</td>
<td>Vitamins and cofactor biosynthesis</td>
<td>Nongene associated reaction</td>
<td>-50.21</td>
<td>-29.87</td>
<td>Strong candidate</td>
</tr>
</tbody>
</table>
Table 10: List of predicted weak candidates reactions subject to regulation, their associated genes and the minimum and maximum model predicted free energy changes (kcal/mol).

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Metabolic Pathway classification</th>
<th>Associated gene</th>
<th>Minimum model predicted $\Delta_r G'$ (kcal/mol)</th>
<th>Maximum model predicted $\Delta_r G'$ (kcal/mol)</th>
<th>Model Predicted Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXPS</td>
<td>Vitamins and cofactor biosynthesis</td>
<td>GSU1764, GSU0686</td>
<td>-15.7</td>
<td>0.047</td>
<td>Weak candidate</td>
</tr>
<tr>
<td>CYSTL</td>
<td>Amino acid Metabolism</td>
<td>GSU0944, GSU0945</td>
<td>-10</td>
<td>0.72</td>
<td>Weak candidate</td>
</tr>
<tr>
<td>PPS</td>
<td>Central Metabolism</td>
<td>GSU0803</td>
<td>-22.21</td>
<td>3.013</td>
<td>Weak candidate</td>
</tr>
<tr>
<td>ME2</td>
<td>Central Metabolism</td>
<td>GSU1700</td>
<td>-4.5</td>
<td>5.67</td>
<td>Weak Candidate</td>
</tr>
<tr>
<td>ASPTA1</td>
<td>Nucleotide Metabolism</td>
<td>GSU1061, GSU1242</td>
<td>-4.68</td>
<td>5.8</td>
<td>Weak Candidate</td>
</tr>
<tr>
<td>ALAD_L</td>
<td>Amino acid Metabolism</td>
<td>GSU2292</td>
<td>-4.33</td>
<td>5.7</td>
<td>Weak Candidate</td>
</tr>
</tbody>
</table>
Table 11: List of all the metabolites identified using NMR spectroscopy in the cell pellet and their intracellular concentrations.

<table>
<thead>
<tr>
<th>Metabolite Name</th>
<th>Cell Pellet Replicate 1 (mM)</th>
<th>Cell Pellet Replicate 2 (mM)</th>
<th>Cell Pellet Replicate 3 (mM)</th>
<th>1DVisible Cluster</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>oxoglutarate</td>
<td>0</td>
<td>0</td>
<td>0.089</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Acetate</td>
<td>15.59</td>
<td>11.71</td>
<td>21.52</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Acetone</td>
<td>0.062</td>
<td>0.006</td>
<td>0.013</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.33</td>
<td>0.037</td>
<td>0.131</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Butyrate</td>
<td>4.26</td>
<td>0.242</td>
<td>0.9</td>
<td>3/3</td>
<td>High</td>
</tr>
<tr>
<td>Cytidine</td>
<td>0</td>
<td>0</td>
<td>0.016</td>
<td>4/8</td>
<td>Medium</td>
</tr>
<tr>
<td>DSS</td>
<td>0.447</td>
<td>0.447</td>
<td>0.447</td>
<td>4/4</td>
<td>High</td>
</tr>
<tr>
<td>Ethanol</td>
<td>0.742</td>
<td>1.23</td>
<td>2.45</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Ethanolamine</td>
<td>0.3</td>
<td>0.067</td>
<td>0.18</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Formate</td>
<td>0.034</td>
<td>0.62</td>
<td>0.012</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Fumarate</td>
<td>0</td>
<td>2.85</td>
<td>12.3</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Glucose</td>
<td>0</td>
<td>0</td>
<td>0.081</td>
<td>13/14</td>
<td>High</td>
</tr>
<tr>
<td>Glutamate</td>
<td>2.11</td>
<td>0</td>
<td>0.12</td>
<td>4/5</td>
<td>High</td>
</tr>
<tr>
<td>Glycerol</td>
<td>0.371</td>
<td>0</td>
<td>0.022</td>
<td>2/3</td>
<td>Medium</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.201</td>
<td>8.31</td>
<td>0.023</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Hypoxanthine</td>
<td>0.283</td>
<td>0</td>
<td>0.023</td>
<td>1/2</td>
<td>Low</td>
</tr>
<tr>
<td>Inosine</td>
<td>0.063</td>
<td>0</td>
<td>0</td>
<td>3/8</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>C1</td>
<td>C2</td>
<td>C3</td>
<td>C4</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td>Isobutyrate</td>
<td>0.169</td>
<td>0.005</td>
<td>0.0272</td>
<td>1/2</td>
<td>Medium</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>0.362</td>
<td>0</td>
<td>0</td>
<td>3/6</td>
<td>Medium</td>
</tr>
<tr>
<td>Isopropanol</td>
<td>0</td>
<td>0.32</td>
<td>0.063</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Lactate</td>
<td>0.471</td>
<td>0</td>
<td>0.012</td>
<td>1/2</td>
<td>Medium</td>
</tr>
<tr>
<td>Leucine</td>
<td>0.352</td>
<td>0</td>
<td>0</td>
<td>5/6</td>
<td>High</td>
</tr>
<tr>
<td>Lysine</td>
<td>0.396</td>
<td>0</td>
<td>0.021</td>
<td>2/7</td>
<td>Medium</td>
</tr>
<tr>
<td>Malate</td>
<td>0</td>
<td>2.61</td>
<td>3.2</td>
<td>2/3</td>
<td>Low</td>
</tr>
<tr>
<td>Methanol</td>
<td>0.261</td>
<td>49.5</td>
<td>8.5</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Methionine</td>
<td>0.229</td>
<td>0</td>
<td>0</td>
<td>3/5</td>
<td>Medium</td>
</tr>
<tr>
<td>Nicotinate</td>
<td>0.179</td>
<td>0</td>
<td>0</td>
<td>4/4</td>
<td>High</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>0.122</td>
<td>0</td>
<td>0</td>
<td>4/6</td>
<td>High</td>
</tr>
<tr>
<td>Propionate</td>
<td>0</td>
<td>0.041</td>
<td>0.51</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Propylene glycol</td>
<td>0</td>
<td>0.045</td>
<td>0</td>
<td>2/4</td>
<td>Medium</td>
</tr>
<tr>
<td>Succinate</td>
<td>0</td>
<td>3.9</td>
<td>4.13</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Theronine</td>
<td>0.038</td>
<td>0</td>
<td>0.007</td>
<td>1/3</td>
<td>Low</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>0.071</td>
<td>0</td>
<td>0</td>
<td>2/5</td>
<td>Medium</td>
</tr>
<tr>
<td>Uracil</td>
<td>0.89</td>
<td>0</td>
<td>0.062</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Valine</td>
<td>0.672</td>
<td>0</td>
<td>0.015</td>
<td>3/4</td>
<td>High</td>
</tr>
</tbody>
</table>
Table 12: List of all the metabolites identified using NMR spectroscopy in the cell supernatant and their intracellular concentrations.

<table>
<thead>
<tr>
<th>Metabolite Name</th>
<th>Cell supernatant Replicate 1 (mM)</th>
<th>Cell supernatant Replicate 2 (mM)</th>
<th>Cell supernatant Replicate 3 (mM)</th>
<th>1DVisible Cluster</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>oxoglutarate</td>
<td>11.1</td>
<td>0</td>
<td>0.089</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Acetate</td>
<td>20.6</td>
<td>12.71</td>
<td>22.2</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Acetone</td>
<td>0.003</td>
<td>0.003</td>
<td>0.001</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.065</td>
<td>0.023</td>
<td>0.01</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Butyrate</td>
<td>0.84</td>
<td>0.27</td>
<td>0.76</td>
<td>3/3</td>
<td>High</td>
</tr>
<tr>
<td>Cytidine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4/8</td>
<td>Medium</td>
</tr>
<tr>
<td>DSS</td>
<td>0.447</td>
<td>0.447</td>
<td>0.447</td>
<td>4/4</td>
<td>High</td>
</tr>
<tr>
<td>Ethanol</td>
<td>0.33</td>
<td>2.07</td>
<td>1.02</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Ethanolamine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Formate</td>
<td>0.1</td>
<td>0.49</td>
<td>0.64</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Fumarate</td>
<td>4.45</td>
<td>9.16</td>
<td>27.23</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Glucose</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13/14</td>
<td>High</td>
</tr>
<tr>
<td>Glutamate</td>
<td>0.054</td>
<td>0</td>
<td>0</td>
<td>4/5</td>
<td>High</td>
</tr>
<tr>
<td>Glycerol</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2/3</td>
<td>Medium</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.054</td>
<td>13</td>
<td>0.002</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Hypoxanthine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1/2</td>
<td>Low</td>
</tr>
<tr>
<td>Substance</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
<td>Value 4</td>
<td>Category</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>Inosine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3/8</td>
<td>Low</td>
</tr>
<tr>
<td>Isobutyrate</td>
<td>0.043</td>
<td>0.003</td>
<td>0.026</td>
<td>1/2</td>
<td>Medium</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3/6</td>
<td>Medium</td>
</tr>
<tr>
<td>Isopropanol</td>
<td>0.032</td>
<td>0.02</td>
<td>0.005</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Lactate</td>
<td>0.013</td>
<td>0</td>
<td>0.006</td>
<td>1/2</td>
<td>Medium</td>
</tr>
<tr>
<td>Leucine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5/6</td>
<td>High</td>
</tr>
<tr>
<td>Lysine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2/7</td>
<td>Medium</td>
</tr>
<tr>
<td>Malate</td>
<td>0.72</td>
<td>0</td>
<td>0</td>
<td>2/3</td>
<td>Low</td>
</tr>
<tr>
<td>Methanol</td>
<td>0.045</td>
<td>81.8</td>
<td>0.452</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Methionine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3/5</td>
<td>Medium</td>
</tr>
<tr>
<td>Nicotinate</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4/4</td>
<td>High</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4/6</td>
<td>High</td>
</tr>
<tr>
<td>Propionate</td>
<td>0.05</td>
<td>0.518</td>
<td>0.045</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Propylene glycol</td>
<td>0.06</td>
<td>0.004</td>
<td>0.052</td>
<td>2/4</td>
<td>Medium</td>
</tr>
<tr>
<td>Succinate</td>
<td>12.67</td>
<td>3.64</td>
<td>4.82</td>
<td>1/1</td>
<td>High</td>
</tr>
<tr>
<td>Theronine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1/3</td>
<td>Low</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2/5</td>
<td>Medium</td>
</tr>
<tr>
<td>Uracil</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2/2</td>
<td>High</td>
</tr>
<tr>
<td>Valine</td>
<td>0.034</td>
<td>0</td>
<td>0.003</td>
<td>3/4</td>
<td>High</td>
</tr>
</tbody>
</table>
Table 13: List of simulation parameters used in the lookup table and their associated values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Maximum uptake rate (mmol/gDW hr)</th>
<th>Half saturation constants (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetate (Ac)</td>
<td>$V_{max1}=3$</td>
<td>$K_{Ac1}=0.161$</td>
</tr>
<tr>
<td></td>
<td>$V_{max2}=0.6$</td>
<td>$K_{Ac2}=0.008$</td>
</tr>
<tr>
<td></td>
<td>$V_{max3}=1.8$</td>
<td>$K_{Ac3}=0.034$</td>
</tr>
<tr>
<td>Ferric Iron (Fe(III))</td>
<td>$V_{maxFe}=100$</td>
<td>$K_{Fe}=1$</td>
</tr>
<tr>
<td>Ammonium (NH$_4$)</td>
<td>$V_{maxNH4}=0.1$</td>
<td>$K_{NH4}=0.1$</td>
</tr>
</tbody>
</table>