Applications of Metabolic Modelling: Understanding Energy Production in Electricity-Producing *Shewanella oneidensis* MR-1 and Lipid-Producing *Nannochloropsis* *gaditana*



Student Name: Mengyu Gao

Registration Number: 110264169

Supervisor(s): Josselin Noirel & Jags Pandhal

Department of Chemical and Biological Engineering

The University of Sheffield

Contents

[Chapter I Introduction: From Traditional Biology to Systems Biology 1](#_Toc414717022)

[Section I.1 Systems Biology 2](#_Toc414717023)

[Section I.1.1 Biological Systems 2](#_Toc414717024)

[Section I.1.2 The Era of 'Omics’ Data 3](#_Toc414717025)

[Section I.1.3 The Era of Modelling 6](#_Toc414717026)

[Section I.1.4 Biological Networks 8](#_Toc414717027)

[Section I.2 Metabolic Networks 9](#_Toc414717028)

[Section I.2.1 Introduction to Metabolism: Anabolism, Catabolism 9](#_Toc414717029)

[Section I.2.2 Representations of Metabolism 9](#_Toc414717030)

[Section I.2.3 Existing databases 13](#_Toc414717031)

[Section I.2.4 Kyoto Encyclopaedia of Genes and Genomes (KEGG) 13](#_Toc414717032)

[Section I.3 Collaboration in the Department 15](#_Toc414717033)

[Chapter II Model-driven Metabolic Analysis of *Shewanella oneidensis* MR-1 17](#_Toc414717034)

[Section II.1 Microbial Fuel Cell (MFC) using *Shewanella oneidensis* MR-1 18](#_Toc414717035)

[Section II.1.1 Introduction to the MFC Culture 18](#_Toc414717036)

[Section II.1.2 Electron Transfer Mechanisms 19](#_Toc414717037)

[Section II.2 Materials and Methods 21](#_Toc414717038)

[Section II.2.1 The Principle of the Flux Balance Analysis 21](#_Toc414717039)

[Section II.2.2 Saving and Loading of the Model 23](#_Toc414717040)

[Section II.2.3 Calculating the Growth Rate 24](#_Toc414717041)

[Section II.2.4 Calculating the NADH Productivity 24](#_Toc414717042)

[Section II.2.5 Alternative Optimal Value 24](#_Toc414717043)

[Section II.3 Model Reproduction 25](#_Toc414717044)

[Section II.3.1 Modelling metabolism of a Cell 25](#_Toc414717045)

[Section II.3.2 The Reconstruction Process 28](#_Toc414717046)

[Section II.3.3 The Metabolite Function of the Model 29](#_Toc414717047)

[Section II.3.4 Model adjusting 33](#_Toc414717048)

[Section II.4 Maximisation of electron fluxes 37](#_Toc414717049)

[Section II.4.1 Electron transport chains in direct electron transport (DET) mode 37](#_Toc414717050)

[Section II.4.2 DET mode under aerobic conditions 38](#_Toc414717051)

[Section II.4.3 DET mode Under Anaerobic Conditions 48](#_Toc414717052)

[Section II.4.4 MET mode and NADH productivity 51](#_Toc414717053)

[Section II.5 Futile Cycles is plural 55](#_Toc414717054)

[Section II.6 The Maximum Electron Productivity by Carbon Source 58](#_Toc414717055)

[Section II.7 The simulation of 13C labelled experiment 61](#_Toc414717056)

[Section II.8 Comparison of electron productivity between predicted and experimental data 65](#_Toc414717057)

[Section II.9 Conclusion 67](#_Toc414717058)

[*Chapter III* Metabolic reconstruction of marine microalgae: *Nannochloropsis gaditana* 69](#_Toc414717059)

[Section III.1 Nannochloropsis gaditana 70](#_Toc414717060)

[Section III.1.1 Introduction to Nannochloropsis gaditana 70](#_Toc414717061)

[Section III.2 Metabolic Reconstruction 70](#_Toc414717062)

[Section III.2.1 Procedure of KAAS 71](#_Toc414717063)

[Section III.2.2 MetNetMaker 71](#_Toc414717064)

[Section III.2.3 Network reconstruction of Nannochloropsis gaditana 72](#_Toc414717065)

[Section III.3 Metabolic reconstruction 74](#_Toc414717066)

[Section III.4 Conclusion 83](#_Toc414717067)

[Chapter IV Conclusions and Future Work 84](#_Toc414717068)

[Section IV.1 Conclusion 85](#_Toc414717069)

[Section IV.2 The future plan of FBA simulation of *Shewanella oneidensis* MR-1 86](#_Toc414717070)

[Section IV.2.1 The gene knockout experiment 86](#_Toc414717071)

[Section IV.2.2 Carbon labelling metabolic flux analysis 86](#_Toc414717072)

[Section IV.3 The future plan of metabolic reconstruction of *Nannochloropsis gaditana* 86](#_Toc414717073)

[Section IV.3.1 The biomass reaction and model prediciton 86](#_Toc414717074)

[Section IV.3.2 Improvement of FBA method: the Dynamic Flux Balance Analysis 86](#_Toc414717075)

Abstract

With the improvement of computational technology in recent years, new research fields such as systems biology have been developed thanks to the interdisciplinary interface between computer science and traditional biology. Distinct from traditional biology, systems biology focuses on cellular activity at a systematic level rather than individual molecular scales. A new technique called ‘Omics’ data analysis has been introduced to systems biology to help understand bio-activities on a greater scale. For instance, proteomics is the study of various protein levels simultaneously. This type of research provides an overall picture of the organism, helping us understand how cellular activities interact with each other.

To further understand subcellular activities, computational modelling was developed with techniques including elementary mode analysis, flux balance analysis, metabolic flux analysis, et cetera.

In this report, two projects related to systems biology have been carried out. The first project is a model-driven metabolic analysis of electron-producing bacteria, called *Shewanella oneidensis* MR-1. In this project, the aerobic and anaerobic respiration was studied. The relation between electron productivity and carbon source has been described. A gene-knockout simulation was also carried out. It was found that the knockout of two ubiquinone-8 related reactions increased the total electron productivity by about 31%. This increase may be because with two knockouts, the flux through the tricarboxylic acid cycle (TCA) cycle maintains a low level, reducing cell growth. Thus, more energy can be converted into electricity. The main electron donor in the electron transport chain is nicotinamide adenine dinucleotide + hydrogen (NADH).

The second project is a metabolic reconstruction of *Nannochloropsis* *gaditana.* As a result,over 300 reactions were included in the model reconstruction of *Nannochloropsis* *gaditana* and the biomass reaction is needed for further predictions. Together with the biomass reaction, this model can be further used for prediction via flux balance analysis (FBA). In the FBA model of *S. oneidensis*, it was found that the model had a better performance under carbon-limited conditions rather than oxygen-limited conditions. The theoretical electron transfer efficiency to the anode was found to be extremely low (less than 0.01% in direct electron transfer (DET) mode or 20% in mediated electron transfer (MET) mode).

# Introduction: From Traditional Biology to Systems Biology

## Systems Biology

### Biological Systems

The fundamental building block of most organisms are cells. Complex organisms such as mammals possess billions of cells. However, the simplest organism may consist of just a single cell.

A cell is complex system. Living cells, either eukaryotic or prokaryotic cells, require the coordination of different cell organs. For example, in the biological system of *Chlamydomonas reinhardtii*, crucial metabolism processes including oxidative phosphorylation, photosynthesis, amino acid synthesis, DNA synthesis, et cetera, take place in different sections of the chloroplast, mitochondria and cytosol (Figure I‑2). Oxidative phosphorylation and photosynthesis are energy generating metabolism processes. The former consumes glucose as an energy source, converting it into ATP (adenosine triphosphate). The latter synthesises ATP from photons and produces oxygen (Thomas D. P. and Willam C. E. 2002). The energy produced can then be used for cell bio-activities such as protein and DNA synthesis. In complex organisms which consist of billions of cells, some cell activities are controlled by neural cells, adding a layer of complexity to the biological system.



Figure I‑2: Biological systems in C. reinhardtii (source from Boyle, N. R. and Morgan, J. A. 2009)

This system consists of a nucleus, cytosol, chloroplast, mitochondria and pyrenold. The TCA cycle, a vital metabolism of the organism, is located in the mitochondria. Amino acid synthesis may take place in three cell organs: the cytosol, mitochondria and chloroplast. Fatty acids are synthesised in the chloroplast (Boyle, N. R. and Morgan, J. A. 2009).

### The Era of 'Omics’ Data

As the result of new technologies, the sequences of informational macromolecules can now be more easily determined. This gives the opportunity of understanding the biological function of organisms at different levels (Kandpal, R. Saviola, B. et al. 2009). The new research field ‘omics’ became popular based on this contribution. The ‘omics’ field includes proteomics, genomics, transcriptomics and so on. Some popular ‘omics fields are listed in Table I‑1.

The sequence of progression for ‘omics’ levels is as follows (see Figure I‑3):

DNA

Genomics

RNA

Transcriptomics

Protein

Proteomics & Interactomics

Metabolites

Metabolomics

Figure I‑3 The ‘Omics’ flux (source from Zhang, W. Li, F. et al. (2010))

Each stage of ‘omics’ is essential in that the reconstructed ‘omics’ processes gives a full view of the cellular activities. Some ‘omics’ fields are introduced follow.

Table I‑1 A description of various ‘Omics’ (Kandpal, R. Saviola, B. et al. 2009 and Zhang, W. Li, F. et al. (2010)

|  |  |
| --- | --- |
| ‘Omics’ Field | Object of Study |
| Proteomics | Protein abundance |
| Metabolomics | Small cellular metabolites abundance |
| Transcriptomics | mRNA transcription |
| Interactomics | Molecular interactions |
| Fluxomics | Dynamic changes of cellular molecules |
| Genomics | Determining genome sequences |
| MicroRNomics | microRNAs (miRNA) |

Some ‘omics’ fields and descriptions are listed in this Table I‑1. ‘Omics’ is actually a suffix added at the end of a word to mean the study of the field. For instance, proteomics is the study of proteins and metabolomics is the study of metabolism.

**Genomics:**

Genomics impacts the family’s health. Recently, it focuses on the study of all the genes in a person. This includes the interaction between genes and the study of harmful diseases like diabetes and cancer (National Human Genome Research Institute 2010). The cost of studying genomics is decreasing due to the development of new DNA sequencing technologies (Zhang, W. Li, F. et al. (2010).

**Transcriptomics:**

Following genomics, transcriptomics is closely linked with sequencing genomes (Kandpal, R. Saviola, B. et al. 2009). It studies the process of gene expression. The key elements in transcriptomics are RNAs, including mRNAs, non-coding RNAs and small RNAs. RNA-Seq, the high-throughput method employed in transcriptomics was published in 2009 (Wang, Z. Gerstein, M. et al. 2009).

**Proteomics and Interactomics:**

Proteins play a crucial role in living cells or organisms, not only in cellular structure, but also in enzymatic reactions. Proteomics studies the structures and the functions of proteins on a large scale (Anderson, N. L. and Anderson, N. G. 1998). Two main methods, the two-dimensional PAGE (2D-PAGE) with mass spectrometric identification and gel-free profiling procedures followed by automated tandem mass spectrometry (LC-MS/MS) are employed in proteomics (Zhang, W. Li, F. et al 2010).

In a cellular system, one function is achieved through to the co-working of many proteins. These co-working proteins are functionally linked to each other. The study of this field is called interactomics (Zhang, W. Li, F. et al 2010 and Kiemer, L. and Cesareni, G. 2007). Three significant approaches in interactomics are automatically identifying the protein interaction in bacteria, protein complexity analysis and large-scale genetic interaction screening (Zhang, W. Li, F. et al 2010).

**Metabolomics and Fluxomics:**

With the information gathered from genomics, transcriptomics and proteomics, the story in a cell can be recombined using metabolomics and fluxomics. The former studies the cellular metabolites while the latter observes the integration between metabolites and dynamic metabolic flux modelling. Several analytical methods are utilised in metabolomics, such as high-performance liquid chromatography mass spectrometry (LC-MS), nuclear magnetic resonance (NMR) spectroscopy and so on (Zhang, W. Li, F. et al 2010).

Corresponding to gene expression pathway, the various ‘Omics’ study cellular activities at different levels. Genomics corresponds to DNA, while transcriptomics studies RNA. Enzymes are studied by proteomics and interactomics, while reactant and enzyme products are the domain of metabolomics.

### The Era of Modelling

The modelling of biological systems plays a significant role in systems biology. The processing power of computers offers a platform for computational models of biological systems. On these platforms, mathematical operations such as algorithms and matrix calculations can be conveniently performed. The *in silico* biological model reflects the biological networks in a cell.

Utilising specific equations allows the simulation of biological networks and the characterisation of their performance. While the simulation does not necessarily represent the actual cell perfectly, it is sufficient to provide an accurate picture of cell networks and their activities under different conditions.

In addition, the mutants of a specific organism can be reconstructed using the computational model. The simulation in computational programs allows the performances of mutant organisms to be predicted. This provides the benefit of creating a general direction to guide laboratory experiments.



Figure I‑4: Integrating the ‘Omics’ data of a cell (source from Joyce, A. R. and Palsson, B. O. 2006)

In order to initialise a biological model, it is necessary for the ‘Omics’ data sets to be integrated. Integrating different ‘Omics’ data sets together allows the creation of a unique biological model (Figure I‑4). The more ‘Omics’ data sets are collected, the more complete the model is.

The ‘omics’ data sets are divided into three classes: components, interactions and functional states. The overall cellular processes are revealed by integrating across all three categories. The picture in the middle of Figure I‑4 illustrates the cellular process from DNA to metabolites. From top to bottom, the gene sequence is transcription to mRNA, followed by the translation process. Proteins produced by the gene are then collected in the metabolic network, catalysing the biochemical reactions in the cell. Metabolites, oligosaccharides, lipids and glycoproteins become activated under the catalysation of enzymes. These processes involve genomics, transcriptomics, proteomics, metabolomics, glycomics and lipidomics. Localizomics is used to characterise the location of these components. Finally, the knowledge of fluxomics comprises the metabolic networks. This allows the visualisation of cell behaviour (Joyce, A. R. and Palsson, B. O. 2006).

### Biological Networks

As an application to biological systems, biological networks are the visualisation of the complex networks in cells. This network is a data set that includes several sub unites, for instance, gene and protein networks, protein-protein interaction networks, metabolic networks and gene-regulatory networks (Proulx, S. R. Promislow, D. E. L. et al 2005). The biological networks are the topology of the ‘omics’, each network represents the complex relationship between cellular molecules in a specific ‘omics’ system.

A network consists of small units which can be considered as metabolites, linked by a number of reactions. This collection of reactions and molecules is considered a system. Most levels of biochemicals can be represented by nodes, not only as genes and proteins, but also neurons and organs. Edges link nodes together representing the cellular interface between nodes at different levels. A simple example from Stephen R. Proulx et al (2005) is introduced in Figure I‑5. Mathematical methods are employed to study network behaviour. The accuracy of the study depends on the resolution level and the degree quantification (Newman, M. 2003).



Figure I‑5: An example of a simple biological network (source from (Proulx, S. R. Promislow, D. E. L. et al 2005))

The green circles and blue squares are the nodes with different degrees. All nodes are linked by dashed lines which are edges demonstrating the interface between nodes. Component C is the simplest example with only one node inside. The darkness cycle in the central map of component A and has a degree of 4. The shortest path between square 1 and square 2 is shown the by dashed arrows, taking 4 steps from square 1 to square 2. In component A, the dark green node locates on 35 shortest paths which is the highest betweenness number. Since every node is linked to another node via only one edge, the clustering coefficient of component B is high.

## Metabolic Networks

### Introduction to Metabolism: Anabolism, Catabolism

Metabolism is the collection of cellular biochemical transformations, catalysed by enzymes. Conventionally, metabolism is divided into two categories, catabolism and anabolism. The former offers energy for cells or organisms by breaking down molecules. For instance, in the glycolysis pathway, glucose is catalysed into pyruvate by enzymatic reactions, then degraded in a step-by-step process into smaller units in the trbicarboxylic acid (TCA) cycle, releasing energy for other cell activities, including anabolism. The latter, anabolism, is the set of reactions that produce “cellular structures” for building the cell.

### Representations of Metabolism

Similar to the representation of the biological network, metabolism can be illustrated via a map of metabolic networks. Considering nodes in Figure I‑5 as metabolites, and edges as enzyme catalysed reactions, a component thus can be recognised as a metabolic network. The TCA cycle is demonstrated in Figure I‑6 as an example. The TCA cycle starts from Condensation, acetyl CoA (2 carbon molecules) and oxaloacetate (4 carbon molecules) catalysed by Citrate synthetase into citrate (6 carbon molecules). The equation of this reaction is:

Acetyl Co-A + Oxaloacetate + H2O 🡪 Citrate + CoA

Citrate synthetase sssssssssynsynthetase

A metabolic network is an assembly of metabolic reactions including reactants, products and enzymes. The TCA cycle, in this case, consists of 9 reactions (Table I‑2) which are collected and represented in a metabolic pathway (Figure I‑6). This network is a small part of the whole metabolic network of an organism. Similarly, any catabolism or anabolism process can be illustrated in this way. The united metabolic pathway of a specific cell contains the most different types of metabolism in the cell. With this demonstration, the complex relationship among the metabolites and enzymes is clearly shown, easing the of study metabolisms.

Tricarboxylic (TCA) cycle is a crucial metabolic pathway that can be seen in cells of almost all organisms, manufacturing energy carriers (ATP/NADP). The cycle begins with the reaction Condensation, which produces 6-c biochemical component. Then 6-c component is degraded into 5-c component in step 4, decarboxylation, followed by oxidative decarboxylation catalysing the 5-c component to 4-c component. Lastly, the 4-c component once again returns to reaction 1, condensation, reacting to 6-c component and starting another iteration of the TCA cycle. In this process, 5 units of energy carriers are produced for every corresponding unit of consumed acetyl CoA.



Figure I‑6: The example of representation of metabolic network: TCA cycle

Table I‑2: Reaction list of TCA cycle (Xamplified Free Online Education Resource 2010)

|  |  |  |
| --- | --- | --- |
| Reaction | Equations | Enzymes |
| 1.Condensation | Acetyl CoA + Oxaloacetate + H2O 🡪 Citrate + CoA | Citrate synthetase |
| 2.Isomerisation | Dehydration:Citric acid🡪Cis-Aconitate + H2O Rehydration:Cis-Aconitate + H2O🡪 Isocitrate | Aconitase, Fe |
| 3.Dehydrogenation | Isocitrate + NAD+ 🡪Oxalosuccinate + NADH2 | Isocitrate dehydrogenase |
| 4.Decarboxylation | Oxalosuccinate 🡪 alpha-Ketoglutarate +CO2 | Oxalosuccinate decarboxylase |
| 5.Oxidative Decarboxylation | alpha-Ketoglutarate + NAD+ + CoA🡪Succinyl CoA + NADH2 + CO2 | Alpha-ketoglutaratedehydrcgenase |
| 6.Substrate level ATP/GTP synthesis | Succinyl CoA + ADP/GDP/IDP🡪Succinate + ATP/GTP/ITP + CoA | Succinyl thiokinase |
| 7.Dehydrogenation | Succinate + FAD ⇌ Fumarate + FADH2 | Succinate degydrogenase |
| 8.Hydration | Fumarate + H2O ⇌ Malate | Fumarase |
| 9.Dehydrogenation | Malate + FAD+ ⇌ Oxaloacetate + FADH2 | Malate dehydrogenase |

There are 9 reactions in the TCA cycle, generating energy and producing H2O and CO2. The main product of one reaction serves as the main reactant for the subsequent reaction. For instance, the product of the condensation reaction, citrate, is the reactant in the isomerisation reaction (citric acid). For the rest of the steps, the main components are Isocitrate, Oxalosuccinate, alpha-Ketoglutarate, et cetera. This is quite normal in metabolic pathways because metabolism is a continuous process that catalyses the chemical components from outside of the organism in order to produce energy or a biochemical structure. This Table I‑2 provides more detail about Figure I‑6, which is more representative of relationships than reactions. Some metabolites such as CO2, H2O and ATP (usually carbon dioxide, water and energy carriers), may be ignored in some metabolic pathways. Some reactions are reversible and have been shown by two-way arrows, for example, dehydrogenation, hydration and dehydrogenation.

### Existing databases

For the convenience of researchers, there are existing databases which contain different kinds of ‘Omics’ data such as the National Centre for Biotechnology Information (NCBI) (Geer, L. Y. Marchler-Bauer, A. et al. 2010) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG). (Kanehisa, M. and Goto, S. 2000) Both the NCBI and KEGG offer a platform for systems biology researchers studying the ‘Omics’ area of interest. Many useful programs and data sets can be found in the two databases, including the Basic Local Alignment Search Tool (BLAST), protein database and DNA databases. The function of a part of the gene can be determined using BLAST and the production can be predicted. BLAST also has the ability of comparing the similarity of proteins and DNA sequences, a useful feature which can be used in other fields. Finally, the structural tools in NCBI can be utilised to forecast the secondary structure of a protein.

For metabolic reconstruction, the targeted gene in the initial step can be easily found using the NCBI if it is sequenced. This gene would then be automatically annotated by the web service on NCBI or KEGG. A draft map can then be produced which provides an overall opinion of what metabolisms of a target organism may consist of, in spite of any missing information in the draft. The reference genes can be chosen from an existing database since there are a large number of genes previously studied.

### Kyoto Encyclopaedia of Genes and Genomes (KEGG)

KEGG is a useful database that is managed by Kyoto University. Enzymes and metabolites are grouped in this database and are used as references in mapping a metabolic pathway. In addition, hundreds of genes of organisms in KEGG can be utilised in automatic annotation. KEGG database includes not only the information of ‘Omics’ data, but also the annotative tools: KEGG Automatic Annotation Server KAAS and BLAST (Moriya, Y. Itoh, M. et al. 2007). KAAS is a web server on KEGG with the capability of annotating the gene of interest. Any target genes can be uploaded to the KAAS server and analysed against a set of chosen organisms as references.

The results of annotations are assembled in a KEGG map (Figure I‑8) with links between genes and enzymes. In an annotation result by KAAS, every hit is shown by a green box for the enzyme. In other words, if an enzymatic-catalysed reaction is found, the corresponding box of enzyme will be green as opposed to white in the case of non-hit reactions. The enzymes can be viewed by tracking the numbers in boxes. For instance, the number 1.1.1.37 is the enzyme Malate dehydrogenase, catalysing the reaction Dehydrogenation:

Malate + FAD+ 🡨🡪Oxaloacetate + FADH2

The reactants and products involved in this reaction are shown in nodes namely Oxaloacetate and (S)-Malate. The two-way arrow between the nodes illustrates the reversibility of the reaction. Clearly, the involution of Oxaloacetate in other metabolic reactions can be observed at the same time. Additionally, the FAD+ and FADH2 in reactions are not shown in the Figure I‑8.

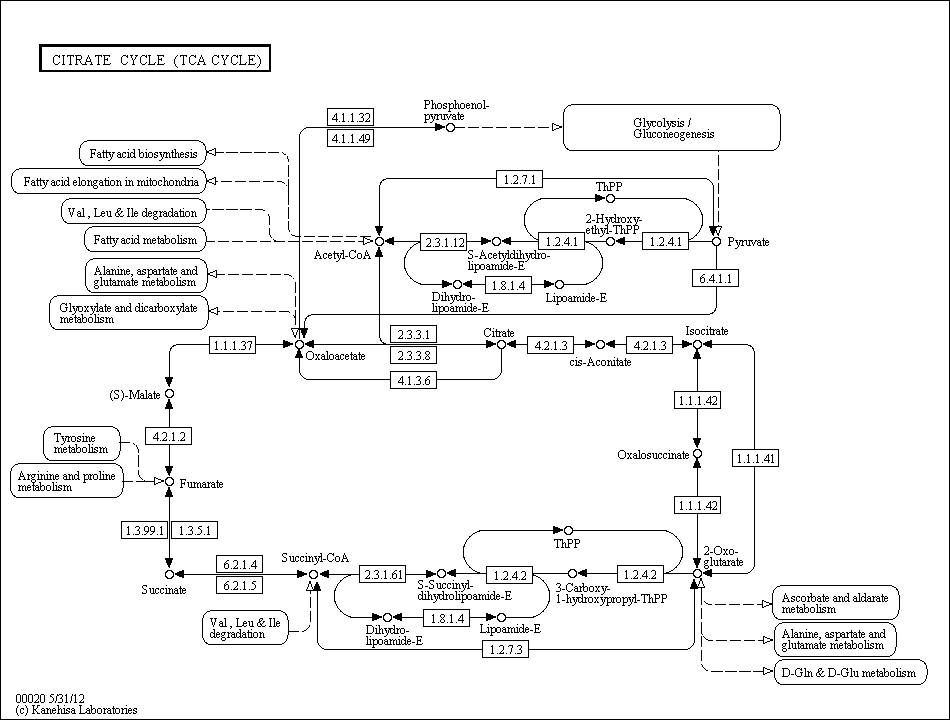


Figure I‑8: An example of a KEGG map

This is a sample KEGG pathway map source from the KEGG website. The representation of this map is the metabolic network of the TCA cycle. Every node in the map represents a metabolite and every square with a number inside illustrates an enzyme. Arrows refer to enzymatic reactions. Reversible reactions are demonstrated by two-way arrows. Dashed arrows show the link between the maps. For instance, the metabolite oxaloacetate also takes part in two metabolic networks: the pathway of Alanine, aspartate and glutamate metabolism and the pathway of Glyoxylate and dicarboxylate metabolism.

Figure I‑8 is only a small part of the metabolic network. An improved reconstruction is the collation of each small parts of the maps. Following the green boxes which link each other, the metabolic flux can go through that means the end product on maps can be produced. The curation process is also required to fix the missing information during annotation process. The representation of these metabolic networks is within KGML files. This format is different from the standard format SBML. Therefore, a format conversion from KGML to SBML is required if FBA is applied to the reconstruction.

## Collaboration in the Department

The development of systems biology is reliant on the improvement of the computational technologies. With studies in systems biology growing rapidly in recently years, numerous software have been established serving these studies. As a result, building a computational model for a cell has become widely available, expanding its usefulness. The models for E. *coli* and yeast, two of the most studied organisms, are well built because the knowledge of these two organisms is sufficient. However, the diversity of organisms results in a difficulty when building models for unknown organisms.

First of all, defining the sequence of an organism is a time consuming work. As an example, the determination of the human genome took 13 years (1990 to 2003). (The Human Genome Management Information System (HGMIS) 2011); this DNA sequencing process required coordination between different groups of people. Secondly, the limitation of experimental data lacks the curation step. Thirdly, the involvement of computational techniques in metabolic reconstruction leads to a difficulty for biologists in the reconstruction process especially using a model.

If the difficulties can be overcome, a reconstruction of metabolic network can be completed with the use of determined gene sequences and experimental data (from laboratory results or published results). As an *in silico* biological model is helpful to learn about the organism, building a computational model for a cell has become a popular objective in biological studies.

Currently, two interesting organisms are studied at the University of Sheffield, *Nannochloropsis gaditana* and *Shewnalla Oneidensis MR-1*. Building computational models of these organisms would be beneficial for guiding laboratory work while contributing to the metabolic engineering community as a whole.

# Model-driven Metabolic Analysis of *Shewanella oneidensis* MR-1

## Microbial Fuel Cell (MFC) using *Shewanella oneidensis* MR-1

*Shewanella* is a microbial bacteria normally living in chemically stratified mediums. It is of interest because of its anaerobic metal-reducing capability. Lactate can be oxidised to reduce the insoluble electrons in *Shewanella*. In other words, metabolism process related to electron transfer and generation is included in *Shewanella* (Fitzgerald, L. A., et al 2012 and Pinchuk, G. E., et al 2010).

### Introduction to the MFC Culture

The culture of MFC bacteria is placed within a tank with a proton exchange membrane (through which only protons can flow pass) (Figure II‑1). The cells are cultured at the anode and the electrons produced by the cells can be transferred to cathode via an electrical connection. At the anode, the typical electrode reaction (using acetate) is CH3COO- + 2H2O 🡪 2CO2 + 7H+ +8e-. The corresponding reaction at the cathode is O2 +4e- + 4H+ 🡪 2H2O, if oxygen is used as the terminal electron acceptor. The reaction changes depending on the choice of terminal electron acceptors (Du, Z., Li, H. & Gu, T. 2007 and Schaetzle, O. Barriere, F. Baronian, K. 2008).



Figure II‑1: A typical two-chamber MFC tank (source from Du, Z., Li, H. & Gu, T. 2007)

### Electron Transfer Mechanisms

At the anode, electrons are produced and transferred by fuel cells via two types of mechanisms, direct electron transfer (DET) mode and mediated electron transfer (MET) mode. Theoretically, the source of electrons is from an organic substrate in the culture medium e.g. lactate. The electrons in these substrates are transferred to the MFC step-by-step to the anode and before finally being accepted by electron acceptors like oxygen (Figure II‑2).



Figure II‑2: Intracellular and extracellular electron transfer paths (source from SCHAETZLE, O. BARRIERE, F. BARONIAN, K. 2008)

DET mode is the transfer of electrons by direct physical contact between cells and the anode or through natural nanowires (known as ‘pili') (Schaetzle, O. Barriere, F. Baronian, K. 2008). The electrons produced in this mode originates from cell respiration. The electrons are firstly obtained by Complex I/II substrates such as formate dehydrogenase and succinate dehydrogenase. They are then passed to quinone (e.g. ubiquinone-8 and menaquinone-7) on the inner membrane. The quinone delivers electrons to Complex III, cytochrome bc1, CymA and DMS/Mtr. These electrons can be transferred to Complex IV via cytochrome c on the outer membrane to catalyse the final electron transfer process (Figure II‑3). In DET mode, the electrons in complex III or cytochrome c can be passed to the anode through direct contact or nanowires (Schroder, U. 2007)

Outer membrane

Inner membrane

Complex I/II

Complex III

Complex IV

Anode

Formate DH

Succinate DH

Nadh DH

Etc.

Ubiquinone8/ubiquinol8

Menaquinone7)/menaquinol7

Methymenaquinone7/methymenaquinol7

Cytochrome bc1

CymA

CymA + Dms/Mtr

Cytochrome oxidase

Figure II‑3: DET mode DET mode (source from Mcmillan, D. Marritt, S. J. et al. 2012)

Unlike DET, in mediated electron transfer (MET) the electron is donated by a natural or artificial redox mediator. The electrons finally contribute to direct redox transformation of excreted metabolites at the electrodes (Schaetzle, O. Barriere, F. Baronian, K. 2008).

Because of the potential in making future electron manufacture material, this organism, Shewanella oneidensis MR-1, is also being studied by another group at the University of Sheffield. It is useful to predict the phenotypes of the cell from computational simulations. The predictions arising from the cell phenotypes may provide general ideas to researches who work in the wet lab.

## Materials and Methods

### The Principle of the Flux Balance Analysis

The method used to analyse the model is called the flux balance analysis (FBA). The principle behind FBA is to convert a metabolic pathway into a stoichiometric matrix in order to solve the problem. For instance, in the metabolic pathway shown in (Figure II‑5), reaction v1 consumes A and D to produce B and E, and can be written as A + D 🡪 B + E. Thus, in a stoichiometric matrix, the parameters for A, B, D and E in reaction v1 are -1,-1, 1 and 1 respectively (negative for reactants, positive for products). Similarly, other reactions can be represented in this way allowing any metabolic pathway to be represented by a stoichiometric matrix. Under steady state, the consumed and produced metabolites are in a balance, which can be written as:

S\*V=0

Where S is a stoichiometric matrix of size m × n, representing metabolic reactions. V is the assembly of vectors which are fluxes through the metabolic network. Vector V has the length of n (Figure II-6, stage b). Compounds (with the number m) are illustrated in rows while reactions (with the number n) are shown in columns. Negative coefficients are assigned to the reactions which consume metabolite while positive coefficients are assigned to the reactions producing metabolites. Under steady state, the solution of this stoichiometric matrix reflects the cell phonotypes (Figure II‑4).



Figure II‑4: The concept of solution space (source from (Orth, J. D. Thiele, I. et al. 2010))

Solution space refers to the entire coordinate system under an unconstrained context. An allowable solution space is created by adding the constraint SV=0 in addition to upper and lower bounds to vectors. Any distribution of flux in a network can be acquired in this solution space. Additionally, a single-flux distribution can be identified by optimising to an objective function. The optimised solution of single flux is located on the edge of the allowable solution space (Orth, J. D. Thiele, I. et al. 2010).



Figure II‑5: Principle of Flux Balance Analysis (source from WIECHERT, W. 2001)

In the picture, the metabolites and reactions are solved by FBA. ‘b’ represents exchange fluxes across cell membranes, while ‘v’ illustrates intercellular fluxes. The metabolites and reactions are represented in a stoichiometric matrix S. Solving the linear problem gives the values of fluxes which equates to the speed of the reactions. The value of the biomass reaction demonstrates the cell growth rate.

### Saving and Loading of the Model

There are two ways to store the information from models: with the SBML format or with the within a Microsoft Excel format file (\*.xlsx). The former is a systems biology mark-up language, which is the standard format that is used in systems biology to save model information. It can be quickly loaded by a computer but is difficult to read by a human. For that reason, in this research, the latter is used. The abbreviation, names, reactions, upper bounds, lower bounds, et cetera are inputted into different columns in Excel, while each row represents a single reaction. An important advantage is that the model can be edited directly using Excel (for example, changing a parameter in a reaction or adding a new reaction).

### Calculating the Growth Rate

To analyse the model, the COBRA toolbox (cobratoolbox) based on MATLAB is used. Before calculating the growth rate, the initial parameters as well as an objective function (usually biomass reaction) have to be set. A MATLAB function is written to export the result to Excel format for convenience.

### Calculating the NADH Productivity

The calculation of growth rate is the basic function of ‘cobratoolbox’, furthermore there are advanced uses of ‘cobratoolbox’ such as calculating the production of co-factor (ATP and NADH), gene-knockout simulation and alternative optimal growth. Due to NADH being related to electron transport chain, the productivity of NADH needs to be calculated. To perform this, a reaction is added to consume free NADH in model, namely ‘NADH 🡪 NAD + H’. Setting the growth rate to a specific value and change objective function to this reaction, the optimal value will be the free NADH in the cell.

### Alternative Optimal Value

Because the FBA result is not unique, it is necessary to calculate alternative optimal values. To perform this, the growth rate has to be set to a specific value. Then, the objective function is changed to the reaction which need to be optimised and calculate the maximum and minimum values of this reaction. A range of optimal values of the target reaction is then obtained.

However, since there are over 700 reactions in the model, manually testing all alternative optimal values are impossible. Thus, a MATLAB function is created to help test the alternative optimal values from the first reaction to the last. The result can then be outputted to Excel format.

## Model Reproduction

### Modelling metabolism of a Cell

Theoretically, modelling metabolism of a cell or organism can be divided into four steps: automatic reconstruction, curating the draft reconstruction, converting to computational model and flux balance analysis.

**Automatic reconstruction**

In the first step, genes of the target organism are collected and analysed automatically using computational algorithm programs. The program compares the similarity between the target genes and a chosen set of genes whose functions are already known. Usually, an algorithm program gives a correspondent score (also known as a ‘hit’) to the result in which the target gene has the probability of functioning in relation to a gene in the chosen set. A higher score correlates to a higher probability. In the final results, the target genes are annotated with different functions by selecting the highest score from their first algorithm results. The assembly of these gene annotations represents an overview of metabolites and enzymatic reactions (which have a probability of occurring) in the target organism.

On the other hand, errors are included in the automatic reconstruction due to the diversity of organisms. Even organisms which appear to be highly similar have differences. This results in no 'hits' for some of the target genes in the reference gene group, or alternatively, wrong ‘hits’. Consequently, the draft reconstruction has to be curated (Feist, A. M. Herrgard, M. J. et al. 2009).

**Curating the draft reconstruction**

There are primarily two types of missing information in the draft reconstruction: gaps and orphans. Gaps are the missing reactions between existing reactions while orphans are reactions without any links to other existing reactions (Orth, J. D. and Palsson, B. Ø. (2010).

In order to improve the reconstruction, the second step of curating is introduced. Some reactions, which are gaps, have a high probability of existence that can be filled manually while some orphans, which are known to not exist in the organism, can be subsequently removed from the pathway maps. Organism specific databases, textbooks, literature and knowledge from experts are used as references for the curation process (Feist, A. M. Herrgard, M. J. et al. 2009).

(Orth, J. D. and Palsson, B. Ø. (2010) Software have also been developed for curating the reconstruction. For instance, BNICE GapFil and SMILEY are used for filling gaps; PHFiller-GC, SEED and ADOMETA are used to fix orphans. (Orth, J. D. and Palsson, B. Ø. 2010) Another step in draft curation is to add organism-specific pathways as some metabolisms only exist in certain organisms.

**Converting to *in-silico* model**

Following curation, the model is transferred into a computer for analysis the in next step (Feist, A. M. Herrgard, M. J. et al. 2009). The information about the target cell or organism is saved in the format, Systems Biology Markup Language (SBML) (Hucka, M. Finney, A. et al. 2003). This is a standard language used in systems biology for representing biological information. Another standard language used in systems biology is the KEGG Markup Language (KGML). (KGML (KEGG Markup Language) 2013) Both SBML and KGML were developed based on the format Extensible Markup Language (XML).

**The usage of reconstruction**

Lastly, the reconstruction process can be used for integration of high-throughput data (Feist, A. M. Herrgard, M. J. et al. 2009) or computational analysis. The most popular approach is the flux balance analysis (FBA), which is a mathematical method for analysing metabolic fluxes through a network (Figure II-6).

The FBA method can be carried out using the Constraints Based Reconstruction and Analysis (COBRA) toolbox under a MATLAB system environment. SBML files can be directly read by the COBRA toolbox, representing the biological information in MATLAB structure on which the calculation of FBA is based on (Schellenberger, J. Que, R. et al. 2011).

Due to the simplicity in analysing networks afforded by FBA, various applications can be implemented. For example, simulating the growth rate of a reconstructed organism in different media, gene knockouts of bacteria and mutants of an organism (Orth, J. D. Thiele, I. et al. 2010).

Systems biology offers an alternative way to study organisms with the ability to predict the behaviours of a cell without a lab work. This prediction may not be 100 percent accurate because of knowledge or data limitations. However, it is sufficiently accurate for it to be an economic and rapid way of directing lab studies.

A successful example of metabolic engineering is the manipulation a mevalonate pathway in *E. coli* in order to make terpenoids. The genes expressing the mevalonate-dependent pathway in S.*cerevisiae* were selected and added into *E. coli*, giving the E. *coli* the ability the manufacture terpenoids (Martin, V. J. J. Pitera, D. J. et al. 2003).



Figure II‑6: Metabolic reconstruction and Flux balance analysis (source from Orth, J. D. Thiele, I. et al. 2010)

1. A genome-scale metabolic reconstruction starts with determining the genome sequence. After steps 1, 2 and 3 (which are described above), the metabolic information can be represented mathematically.
2. The core of FBA: S\*V=0. Fluxes are illustrated in V. The relation between metabolites and reactions is represented in the stoichiometric matrix S.
3. The linear equations are defined by a mass balance. This balance is simulated at a steady state.
4. Defining objective function (z=cTv) to predict the maximum growth rate of an organism.
5. The distribution of flux is solved by linear programs. Z is the direction of growth rate and the point of optimal v is one of the best solution in the defined solution space (Orth, J. D. Thiele, I. et al. 2010).

### The Reconstruction Process

The model used in this project was published by Grigoriy E. Pinchuk et al in 2010. The published network consisted of 634 metabolites, 774 reactions and 783 genes, and includes a branch of electron transport chains. Usually a model is written in a SBML file, however, the model published by Pinchuk et al was saved in several Excel files for convenient access and reading. Therefore, a conversion to a computer-friendly format was required.

Key information, metabolic reactions and metabolites, were written in ‘table s2’ and ‘table s3’ in Pinchuk, G. E., et al 2010 work respectively. There are two ways to translate that information. One way is to form the metabolic data in SBML format, which is the standard format of systems biology, another way is to reproduce the data into a machine-readable Excel format. Because a loading function that loads model from Excel format already exists in cobratoolbox, transferring the data to a specific Excel structure was more practical. The conversion from Excel format to SBML format is more complicated than Excel-to-Excel transfer, furthermore, data in presented in Excel format is more convenient for human reading. Thus, reproducing the model in Excel format is faster and more reasonable. Hence, the format can be easily transferred by cobratoolbox.

A cobratoolbox format Excel model includes two spreadsheets, reactions and metabolites. In the spreadsheet ‘reactions’, the title is written in the first row, while abbreviation, name, reaction and the rest of information are written from column A to O. All information about metabolites is included in the separate spreadsheet, ‘metabolites’.

The first reproduced model was successfully loaded into MATLAB, but did not result in any solution. This was caused by a discrepancy in the reaction format: in MATLAB, the reaction is written as A + B --> 2 C, which is different in Pinchuk et al’s reconstruction, whose coefficients are included in brackets, i.e. A + B --> (2) C. The bracket lead to the unexpected mistake of coefficients loading, resulting in no solutions during the optimisation process. Recognising this, the brackets were then removed using MATLAB codes. The second reproduced model, without brackets on the coefficients, showed an acceptable solution. Information on a missed metabolite, chromium (III) hydroxide, was also added to the ‘metabolites’ sheet as the 635-th metabolite. The reproduced model thus consisted of 635 metabolites, 774 reactions and 783 genes.

### The Metabolite Function of the Model

The metabolic function of *Shewanella* model has been checked. Crucially, the central metabolisms: the glycolysis / gluconeogenesis pathway (Figure II-7), the Citrate cycle (Figure II-8) and the Pentose phosphate pathway (Figure II-9) (sourced from the KEGG pathway) were included in the model. This means that the *in silico* cells could be grown under certain conditions. However, there was some missing information on KEGG pathways. For example, in Figure II-7, reaction 1.1.1.27 is missing (not in green box). This missing reaction exists in the Grigoriy E. Pinchuk et al’s reconstruction. Thus, the lactate uptake from culture medium could be used in either the TCA cycle or the Pentose phosphate pathway through the glycolysis / gluconeogenesis pathway.

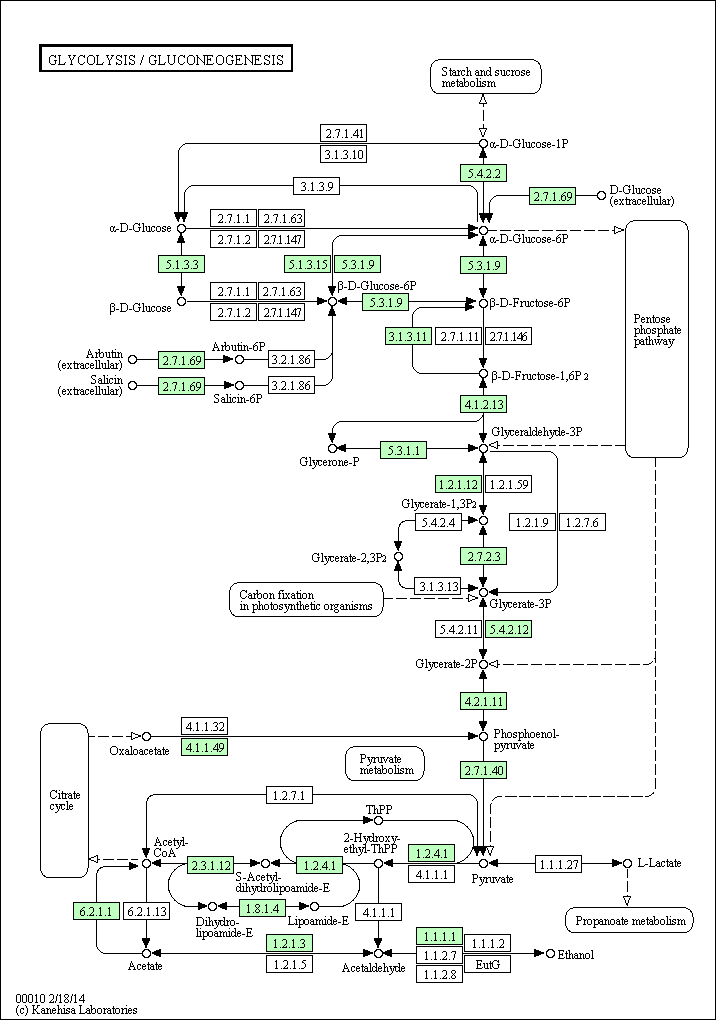


Figure II‑7: Glycolysis / gluconeogenesis pathway

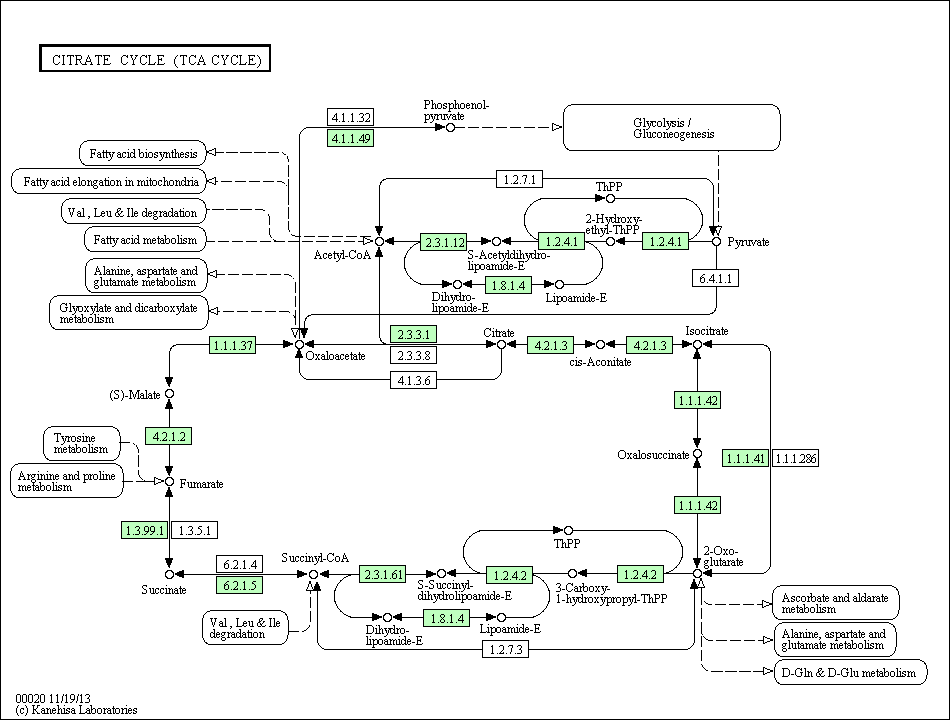


Figure II‑8: Citrate cycle

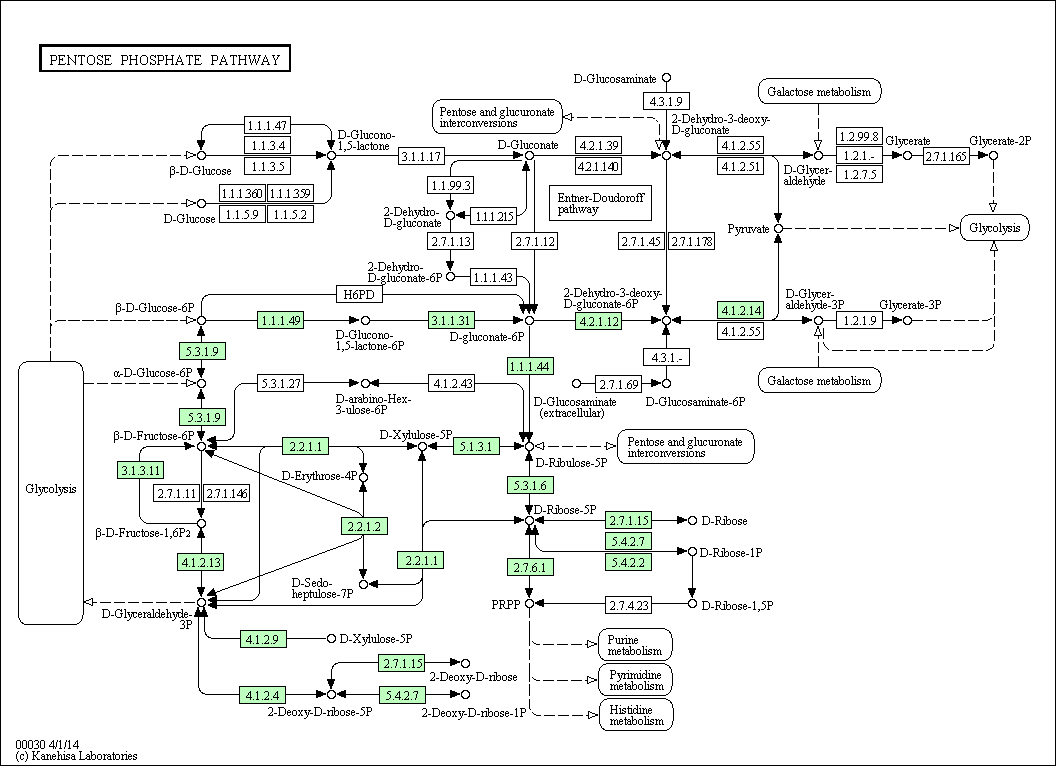


Figure II‑9: Pentose phosphate pathway

### Model adjusting

The next step in reproducing the model was to adjust errors within the model reproduction process or the model itself. Because a model is never completely identical to actual cells, the reliability of models is relative lower to experimental data.

Three oxidases, two cytochrome c oxidases (Cco, SO2361-2364; Cox, SO4606-4607, SO4609) and one cytochrome d ubiquinol oxidase (Cyd, SO3285-3286), using oxygen as the terminal electron acceptor in electron transport chain were observed for electron productivity under aerobic conditions. Two Fe (III) Reductases (menaquinone-7/methylmenaquinone-7), which use Fe3+ as their terminal electron acceptor, were observed for electron productivity under anaerobic conditions. Because there is a maximum electron iron translocation ratio of 2.8H+/2e-, the flux ratio of Cyd and Pet-Cco/Pet-Cox was 1:4. The published growth rates were considerably well matched with experimental data under the growth rate of 0.085/h. The experimental data was slightly higher than the model after 0.085/h.

A growth rate was calculated by setting all negative boundaries to -1000 for reversible reactions, 0 for irreversible reactions and all positive boundaries to 1000 for both reversible and irreversible reactions. This growth is extremely large and meaningless but the model is mathematically correct. The initial parameters for the model were obtained from the M1 medium (for composition see Supplemental data I).

Because the standard unit for growth rate in cobratoolbox is mM/gDW/h (milli-mole per grams dry weight per hour), the unit ‘grams per litre’ was converted to ‘milli-mole per litre’. The total iron was separately calculated to fit in the model (for more details see Supplemental Data I & II).

The negative bounds of the corresponding exchange reactions in the model were set to specific negative values for the uptake rate. The positive bounds for all exchange reactions were set to 1000 assuming all molecules can be secreted out of cells but only a limited amount of nutrients can be up-taken by cells. The set of opposite way of parameters leads to a large growth having the same slop of growth with experimental data.

The simulation of growth under various lactate uptake rate is shown in Figure II-10. There is a slight difference in the simulation between this research and published model. The simulation found that the growth rate was 0.0874/h when the lactate uptake rate was 4.08 mmol/h under aerobic conditions. Meanwhile, the growth rate for the published model was 0.085/h under the same conditions. The difference in growth rates could be explained by the use of different analytical software (GAMS was used for the published work).

Figure II‑10: The maximum growth rate vs lactate consumption rates. The horizontal axis is lactate consumption rate, vertical axis is the maximum growth rate that the cell can be reached.

Table II‑1: Simulation Comparison

|  |  |  |
| --- | --- | --- |
| Result Source | Lactate consumption rate (mM/gDW/h) | Growth rate  (/h) |
| G. E. Pinchuk et al 2010 | 4.08 | 0.085 |
| Experimental measurements | 4.06 | 0.085 |
| Present simulation | 4.08 | 0.0874 |

A comparison of maximal biomass yield and O2 requirements with Pinchuk et al’s results was also carried out. It was found that in addition to the growth rate, the biomass produced by all carbon sources was marginally higher in the present work (Table II-2), while the oxygen requirements in this study were found to be slightly smaller. Despite these differences, the results of this study could be considered to be in agreement with Pinchuk et al’s work.

One possible reason might be the use of different software causing the difference, while the difference in initial parameters is another probability (for details of the parameters used, see the Supplemental File). The parameters used in this study were directly obtained from recipe of media, the differences possibly occuring between parameters of this study and Pinchuk et al’s work.

Glycerol in Table II-2 and Figure II-11 were predicted in this work only, therefore a comparison was not performed. Two reactions, one exchange and one transport reaction for glycerol, were introduced into the model for this prediction.

Table II‑2: Comparison of O2 requirement and biomass yield. \*Glycerol was studied in this work only.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Pinchuk et al 2010** | | **Present work** | |
| Component | O2 Requirement (mmol O2/ mmol carbon source) | Biomass Yield (gAFDW/ mmol carbon source) | O2 Requirement (mmol O2/ mmol carbon source) | Biomass Yield (gAFDW/ mmol carbon source) |
| Acetate | 1.4034 | 0.0132 | 1.3926 | 0.0134 |
| N-Acetyl-D-glucosamine | 5.0894 | 0.0643 | 5.0783 | 0.0646 |
| Formate | 0.3401 | 0.0035 | 0.3312 | 0.0037 |
| Fumarate | 1.9162 | 0.024 | 1.905 | 0.024 |
| L-Glutamine | 2.8339 | 0.0368 | 2.8277 | 0.037 |
| D-Lactate | 2.042 | 0.0212 | 2.031 | 0.0214 |
| Propionate | 2.4674 | 0.0228 | 2.4566 | 0.023 |
| Pyruvate | 1.6157 | 0.0195 | 1.6049 | 0.0198 |
| L-Serine | 1.6009 | 0.0199 | 1.5899 | 0.02 |
| Succinate | 2.341 | 0.0256 | 2.33 | 0.0258 |
| Glycerol\* | N/A | N/A | 1.815 | 0.0372 |

The biomass yield and corresponding oxygen requirement rises as the number of carbon molecules increases in the structure of the carbon sources. Formate (CHO2), which has only one carbon atom and the simplest structure within the 11 carbon sources, shows the smallest biomass yield and oxygen requirement. Acetate (C2H3O2) has two carbon molecules that can produce a considerably higher biomass yield than formate. The largest biomass yield was produced by N-Acetyl-D-glucosamine, a large eight carbon molecule structure. Interestingly, glycerol has a similar biomass yield than fumarate and D,L-Lactate, but requires a lot more O2.

Figure II‑11: Biomass yield by varies O2 requirement referring to Table II-2. The maximum growth is reached under a 10 mmol/h carbon uptake rate and its corresponding oxygen requirements are illustrated for 11 model-predicted carbon sources in this research. The 10 blue points are results by Pinchuk et al, while the sole red point (glycerol) was predicted by this research only.

Table III‑3: Two reactions for glycerol prediction. glyc, [e], and [c] represent glycerol, exchange, and cytoplasm respectively.

|  |  |
| --- | --- |
| Reaction Name | Reaction Formula |
| Glycerol exchange reaction | [e]: glyc <==> |
| Glycerol transport reaction | glyc[e] <==> glyc[c] |

Figure II‑12: Maximum biomass yield against oxygen consumption rate. The horizontal axis represents the oxygen consumption rate (mmol O2/gAFDW/h), and the vertical axis represents the maximum biomass yield (mgAFDW/mmol carbon source) that the cell can reach based on three carbon sources. The carbon uptake rates were set to 10 mmol/gAFDW/h.

The graph of maximum biomass yield Figure II-12 using the three carbon sources, pyruvate, lactate and propionate was reproduced. Similar to Pinchuk et al, propionate had the highest biomass yield within the three carbon sources and the lowest maximum biomass yield was achieved by pyruvate.

The results of this work illustrates a high degree of similarity with Pinchuk et al’s work as well as to experimental data, validating the reproduction of the model. Hence, the basis for the rest of the research work can be considered reliable.

## Maximisation of electron fluxes

### Electron transport chains in direct electron transport (DET) mode

Following model adjustment, the phenotype of MR-1 can be explored. The most attractive aspect in the phenotype of MR-1 is electron productivity because MR-1 is one of the MFC organisms geared specifically for electricity. There are two ways to donate electrons from MFCs that were mentioned in the introduction: MET and DET. Therefore several reactions in the model reconstruction have to be observed.

In DET mode, two different reactions (Table II-7) donate their electrons to the terminal electron acceptor, the enzymes of the cytochrome family for aerobic respiration and other enzymes relating to menaquinone-7 and methylmenaquinone-7 for anaerobic condition (Table II-8). The enzymes cytochrome-c oxidase and cytochrome oxidase bd (ubiquinol-8: 2 protons) were selected for aerobic conditions because oxygen is involved in the reaction, producing water as the final product. Fe3+ has been used for anaerobic conditions, thus, Fe (III) Reductase was selected.

The terminal electron acceptor varies in *Shewanella* (there are 12 terminal electron acceptors in the model of *Shewanella* under anaerobic conditions (Table II-8), such as Fe3+, NO3-, etc., each with different redox potentials. The electron is passed from a high-redox potential to a low-redox potential through a step-by-step process, thus, the energy in the chemical compound can be fully transferred to currency metabolites such as ATP and NADH.

In the stoichiometric matrix, there is one reaction which corresponds to each electron acceptor. Because the rest of the reactions in the metabolic pathways are the same, and only the last reactions related to the acceptor are different, the calculated results would be the same if the molecules of the acceptor are sufficient in the media under anaerobic conditions. In other words, the redox potential does not have influence on the simulation results for electron productivity. Therefore, the choice of a single terminal electron acceptor can be used to simulate all anaerobic conditions. Hence, in this study, Fe3+ was selected as the terminal electron acceptor to simulate the growth of *Shewanella* under anaerobic conditions.

### DET mode under aerobic conditions

In the DET mode, the electron transfer by cytochrome c and ubiquinol-8 under aerobic conditions is through physical contact between the cell membrane and the anode. Ubiquinone-8 /ubiquinol-8 play a major role in electron transportation under aerobic conditions. It delivers electrons from complex I and complex II, such as formate dehydrogenase, to complex III (Duncan G. et al 2012). Next, the electron is passed to the cytochrome c or cytochrome oxidase bd and then released to the anode by physical contact.

The metabolites that donate their electrons to the electron shuttle, ubiquinone-8 vary. These metabolites are also involved in different metabolisms. For example, succinate can donate its electron to ubiquinone-8 to produce fumarate. This is an important reaction in the Citrate Cycle. The currency metabolite NADH is another electron donor. It is assumed that obtaining the electron from a different metabolism leads to a different electron flux as well as a different growth rate. For instance, more electrons obtained from succinate will increase the flux values within the TCA cycle, possibly leading to an increase of growth rate. Therefore, investigating the efficiency of electron donors’ gain is of interest.

To obtain the total electron donors, the model has been investigated. Nine metabolites were found to be able to donate their electrons to ubiquinone-8 under aerobic conditions (Table II-3). During experimentation, NADH dehydrogenase (Ubiquinone-8 & 4 protons) showed the highest efficiency (most electrons were donated by NADH) (Figure II-13). Simulation of *Shewanella* growth in M1 media under aerobic conditions resulted in a lactate uptake rate of 4.08 mmol/h and a growth rate of 0.0874 /h. Under this context, flux through NADH dehydrogenase (Ubiquinone-8 & 4 protons) was 9.7741 mmol/h. Meanwhile, succinate dehydrogenase had the second best efficiency with a flux of 6.799 mmol/h. Because of the enzymatic limitation, the flux through the two electron reactions: Cytochrome-c oxidase (2 protonstranslocated) and Cytochrome oxidase bd (ubiquinol-8:2 protons) were set to 4:1. Thus, the electron fluxes for those reactions were 3.3145 mmol/h and 13.2584 mmol/h. Then, as NADH dehydrogenase (Ubiquinone-8 & 4 protons) had the highest flux value, it can be seen that NADH provided most electrons via this reaction.

In the next step, both lower and higher bounds of this NADH reaction were set to zero to simulate the knockout of the reaction. The knockout of NADH dehydrogenase (Ubiquinone-8 & 4 protons) reaction resulted in a nil flux pass for this reaction. However, the growth rate remained the same, at 0.0874 /h. Succinate in the TCA cycle provided most electrons via succinate dehydrogenase, with a flux value of 16.573 mmol/h. The electron fluxes remained at the same levels. This shows that the reaction succinate dehydrogenase can work individually in electron donor reactions, maintaining the growth and electron reactions at the same levels as wild type cells.

Next, the knockout of the succinate dehydrogenase reaction was tested. The upper and lower boundaries of NADH dehydrogenase (Ubiquinone-8 & 4 protons) were set back to 1000 and 0 respectively and the succinate dehydrogenase reaction was shutdown. The growth was slightly decreased (0.0733 /h) in the succinate dehydrogenase mutant but the electron fluxes were slightly higher than the wild type and NADH dehydrogenase (Ubiquinone-8 & 4 protons) mutants. Most electrons were donated by NADH (13.7482 mmol/h) while 4.08 mmol/h of electrons were provided by D/L-lactate via D/L-lactate dehydrogenase.

In contrast, the double knockout of enzymes NADH dehydrogenase (Ubiquinone-8 & 4 protons) and succinate dehydrogenase were the most interesting mutants. The growth rate of this mutant dropped to 0.0289 /h but with a significant increase of electron product: 21.8536 mmol/h electrons in total. 14.8729 mmol/h of electrons were provided by D/L-lactate, while 6.9737 mmol/h of electrons were donated by L-Proline to produce 1-Pyrroline-5-carboxylate.

When three enzymes NADH dehydrogenase (Ubiquinone-8 & 4 protons), succinate dehydrogenase and D/L-lactate dehydrogenase were knocked out from the model, the cell showed an extremely low growth rate (also known as ‘zero growth’). In other words, the cell would be dead if the three main enzymes related to the electron donors were deleted because the remaining electron donors cannot provide sufficient electrons to the electron transport chains. This leads to fewer H+ ions are pumped to the outer membrane, decreasing the efficiency of ATP synthase.

Table II‑3 Metabolites involved in ubiquinol-8 producing reactions. There are 9 different electron donors in the reconstruction. NADH is the currency metabolite that provides most of the electrons during simulations. Succinate is essential in the TCA cycle. Three fatty acid precursors are also included: tetradecanoate, hexadecanoate and octadecanoate.

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Reactants | Productions | Subsystems |
| 1 | L-Aspartate | Iminasspartate | Cofactor and prosthetic group biosynthesis |
| 2 | (s)-Dihydroorotate | Orotate | Purine and Pyrimidine Biosynthesis |
| 3 | tetradecanoate | Fatty acid | Alternate Carbon Metabolism |
| 4 | hexadecanoate | Fatty acid | Alternate Carbon Metabolism |
| 5 | octadecanoate | Fatty acid | Alternate Carbon Metabolism |
| 6 | D/L-lactate | Pyruvate | Alternate Carbon Metabolism |
| 7 | L-Proline | 1-Pyrroline-5-carboxylate | Arginine and Proline Metabolism |
| 8 | Succinate | Fumarate | Citrate Cycle |
| 9 | NADH | NAD | Energy metabolism |

To further investigate what happens inside of the cell, the distribution of wild type cells and the double knockout cell of simulation 4 in Table II-3 have been selected to study (see Figure II-13). The distribution of values is TCA cycle, which is important in aerobic respiration and a key metabolism in most organisms in the world. In this metabolism, lactate-D is firstly converted to pyruvate, and the electrons are accepted by methymenaquinone-7 in wild type c

ells.

Under aerobic conditions, the major electron shuttle is ubiquinone-8 which delivers the electron to the cytochrome family. Thus, the electrons accepted by methymenaquinone-7 are not essential under these conditions. In contrast, in the double knockout cell, lactate-D donated their electrons to uniquinone-8, resulting in a flux value of 14.8729 mmol/h. This flux value was much larger than the lactate uptake rate, 4.08 mmol/h, because the excess 10.7929 mmol/h lactate-D was produced from pyruvate with electrons provided by NADH. This may mean that the enzyme NADH dehydrogenase (Ubiquinone-8 & 4 protons) was being knocked out of the cell, preventing NADH donating their electrons to unibquinone-8 via this reaction, thus it repeatedly produced lactate to increase electron donation. This could be the reason that growth dropped dramatically under double knockout cell: more NADH was consumed to produce lactate instead of being used for growth.

Then, the flux flow from pyruvate to acetyl-CoA enters the TCA cycle. In the TCA cycle, the overall fluxes of wild type cells are larger than the fluxes of double knockout cells. The fluxes of double knockout are close to zero, because the reaction succinate dehydrogenase is one of the double knockout reactions. Succinate as a reactant in this reaction is catalysed by succinate dehydrogenase to produce fumarate: the electrons in succinate are accepted by ubiquinone-8 to be further utilised in the electron transport chain. Due to the shutdown of this reaction, it is unable to generate electrons in the TCA cycle and the flux is stopped at succinate because no more fumarate can be converted. Thus, the electron donating process is shifted to the lactate dehydrogenase reaction, resulting in a high flux value through lactate to pyruvate and low flux values in the TCA cycle. Referring to Table II-4, the reduction of flux values in the TCA cycle leads to a decrease in growth rate (0.0289 /h growth rate for double knockout and 0.0874 /h for wild type), but increases the overall flux value of electron reactions (21.8536 mmol/h electron flux for double knockout and 16.573 mmol/h electron flux for wild type).

pyruvate

Oxaloacetate

Citrate

Isocitrate

2-Oxoglutarate

Succinyl-CoA

Succinate

Fumarate

L-Malate

ubiquinone8

ubiquinone8

ubiquinol8

ubiquinol8

methymenaquinone7

methymenaquinol7

Coenzyme A

Lactate-D

0/0/(14.8729)

Acetyl-CoA

7.196/. 3.870/(7.3910)

0.0401/2.7956/(0.0320)

0.0401/2.7956/(0.0320)

0/1.9883/(0)

0.0401/2.7956/(0.0320)

0.0145/1.9534/(0.0116)

0/6.7990/(0)

0.0255/2.7606/(0.0204)

0.0273/3.4759/(0.0218)

0/0/(10.7929)

NADH

NAD

4.08/4.08/(0)

Figure II‑13 The comparison of the flux distribution in TCA cycle between wild type growth under anaerobic and aerobic conditions, and the double knockout growth. The left value in yellow box is the distribution value of wild type growth anaerobically, the value in the middle is the aerobic growth of wild type cells, the right value in bracket is the distribution value of double knockout. The units are mmol/h.

Table II‑4 Gene knockout simulations of ubiquinone-8 producing reactions. The red x is the knockout of the gene, the units of growth are /h, the rest units are mmol/h.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **Simulations and values** | | | | |
| **Names** | **Reactants** | **Productions** | **1** | **2** | **3** | **4** | **5** |
| NADH dehydrogenase  (Ubiquinone-8 & 4 protons) | NADH | NAD | 9.7741 | X | 13.7482 | X | X |
| Succinate dehydrogenase | Succinate | Fumarate | 6.799 | 16.573 | X | X | X |
| D/L-lactate dehydrogenase | D/L-Lactate | Pyruvate | 0 | 0 | 4.08 | 14.8729 | X |
| NADH dehydrogenase  (ubiquinone-8) | NADH | NAD | 0 | 0 | 0 | 0 | 0.0138 |
| Proline oxidase | L-Proline | 1-Pyrroline-5-carboxylate | 0 | 0 | 0 | 6.9737 | 0.009 |
| Dihydoorotic acid dehydrogenase (ubiquinone-8) | (S)-Dihydroorotate | Orotate | 0 | 0 | 0 | 0.0065 | 0 |
| NADH-Ubiquinone Oxidoreductase (Na+ translocating) | NADH Na[c] | NAD Na[e] | 0 | 0 | 0 | 0 | 0 |
| L-aspartate oxidase | L-Aspartate | Iminasspartate | 0 | 0 | 0 | 0 | 0 |
| **Electron reactions** | | |  |  |  |  |  |
| Cytochrome-c oxidase (2 protonstranslocated) | | | 3.3146 | 3.3146 | 3.569 | 4.3707 | 0.0046 |
| Cytochrome oxidase bd (ubiquinol-8:2 protons) | | | 13.2584 | 13.2584 | 14.2578 | 17.4829 | 0.0182 |
| **Sum of electron producing reactions** | | | **16.573** | **16.573** | **17.8268** | **21.8536** | **0.0228** |
| **Growth rates** | | | **0.0874** | **0.0874** | **0.0733** | **0.0289** | **2.42E-05** |

Table II‑5 gene knockout simulations of menaquinone-8 producing reactions. The red x is the knockout of the gene, the units of growth are /h, the rest units are mmol/h.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **Simulations and values** | | | | |
| **Names** | **Reactants** | **Productions** | **1** | **2** | **3** | **4** | **5** |
| NADH dehydrogenase (methylmenaquinone-7 & 4 protons) | NADH | NAD | 17.108 | X | 0 | X | X |
| D/L-lactate dehydrogenase | D/L-Lactate | Pyruvate | 4.08 | 0 | X | X | X |
| dihydroorotate dehydrogenase | (S)-Dihydroorotate | Orotate | 0.0081 | 0 | 0 | 0 | X |
| L-aspartate oxidase | L-Aspartate | Iminoaspartate | 0 | 0 | 0 | 0 | 0 |
| Formate Dehydrogenase (methylmenaquinone-7: 1 protons) | Formate | CO2 | 0 | 0 | 0 | 0 | 0 |
| glycerol-3-phosphate dehydrogenase (methylmenaquinone 7) | sn-Glycerol 3-phosphate | Dihydroxyacetone phosphate | 0 | 0 | 0 | 0 | 0 |
| Hydrogenase (methylmenaquinone-7) | H2 |  | 0 | 0 | 0 | 0 | 0 |
| NADH dehydrogenase (methylmenaquinone7 & no proton) | NADH | NAD | 0 | 0 | 0 | 0 | 0 |
| **Electron reactions** | | |  |  |  |  |  |
| Fe (III) Reductase (methylmenaquinone 7: 2H per 2 e-) | | | 21.1962 | 0 | 0 | 0 | 0 |
| **Sum of electron producing reactions** | | | **21.1962** | **0** | **0** | **0** | **0** |
| **Growth rates** | | | **0.0363** | **0** | **0** | **0** | **0** |

Because the flux distributions by FBA are not unique, it is possible that one flux has a multitude of values under one condition. Thus, the flux variability analysis (FVA) is introduced. This algorithm identifies the minimum and the maximum possible fluxes for one particular reaction when the objective value is set to a constraint (ORTH JEFFREY D et al 2010).

Next, the FVA tests for both wild type and double knockout were completed. To perform the FVA, the growth rate as an objective function was firstly constrained to 0.0873 /h and 0.0288 /h respectively. The reason that the growth rate in FVA was slightly smaller than the optimal value in FBA is that the actual optimal value in FBA was not as large as the value printed on screen, it possibly close to printed value and be added to this value automatically by program due to the limited display space. Thus, the printed value was slightly larger than the actual value. The use of the printed value will have no solutions as a result.

Secondly, the objective function was set to the particular reaction. Next, the optimisation function from cobratoolbox was run to calculate the minimum and maximum fluxes of the target reaction. In order to improve the efficiency of the FVA calculation, the cobratoolbox has a function to run the FVA test for the whole model. However, this function destroys the structure of the model saved in MATLAB. Therefore, it was necessary to save the model before running this function. The function has been re-written in this research project to simplify the process of reproducing the model into an Excel format. The flux values could then be easily selected and picked out.

For comparison with the FBA distribution, the values in the TCA cycle for two FVA tests of wild type and double knockout were selected and recorded in Figure II-14 and Figure II-15. From the FVA test, it can be seen that most of the flux values in the TCA cycle maintain a small range except for the electron donation reactions. For example, in wild type cells (Figure II-14), the flux value from fumarate to L-Malate was between 2.68 mmol/h and 2.847 mmol/h. The two exceptions were the reaction from Succinyl-CoA to succinate and Isocitrate to 2-Oxoglutarate. The former has a negative flux value that means Succinyl-CoA can be produced from both Succinate and 2-Oxoglutarate. The latter has a small, minimal value of 0.64 mmol/h. The maximum value of the latter was at the same level with other values within TCA cycle. This was because the 2-Oxoglutarate can be produced from either Isocitrate or Citrate, the production from Citrate will lead to a decrease of the production

Pyruvate

Oxaloacetate

Citrate

Isocitrate

2-Oxoglutarate

Succinyl-CoA

Succinate

Fumarate

L-Malate

ubiquinone8

ubiquinone8

ubiquinol8

ubiquinol8

methymenaquinone7

methymenaquinol7

Coenzyme A

Lactate-D

0/4.232

Acetyl-CoA

3.203/4.022

2.724/2.882

2.724/2.88

1.844/2.613

0.64/2.7

1.81/-2.57

2.635/16.605

2.68/2.847

2.865/3.630

0/0.152

NADH

NAD

0/4.232

0.64/2.709

NADP/NADPH

Figure II‑14 FVA test of wild type cells. The left values in yellow box are the minimum flux values, the right values are maximum flux values. Positive values is the reaction acts the same way of arrow and the negative value is the reaction acts the opposite way of the arrow. The units are mmol/h.

from Isocitrate. The overall production of 2-Oxoglutarate was less than the productivity of Isocitrate and Citrate (2.88 mmol/h).

In the double knockout simulation (Figure II-15), the value in TCA cycle maintains in a low level due to the knockout of the reaction from Succinate to Fumarate. Both minimum and maximum values in the Krebs cycle were close to or equal to zero. This is the main reason that the growth rate decreased dramatically. One exception is that over 7.019 mmol/h, Isocitrate can be produced from 2-Oxoglutarate. This might be because of the link of 2-Oxoglutarate with other metabolisms such as Ascorbate and aldarate metabolism and

Pyruvate

Oxaloacetate

Citrate

Isocitrate

2-Oxoglutarate

Succinyl-CoA

Succinate

Fumarate

L-Malate

ubiquinone8

ubiquinone8

ubiquinol8

ubiquinol8

methymenaquinone7

methymenaquinol7

Coenzyme A

Lactate-D

4.068/21.857

Acetyl-CoA

7.355/7.398

0.03/0.036

0/0

-7.019/0.036

0.011/0.013

0/0

0.02/0.038

0.012/0.04

0/17.777

NADH

NAD

0/0

0.03/0.036

-7.02/0.036

NADP/NADPH

Figure II‑15 FVA test of double knockout cells. In yellow box, the maximum and minimum flux values are illustrated in the right and left respectively. The negative values are the reactions through the opposite way of the flux arrow. The units are mmol/h.

alanine, aspartate and glutamate metabolism (the flux may have come from these metabolisms).

With regards to quinone dependent reactions, the flux distribution shows different results. In the wild type simulation, the electron donated by lactate can be either accepted by uniquinone-8 and methymenaquinone-7. Thus, the minimum and maximum flux values for these two reactions are 0 mmol/h and 4.232 mmol/h respectively. The maximum value of 4.232 mmol/h is greater than the lactate uptake rate of 4.08 mmol/h because 0.152 mmol/h of lactate can be produced by pyruvate via the NADH reaction. The value of ubiquinone-8 related reaction in the Krebs cycle has a range of values, from 2.635 mmol/h to 16.605 mmol/h. The value 2.635 mmol/h is the minimum required value to keep the TCA cycle in a common value, this value is similar to the other value in the Krebs cycle. However, the maximum value 16.605 mmol/h is much higher than the average value in the Krebs cycle. This is possibly because the reaction succinate dehydrogenase can be taken over by the electron donation process when the reaction NADH dehydrogenase (Ubiquinone-8 & 4 protons) is shut down. In contrast, in the double knockout simulation, lactate dehydrogenase dominates the whole electron donation process due to the remaining two efficient ubiquinone-8 reactions being shut down. From 4.068 mmol/h to 21.857 mmol/h, lactate can be catalysed to pyruvate under this condition, meanwhile, from 0 mmol/h to 17.777 mmol/h NADH is used to produce lactate from pyruvate depending on how much lactate are produced.

Similar to FBA, FVA also simulates what could possibly happen inside the cells, but with a range of scenarios. Obtaining the values in FVA is more complex than in FBA, but the result is generally more reliable than FBA. For example, the reaction catalysed by succinyl-CoA synthetase produces succinate from succinyl-CoA. The calculated flux value through this reaction by FBA is, however, in FVA calculation this value can be reversed, meaning the opposite direction of the reaction.

There is a slight difference between the maximum FVA and FBA flux value of the succinyl-CoA synthetase reaction. The calculated maximum flux through succinyl-CoA synthetase is 1.81 mmol/h but the FBA value is 1.9534 mmol/h which is larger than the calculated maximum value. This was possibly because of the difference in optimal growth rate. In FBA, the growth rate was 0.0874 /h which was the objective function that to be optimised. However, in FVA the growth rate was 0.0873 /h to avoid the problem of having no solution, thus the calculated maximum value for succinyl-CoA synthetase was slightly smaller than the value of FBA.

### DET mode Under Anaerobic Conditions

Under anaerobic respiration, the electrons are accepted by menaquinone-7 or methylmenaquinone-7 instead of ubiquinone-8. The terminal electron acceptor varies (see Table II-8) depending on anaerobic conditions or aerobic conditions. There are seven electron donors that can donate their electrons to menaquinone-7 and methylmenaquinone-7 (Table II-5).

Table II‑6 Seven electron donors under anaerobic conditions. [e]: The reaction from formate to CO2 is generated at the outer membrane. When hydrogen is involved in an electron donation reaction, both electrons and hydrogen ions are accepted by quinone so that there are no products produced.

|  |  |  |  |
| --- | --- | --- | --- |
| No | Reactant(s) | Product(s) | Subsystem(s) |
| 1 | L-Aspartate | Iminasspartate | Cofactor and Prosthetic Group Biosynthesis |
| 2 | (S)-Dihydroorotate | Orotate | Purine and Pyrimidine Biosynthesis |
| 3 | Formate [e] | CO2 [e] | Energy Metabolism |
| 4 | sn-Glycerol 3-phosphate | Dihydroxyacetone phosphate | Energy Metabolism |
| 5 | H2 |  | Energy Metabolism |
| 6 | D/L-Lactate | Pyruvate | Alternate Carbon Metabolism |
| 7 | NADH | NAD | Energy Metabolism |

Similar to aerobic conditions, the gene knockout simulation is done under anaerobic conditions (see Figure II-15). However, only the wild type cells (the first) can live anaerobically, the knockout of either NADH dehydrogenase (methylmenaquinone-7 & 4 protons) or D/L-lactate dehydrogenase results the in death of the cells. The growth rate of the wild type cells is 0.0363 /h, with the corresponding value of electron flux as 21.1962 mmol/h. The electrons are mainly donated by two reactions, NADH dehydrogenase (methylmenaquinone-7 & 4 protons) and D/L-lactate dehydrogenase. Over 17 mmol/h of electrons are provided by NADH under anaerobic conditions. Lactate provides 4.08 mmol/h of electrons which equals to the uptake rate of lactate.

From the FVA test of the anaerobic growth of wild type cells (Figure II-16), the fluxes within the TCA cycle remain at a low level, from around 0.036 mmol/h to 0.08 mmol/h. one reaction that can be reversed in the TCA cycle is the reaction from 2-Oxoglutarate to Isocitrate. Over 6.7 mmol/h Isocitrate can be produced from 2-Oxoglutarate. Under anaerobic conditions, the electrons in lactate are mainly accepted by menaquinone-7 and methymenaquinone-7, so that the maximum of 4.1427 mmol/h of electrons can be transferred. The flux through pyruvate and Acetyl-CoA (over 7 mmol/h) is higher than the average flux value. The flux then from Acetyl-CoA does not enter the TCA cycle. It has been discovered that the Acetyl-CoA is catalysed to Acetylphosphate for energy conservation

Pyruvate

Oxaloacetate

Citrate

Isocitrate

2-Oxoglutarate

Succinyl-CoA

Succinate

Fumarate

L-Malate

ubiquinone8

ubiquinone8

ubiquinol8

ubiquinol8

methymenaquinone7

methymenaquinol7

Coenzyme A

Lactate-D

0/0.0362

Acetyl-CoA

7.1819/7.2423

0.0737/0.0804

0.0737/0.0804

0.0362/0.0362

-6.7071/0.0804

0.0506/0.0540

0 /0.0362

0.0617/0.0862

0.0502/0.088

0.0362/0.0627

NADH

NAD

0.0362/4.1427

Figure II‑16 FVA result of anaerobic wild type growth. The flux values are illustrated in yellow box, the minimum values in the left and maximum values in the right. The units are mmol/h.

during anaerobic conditions. The Acetylphosphate then is converted to acetate to produce more ATP (HUNT, K. A. FLYNN, J. M. et al 2010).

The FBA distributions of three simulations (both anaerobic and aerobic growth of wild types together with double knockout) are represented in Figure II-13. In the TCA cycle, the fluxes of anaerobic wild type cells are similar to the fluxes of the double knockout in that the values are remaining in same level (close to zero). The fluxes values of anaerobic wild type are greater than double knockout by a narrow margin. This was possibly the reason for the final growth rate of the anaerobic wild type (0.0363 /h) being a little bit bigger than that for the double knockout cells (0.0289 /h). In contrast, the average flux value of aerobic growth of wild type is at a high level (over 2 mmol/h on average), leading to the largest growth rate (0.0874 /h).

The maximum growth rates under different conditions and their electron productivities are shown in Figure II-17. The green triangle is the growth of the wild type with and without growth limitations. When the reaction ratio was not applied to the reactions cytochrome c and cytochrome bd, more H+ ions could be pumped outside of the membrane in the model, the growth rate of the cell thus is larger (over 0.011 /h). When the ratio was applied, the growth declined to 0.0874 /h, which is the normal level (the blue triangle in Figure II-17). The double knockout has the smallest growth rate and the maximum electron flux. The anaerobic growth of wild type shares the same growth level with the double knockout.

Figure II‑17 the growth rate verses electron fluxes. The graph illustrates the maximum electron fluxes under different growth rates. Ubq8\_cytm-c is the ubiquinone-8 related to the cytochrome c reaction; ubq8\_cytm-bd is the ubiquinone-8 related to the cytochrome bd reaction. The ratio of these two reactions was set to 4:1 to match the enzymatic limitation. The sum of the flux values of these two reactions equals the total electron flux (the top curve in the graph).

### MET mode and NADH productivity

In mediated electron transfer (MET) mode, the mediator gains electrons from NADH which turns into NAD+. The drain of NADH and increase of NAD+ molecules changes the ratio of NAD+ to NADH, disrupting the equilibrium. Therefore, more NADH is generated due to the adjustment of cellular metabolism to maintain a redox balance (Longfei, M & Wynands, V. 2013). The simulation of the *Shewanella* MET model is thus based on the detection of NADH production and consumption by normal growth demands.

The NADH reactions including both NADH production and consumption reactions were picked from the model. There are 45 reactions that can produce NADH and 19 reactions that consume NADH. Most NADH producing reactions are in amino acid metabolism (19 reactions), while the 9 reactions consuming NADH are from energy metabolism. Before the calculation of NADH productivity, the futile cycle within NADH metabolisms had to be removed according to the previously mentioned method.

Because the result of the FBA was non-unique, the simulation of NADH productivity was done using FVA. To drain the NADH to NAD+ that stimulates the electrons gained by mediators, a reaction NADH mediator (nadhm) was introduced. The reaction is relatively simple with the form: nadh[c] --> nad[c] + h[c]. It assumes that in subcellular systems, the electron in NADH is obtained by a mediator and produces NAD+ and H+ ions. Optimisation of this reaction under specific growth rates can be used to obtain the maximum, free NADH productivity.

The electron and NADH productivity under various growth rates are demonstrated in Figure II-18. The simulations were performed under growth rates from 0.01 /h to 0.08 /h with a lactate uptake rate of 4.08 mmol/h. The NADH producing reaction is the sum of the values of all selected NADH producing reactions. Overall, the total electron and NADH productivity decreases as the growth rate increases, with this trend repeating itself as shown in Figure II-18. When the growth is low (at 0.01 /h), nearly all NADH produced (just over 15 mmol/h) can be used in MET electron transfer. Only around 8 mmol/h of electron reaction, which is the ubiquinone-8 related electron transfer, is reacted. The amount of electrons provided by NADH (MET mode) is nearly doubled from the electron reactions (DET mode). As the growth rate increased, more electrons were transferred via DET mode (the curve of electron reactions in Figure II-18) and less NADH was produced. When the growth rate equalled 0.08 /h, almost all NADH was utilised to support cell growth, such as donating their electrons to DET mode and the production of biomass. It was found that very low amounts of NADH could be used in MET mode. The main electron source were the electron donors via DET mode which provided over 90% of electrons.

Figure II‑18 Electron, NADH productivity against growth rate. ‘NADH MET’ is mediated electron transport, ‘NADH producing’ is the maximum NADH productivity, ‘electron reactions (sum)’ is the electron transport chain via aerobic respiration, the ‘sum of electron and NADHm’ is the total electron productivity including NADH and respiration.

From Figure II-18, it can be seen that the electron productivity (including DET and MET mode) shows a decreasing trend with increase of the growth rate. This might be because more energy had to be consumed for the larger growth rates. For example, NADH and ATP were consumed in the biomass reaction which explains the cell growth. The high cell growth is illustrated by the high value of biomass reaction, leading to the high consumption of NADH and ATP. Therefore, less NADH can be used in either DET or MET modes of electron transfer. When the growth rate was low, the energy required for cell growth was maintained at a low level. Thus, more energy could be diverted to electron transfer. Interestingly, while the electrons transferred via MET mode decreased with the increase in growth rate, the electrons transferred by DET mode was raised at the same time. This illustrates that DET is necessary for the cell growth and the competition of electrons between MET and DET mode. When the growth rate was lower than approximately 0.035 /h, the electron productivity of MET was greater than DET; after this level the efficiency of DET was higher than MET.

Table II‑7 Terminal electron acceptors under aerobic condition. Oxygen and Hydrogen peroxide are the terminal electron acceptors under aerobic conditions in the model of Shewanella. The overall products are water. Two shuttle molecules are Ferrocytochrome c(focytcc) and ubiquinol-8 (ubq8h2). The ficytcc is Ferricytochrome c and ubq8 is ubiquinone-8. [c] means cytoplasm, the metabolite is inside the cell. [e] is exchange, the metabolite with [e] is at outside of cell.

|  |  |  |
| --- | --- | --- |
| Reactants | Products | Reactions |
| Hydrogen peroxide | Water | 2 focytcc[c] + 2 h[e] + h2o2[e] --> 2 ficytcc[c] + 2 h2o[e] |
| Oxygen | Water | 2 focytcc[c] + 4 h[c] + 0.5 o2[c] --> 2 ficytcc[c] + 2 h[e] + h2o[c] |
| 2 h[c] + 0.5 o2[c] + ubq8h2[c] --> 2 h[e] + h2o[c] + ubq8[c] |

Table II‑8: terminal electron acceptors under anaerobic condition. Without oxygen, there are 12 molecules can act as terminal electron acceptors. Two shuttles work under anaerobic conditions, menaquinol 7 and methylmenaquinol 7 (mmql7). Because menaquinol 7 and methylmenaquinone 7 act as the same role, only methylmenaquinol 7 is shown in figure as example. Mmqn7 is methylmenaquinone 7.

|  |  |  |
| --- | --- | --- |
| Reactants | Products | Reactions |
| CO3+ | CO2+ | 2 cobalt3[e] + mmql7[c] --> 2 cobalt2[e] + 2 h[e] + mmqn7[c] |
| chromate | chromium (III) hydroxide | cro4[e] + 2 h[e] + 1.5 mmql7[c] --> CrOH3[e] + h2o[e] + 1.5 mmqn7[c] |
| Dimethyl sulfoxide | Dimethyl sulfide | dmso[e] + mmql7[c] --> dms[e] + h2o[e] + mmqn7[c] |
| Fe3+ | Fe2+ | 2 fe3[e] + mmql7[c] --> 2 fe2[e] + 2 h[e] + mmqn7[c] |
| fumarate | Succinate | fum[e] + mmql7[c] --> mmqn7[c] + succ[e] |
| Manganese(IV) oxide | Mn2+ | 2 h[e] + mmql7[c] + 2 mn4o[e] --> 2 h2o[e] + mmqn7[c] + 2 mn2[e] |
| Nitrate | Nitrite | mmql7[c] + no3[e] --> h2o[e] + mmqn7[c] + no2[e] |
| Nitrite | Ammonium | 2 h[e] + 3 mmql7[c] + no2[e] --> 2 h2o[e] + 3 mmqn7[c] + nh4[e] |
| Sulfite | Hydrogen sulfide | 2 h[e] + 3 mmql7[c] + so3[e] --> 3 h2o[e] + h2s[e] + 3 mmqn7[c] |
| Thiosulfate | Sulfite | mmql7[c] + tsul[e] --> h2s[e] + mmqn7[c] + so3[e] |
| tetrathionate | Thiosulfate | mmql7[c] + tttnt[e] --> 2 h[e] + mmqn7[c] + 2 tsul[e] |
| Uranyl | Uranium dioxide | mmql7[c] + urnyl[e] --> 2 h[e] + mmqn7[c] + urdio[e] |

## Futile Cycles is plural

In a constraint-based model, the major constraints are the uptake rates of metabolites and mass balance. However, it is difficult to implement thermodynamic limitations to the reactions. Without these thermodynamic constraints, some metabolic reactions formed in an internal cycle can be computed as physical fluxes. Similar to electrical networks, these reaction cycles are analogous to Kirchoff’s current and voltage laws (Beard, D. A et al 2002).

In FBA calculations, there are upper and lower boundaries for each reaction to prevent numerical problems during the optimisation algorithm. The presence of these futile cycles leads to excessively high values for the reactions inside the cycle. The reason for this, in one case, is that the loop reactions appear in non-futile fluxes and in activated pathways simultaneously. Because of the superimposition of values from futile fluxes and activated productive pathway, the flux values of the loop reactions are high but different. This is usually the case in a currency metabolite like NADH. In other cases, the high values of flux in loop reactions do not contribute to the objective value, possibly due to the loop law of the algorithm (Figure II-19) (Longfei, M & Wynands, V. & Verwoerd, W. 2013).

In this study, a futile cycle was present in NADH and ATP metabolism. The flux vales within the futile cycles were extremely high, sometimes reaching the highest boundary, but not contributing to the objective value such as growth rate. This presents a difficulty in the calculation of NADH production rate. Therefore, reducing the flux value in futile cycles was necessary.



Figure II‑19 The futile cycle. The enzymes in this reaction are fructose-1,6-bisphosphatase (FbPase) and phosphofructokinase (PFK). Two metabolites are fructose-6-phosphate (F6P) and fructose-1,6-bisphosphate(F1,6bP). ADP is released during the PFK catalysed reaction and Pi is released when F6P is produced from F1,6bP.

To remove these futile cycles, many elaborate methods have been developed such as the energy balance equation (EBA) (Beard, D. A et al 2002), known flux directionality (Varma, A. & Palsson, B. O. 1993), predicted thermodynamic parameters (Noor, E. Even, A. B. et al 2012) as well as known thermodynamic parameters (Martino, D. Daniele et al 2012). There are also nonlinear constraints to apply to the FBA distribution to calculate the loop reactions in futile cycles. However, these coefficients and constraints are hard to obtain, making the strategies impossible to apply (Longfei, M & Wynands, V. & Verwoerd, W. 2013).

Other algorithms which can avoid futile cycles such as elementary mode analysis and extreme pathways are not applicable to this model due to the large size of the model. The elementary mode analysis and extreme pathways require all possible active pathways in models. The large number of metabolites has an extraordinary huge number of pathways, creating difficulty in the calculation for the program (Trinh, C. et al 2009). By following Longfei, M & Wynands, V. & Verwoerd, W. 2013, the futile cycle was removed alternatively by split reversible reactions into two irreversible reactions and minimising networks. This method is simple to carry out but has drawback of disrupting the variability of optimal flux distributions.

In this research, futile cycles are represented in four NADH related reactions, glutamate dehydrogenase (NAD), isoleucine dehydrogenase, L-leucine dehydrogenase and valine dehydrogenase (Table II-9). In order to prevent futile cycles, these four reversible reactions were firstly written in eight irreversible reactions. For example, the reaction glutamate dehydrogenase (NAD) in published model as written as [c] : akg + h + nadh + nh4 <==> glu-L + h2o + nad. In the present model, it was written in two forms, glutamate dehydrogenase (NAD), [c] : akg + h + nadh + nh4 --> glu-L + h2o + nad and glutamate dehydrogenase (NAD) backwards, [c] : glu-L + h2o + nad --> akg + h + nadh + nh4.

The reversibility parameters for glutamate dehydrogenase (NAD) and the backward reaction were changed from 1 to 0, meaning this reaction was changed to two irreversible reactions. The lower bounds for the two changed reactions were set to 0 instead of -1000. The same changes were performed on the remaining three reactions. In the second step, the artificial metabolites were added in four, forward NADH reactions (Table II-10). The artificial metabolites (AM) were only added in four, forward NADH reactions, while the backward reactions were maintained the same. Because the AM was only present in these four forward NADH reactions, the minimising of this AM will break the balance in the futile cycle in these NADH metabolism. The last step is adding an artificial reaction to minimize this AM. The name of the artificial reaction is fluxmin, the reaction is [c] : AM --> Arti\_C. Both AM and Arti\_C are artificial metabolites are not present in other reactions. This is a non-reversible reaction that the reversibility parameter of this reaction is 0, the lower boundary is 0 instead of -1000.

Table II‑9: Four futile cycles involving NADH reactions. Where akg: 2-Oxoglutarate, nh4: Ammonium, h2o: water, 3mop: (S)-3-Methyl-2-oxopentanoate, 4mop: 4-Methyl-2-oxopentanoate, 3mob: 3-Methyl-2-oxobutanoate, glu-L: L-Glutamate, ile-L: L-Isoleucine, leu-L: L-Leucine, val-L, L-Valine.

|  |  |
| --- | --- |
| Names | Reactions |
| glutamate dehydrogenase(NAD) | [c] : akg + h + nadh + nh4 <==> glu-L + h2o + nad |
| isoleucine dehydrogenase | [c] : 3mop + h + nadh + nh4 <==> h2o + ile-L + nad |
| L-leucine dehydrogenase | [c] : 4mop + h + nadh + nh4 <==> h2o + leu-L + nad |
| valine dehydrogenase | [c] : 3mob + h + nadh + nh4 <==> h2o + nad + val-L |

Table II‑10 Added artificial metabolites (AM) in four forward NADH reactions. The most of the part of the reactions are the same, AM were added at the production part of the reactions that to make these forward reaction producing an artificial metabolite that do not exist in the model.

|  |  |
| --- | --- |
| Names | Reactions |
| glutamate dehydrogenase(NAD) | [c] : akg + h + nadh + nh4 --> glu-L + h2o + nad + AM |
| isoleucine dehydrogenase | [c] : 3mop + h + nadh + nh4 --> h2o + ile-L + nad + AM |
| L-leucine dehydrogenase | [c] : 4mop + h + nadh + nh4 --> h2o + leu-L + nad + AM |
| valine dehydrogenase | [c] : 3mob + h + nadh + nh4 --> h2o + nad + val-L + AM |

In Longfei, M & Wynands, V. & Verwored, W. 2013’s work, the strategy to remove futile cycles is called the FATMIN algorithm. In this algorithm, the growth rate is maximised initially. The growth rate is then set to the specific maximized value, and the reaction Fluxmin which only comprises the artificial metabolite is set as objective function. Next, as an objective function, the fluxmin is minimised to obtain the minimum values of the reactions in futile cycle.

Surprisingly, in this *Shewanella* model, the high flux value in futile cycle disappeared when the each of the reversible NADH reactions were written separately into two irreversible reactions. The subsequent steps: setting growth rates to a specific value and minimising the artificial flux were not required. This may be due to some metabolites in these four reactions that were also involved in other active reactions. For example, akg (2-Oxoglutarate) is a key molecule in TCA cycle, it produces Succinyl-CoA by active with two molecules of CoA and releases one molecule of NADH. The fluxes in TCA cycle have specific values, which may act as the artificial metabolite in Longfei, M & Wynands, V. & Verwored, W. 2013’s work, reducing the flux value of futile cycle.

## The Maximum Electron Productivity by Carbon Source

Due to the difference in carbon atom numbers and structures between the carbon sources, the maximum electron productivities and growth rates varies by carbon source. For example, with the same amount of carbon source (mmol/h), the maximum growth using lactate (includes 3 carbon atoms per lactate molecule) is larger than that by using acetate (2 carbon atoms per acetate molecule). In graph section III.2.3, varying maximum growth rates are plotted against different carbon sources. Formate has the lowest maximum growth rate while the largest maximum growth rate is represented by N-Acetyl-D-glucosamine. Similar to Figure II-20, a graph that compares maximum electron productivities against maximum growth rate by carbon sources was produced (Figure II-18).

Figure II‑20 Electron potential against growth rate for varying carbon sources. The maximum electron potentials against the largest growth rates of the 9 carbon sources are illustrated. The six carbon sources with blue dots are those which growth rates have been verified by experiments, the three carbon sources with black dots are data that were calculated by FBA modelling.

The relationship of maximum electron potential to carbon source was similar to biomass yield by carbon source. Due to having the simplest structure and only one carbon atom per molecule, formate had the lowest biomass yield and electron productivity. Both biomass yield and electron productivity for succinate was greater than lactate, followed by acetate. Figure II-21 shows the electron productivities under different growth rates by three carbon sources, lactate, glutamate and N-acetyl-glucosamine.

Similar to lactate, the overall trends of electron productivities of glutamate and N-acetyl-glucosamine show decline when growth rate increased. N-acetyl-glucosamine had the largest electron productivity that is, around 65 mmol/h electrons could be produced under a growth rate of 0.01 /h. The maximum growth rate reached 0.26 /h with over 40 mmol/h of electron productivity using the carbon source N-acetyl-glucosamine. In comparison, the maximum growth rate that was reached by glutamine was less than 0.15 /h. The electron potential of glutamine was between approximately 25 mmol/h and 36 mmol/h.

The maximum electron productivities were different when the carbon sources used for growth were different, but they showed an overall trend that higher growth rate in the cells lead to fewer electrons (in carbon sources) being utilised. To produce more electrons, maintaining cell growth at a low level is possible method.

There are two ways to maintain the growth rate at a low level, by growing under anaerobic conditions and by gene knockout. In anaerobic conditions, *Shewanella* remains at a low growth rate due to anaerobic respiration, which use menaquinone-7 and Methymenaquinone7 to deliver electrons; thus the flux value in TCA cycle is reduced.

In the context of gene knockout of ubiquinone-8 reactions, the deletion of single reactions reduces some growth but not to a sufficient extent (the electron flux only increased by few levels). The double deletion of NADH dehydrogenase (ubiquinone-8 & 4 protons) and Succinate Dehydrogenase was a suitable gene knockout simulation in that the growth rate and electron productivity were close to levels seen under anaerobic growth.

Because both anaerobic growth and double knockout simulations reduce flux values in the TCA cycle, it can be inferred that the TCA cycle is important in the control of growth rate. Growth may be controlled by regulating oxygen levels in the culture medium. The highest growth level was supported by the use of ubiquinone-8; the combined use of ubiquinone and memaquinone-7 may reduce the growth level.

Unlike under anaerobic conditions, the growth controlled by gene knockout could only be maintained at a specific level rather than at a chosen level. The two single gene knockouts changed the growth level in a small amount, but the growth of double knockout was maintained in a low level. The difference between these two growth controlling modes, oxygen levels and the gene knockout, lie in the terminal electron acceptors. The former uses Fe3+, NO3-, etc as the terminal electron acceptor while the latter uses oxygen. This means the gene knockout of NADH dehydrogenase (ubiquinone-8 & 4 protons) and Succinate Dehydrogenase keeps the flux in TCA cycle in a low level in aerobic conditions.

Figure II‑21 Maximum electron productivities against growth rates by three carbon sources.

## The simulation of 13C labelled experiment

As a powerful tool, 13C isotopomer analysis is often used in intracellular fluxes mapping. A labelled 13C carbon sole source to subcellular system is distributed in primary metabolites, usually amino acids. This distribution can be measured and utilised for the benefit of understanding subcellular metabolism. With 13C labelling experimental data, the FBA can produce a reliable simulation.

The 13C labelling data for *Shewanella* was published by Yinjie, T. et al in 2007. In their research, lactate was the labelled carbon source. Oxygen levels were measured at 70% for carbon limited conditions and 0% to 10% for oxygen limited level. However, these oxygen levels could not be measured in computational simulations because it requires a specific value. Thus, the oxygen levels in the simulation were set to unlimited, and were determined by the reaction ratios.

The in vivo distribution of flux in central metabolism and the TCA cycle is demonstrated in Figure II-22. The carbon source used was lactate. When lactate was up-taken by *Shewanella* cells, it was converted to pyruvate. Over 70% of pyruvate was catalysed to acetyl-CoA before entering the TCA cycle. In a carbon-limited condition, 9.1% of pyruvate was converted to Phosphoenolpyruvate through phosphoenolpyruvate synthase. The percentage was lower under oxygen limited conditions that 2.4% of pyruvate was converted to phosphoenolpyruvate. 9.1% of carbon source entered into serine metabolism through phosphoglycerate dehydrogenase under carbon-limited conditions and the percentage for oxygen limited was 2% (Figure II-22). Based on Figure II-22, 15 ratio reactions were added on the model to simulate carbon and oxygen limited growths.

The comparison between experimental data and FBA simulations can be seen in Table II-11 and Table II-12. The more reaction ratios were applied to the model, the less growth rate was calculated. For both oxygen-limited and carbon-limited conditions, the FBA results without reaction ratios had a larger growth rate than experimental data. The calculated growth without ratios for oxygen-limited condition was similar to shake flask conditions, which the lactate consumption rate is 7.2 ± 0.7 mM/gDCW/h.

It is notable that the FBA results with 12 ratios that are more close to experimental data. For carbon limited condition, there is narrow difference between experimental data and FBA simulation. When the other three reaction ratios, related to serine metabolism, were applied, the growth rate was reduced but the oxygen uptake rate increased. Simultaneously, more acetate was excreted when these ratios were applied but CO2­ formation rate remained the same. Similarly, in oxygen-limited conditions, the adding of serine ratio reactions lead to the same result with carbon-limited conditions. The FBA simulation with 12 ratio reactions were more close to experimental data. It can be seen that considerable amounts of acetate were excreted under oxygen-limited conditions. However, the excretion of acetate could not be directly compared because of different dimensions. The dimensions used in the experiments was mM but the standard unit in FBA is mM/gDCW/h.

When 15 ratio reactions were applied, the model slightly overestimated that the growth rate was less than the expected level and more excretion of acetate was predicted. The ratios applied to serine metabolism do not improve the performance of the model. In contrast, the ratios formed only in central metabolism, e.g. TCA cycle, resulted in better performance. The growth rate in the model is controlled by respiration, the flux values in TCA cycle. The high growth rate corresponds to more fluxes through the TCA cycle because the energy produced by TCA is essential for cellular growth. Thus, the control of TCA cycle is crucial to predict cell growth in the model. Due to the limitation of enzymatic kinetics, few reaction ratios for TCA are helpful for obtaining a better prediction from FBA. However, the overuse of ratios would lead to an opposite, undesirable outcome.

There are mis-predictions in the model that no acetate was produced under FBA simulation without any reaction ratios and acetate was produced aerobically with 15 ratio reactions added. These contradict published experimental data. To investigate the problem, acetate was used as the sole carbon source and growth rates were calculated under anaerobic condition. The model showed a small growth. Therefore, the acetate can be further utilised for cell growth after being converted from lactate in the model rather than excreted out of the cell. When overmuch ratios added to the model, the acetate is forced to excreted aerobically. The mistakes might have been caused by the faulty reconstruction of acetate metabolism in the model under anaerobic conditions, or the lack of enzymatic kinetic data.

Table II‑11 Simulations of carbon limited condition

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Carbon limited | Growth rate | DO level/O2 uptake rate | Acetate excreted | Pyruvate excreted(mM) | Lactate consumption rate (mM/gDCW/h) | CO2 formation rate  (mM/gDCW/h) |
| Experiment | 0.079 | 70 % | ~0(mM) | ~0 | 4.1 ± 0.3 | 9.8 ± 1.2 |
| FBA with 12 ratios | 0.0794 | 7.67 | 0.09(mmol/h) | 0 | 4.08 | 7.66 |
| FBA with 15 ratios | 0.0619 | 7.93 | 0.2586(mmol/h) | 0 | 4.08 | 7.67 |
| FBA without ratios | 0.0890 | 8.21 | 0(mmol/h) | 0 | 4.08 | 8.512 |

Table II‑12 Simulations of oxygen limited condition

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Oxygen limited | Growth rate | DO level/O2 uptake rate | Acetate excreted | Pyruvate excreted(mM) | Lactate consumption rate (mM/gDCW/h) | CO2 formation rate  (mM/gDCW/h) |
| Experiment | 0.10 | 0 to 10 % | 17.5 ± 0.9 (mM) | ~0.1 | 7.9 ± 0.6 | 13.7 ± 1.5 |
| FBA with 12 ratios | 0.08 | 9.426 | 2.87(mmol/h) | 0 | 7.9 | 13.928 |
| FBA with 15 ratios | 0.0777 | 13.883 | 2.944(mmol/h) | 0 | 7.9 | 13.762 |
| FBA without ratios | 0.1724 | 16.482 | 0(mmol/h) | 0 | 7.9 | 16.482 |

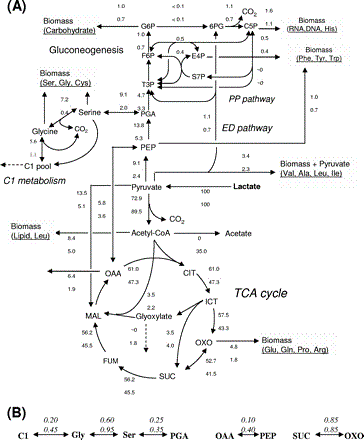


Figure II-22 fluxome of lactate metaboloism in Shewanella. (A) The in vivo flux distribution in TCA cycle and central metabolism of Shewanella is illustrated. The upper numbers is carbon limited condition, the middle numbers in brackets are shake flask growth, the oxygen limited condition is shown by lower numbers. (B) Exchange coefficients for significant reversible fluxes. Upper numbers are carbon limited and lower number are oxygen limited condition. (source from Yingjie J. Tang et al 2007)

## Comparison of electron productivity between predicted and experimental data

In order to further investigate FBA and subcellular fluxome information, proteomics data was compared with experimental data. The experiment was performed under four conditions: slow growth under aerobic and anaerobic conditions as well as fast growth under aerobic anaerobic conditions.

The growth rates for both aerobic and anaerobic slow conditions (Table II-13) were set to 0.172 /h, whose average power output is 11.196 µW. The average power output can be converted into a 0.047 mA current. At the growth rate of 0.172 /h, the calculated lactate uptake rate was 7.8814 mmol/h. Thus, this lactate uptake rate was used during the simulation of slow growth for aerobic conditions. For anaerobic conditions, due to the low carbon utilisation capacity, the calculated lactate uptake rate was 14.7232 mmol/h.

From previous knowledge, electrons are produced either via DET or MET modes. For DET mode, the electron product is through the fluxes of Cytochrome-c and Cytochrome-bd for aerobic condition and Fe reductase for anaerobic conditions. Therefore, in aerobic simulation, the value of two electron fluxes Cytochrome-c and Cytochrome-bd were observed, the sum of which was 31.7195mmol/h, equalling 0.85A. Compared with experimental data, 0.047mA, the maximum of 0.006% of electron via DET mode was converted into current. Under the context of anaerobic conditions, the flux value of Fe reductase was 72.7567 mmol/h, which equals a current of 1.9499 A. Only the maximum of 0.0025% of electron was converted into current via DET mode under anaerobic conditions.

For fast growth simulation (Table II-14), the growth rates for aerobic and anaerobic conditions were 0.667 /h, with 51.656 µW average power output, which is about 0.1 mA current. The calculated lactate uptake rates were 30.5633 mmol/h for aerobic conditions and 57.0954 mmol/h for anaerobic conditions. Under aerobic conditions, when the lactate uptake rate was set to 30.5633 mmol/h, the sum of flux values through Cytochrome-c and Cytochrome-bd was 123.0054 mmol/h (converted to current is 3.2965 A). At the lactate uptake rate of 57.0954 mmol/h, the growth rate was 0.667 /h without oxygen. The sum of flux values through two cytochrome family enzyme catalysed reactions was 282.1841 mmol/h. This values equals a current of 7.5625 A. The maximum of 0.003% of electrons via DET mode were converted into current under aerobic conditions and the maximum percentage for anaerobic condition was 0.001%. The efficiency of electron transfer for both fast and slow growth was surprisingly low via DET mode.

As a key cofactor in metabolic systems, NADH acts as a reducing agent providing electrons. In MET mode, the NADH plays an important role as an electron donor. Mediators such as pyocyanine and ACNO, collect electrons from NADH and deliver them to the anode at the outer membrane. The potential of electron productivity can be reflected by the productivity of NADH, which can be easily calculated via FBA simulations.

The free NADH produced by MR-1 under slow growth, aerobic conditions was 0.01919 mmol/h. This equals 0.515 mA. Under slow growth anaerobic conditions, 0.02269 mmol/h of free NADH was produced, a 0.608 mA current. From the maximum of 7.73% to 9.12% of electrons was turned into current under slow growth mode. Under fast growth mode, the free NADH productivities were 0.01909 mmol/h and 0.02289mmol/h for aerobic and anaerobic conditions respectively, which correspond to 0.511 mA and 0.613 mA of current. The efficiency of current conversion is between the maximum of 16.31% and 19.57%.

From the calculated results, the efficiency of current conversion of free NADH in MET mode is significantly higher than respiration in DET mode.

The huge difference between calculated and experimental current may be due to engineering design of the cultural tank. The calculated current is the theoretical electron productivity assuming all electrons are converted to current, when transferred to the anode. However, in reality, only a small amount of electrons are transferred to the anode due to the small number of cells in physical contact with the anode. Most of electrons are thus accepted by the electron acceptors in the culture medium, such as oxygen and Fe3+. Theoretically, the electrons produced under anaerobic conditions are more than that produced under aerobic conditions regardless of whether the growth rate is slow or fast. This may be explained by the use of energy. As the energy is produced from carbon sources, it can be used for either growth or current production. The larger the growth the cells have, the lower the number of electrons that can be generated. However, there might be misprediction of both electron and growth rates under anaerobic condition. Because acetate is produced using lactate as the carbon source anaerobically (Yinjie, T. et al 2007), this was not represented in the model. Therefore, more research is needed to reveal the anaerobic respiration has occurred.

Table II‑13 Comparison of current under slow growth condition

|  |  |  |  |
| --- | --- | --- | --- |
| Slow growth | Growth rate /h | Current (anaerobic) | Current (aerobic) |
| Experimental data | 0.172 | 0.047mA | 0.047mA |
| Electron reaction | 0.172 | 1.9499A | 0.85A |
| Free NADH | 0.172 | 0.608mA | 0.515mA |

Table II‑14 Comparison of current under fast growth condition

|  |  |  |  |
| --- | --- | --- | --- |
| Fast growth | Growth rate /h | Current (anaerobic) | Current (aerobic) |
| Experimental data | 0.667 | 0.1mA | 0.1mA |
| Electron reaction | 0.667 | 7.5625A | 3.2965A |
| Free NADH | 0.667 | 0.613mA | 0.511mA |

## Conclusion

In this research, a published model of S. o*neidensis* was reproduced and tested for use. The use of the model included exploring the theoretical electron productivities with two electron producing modes (MET and DET), as well as the electron productivities and growth using different carbon sources.

Firstly, the model was reproduced using results from Pinchuk et al 2010 into Excel format. The reproduced results were slightly higher than the paper. The reason for this could be the difference between initial parameters and the use of different analytical software. In direct electron transfer (DET) mode, the main electron donors are NADH, succinate and D/L-lactate under aerobic condition. Theoretically, the electron flux through cytochrome family is 13.2584 mmol/h. However, we found that if two enzymes were knocked out (NADH dehydrogenase (Ubiquinone-8 & 4 protons) and Succinate dehydrogenase), the electron productivity could be improved by over 30% to 21.8536 mmol/h, with a dramatic decrease of growth rate from 0.0874 /h to 0.0289 /h. The knockout of these three main electron donors leads to cell death. Under anaerobic conditions, NADH is the sole electron donor, and the knockout simulation of NADH dehydrogenase causes cell death. In mediated electron transfer (MET) mode, less free NADH donating their electrons can be produced as the growth rate increases.

Secondly, the electron productivities and growth by different carbon sources was then tested. Theoretically, the largest growth rates and electron fluxes can be produced by glutamate, while smallest one would be from formate. However, only the results produced by formate, glycolate and acetate were tested by experimental data. Another find was that the larger the growth rate, the lower the amount of electron fluxes that could be produced. Three carbon sources, D-lactate, L-glutamate and Nacetyl-glucosamine, were tested to that the same trend that electron productivities drop with increase of growth rate.

Finally, the model was used to compare with 13C labelled experimental data. Several ratio reactions were attached to the model to improve the prediction result. The model is in general agreement with the 13C labelled experimental data under aerobic, carbon-limited conditions with or without ratio reactions. However, under oxygen-limited conditions, the model did not predict similar results to the experimental data; especially with regards to acetate excrete predictions. Further comparison between predicted flux and proteomics data illustrates that only a small proportion (less than 0.01% in DET mode or 20% in MET mode) of the electron fluxes were delivered to the anode to produce current. It is reasoned that most electrons were accepted by the electron acceptor in the culture media.

In conclusion, an organism with fast growth and high electricity production is impossible. The energy has to be balanced between supporting growth and increasing electricity production.

# Metabolic reconstruction of marine microalgae: *Nannochloropsis gaditana*

## Nannochloropsis gaditana

### Introduction to Nannochloropsis gaditana

*Nannochloropsis gaditana* is a marine microalgae with the capability of generating NH3 as an energy resource (Abdou, R. Rekia 2012). It draws biologists’ attentions because of its high yield of lipid content and ability to live in salt water. With the reducing of fossil fuels and increasing energy demand on the earth, *Nannochloropsis gaditana* has been studied as a biological energy source for the future due its reproducible biofuel products.

In previous studies, *Nannochloropsis gaditana* shows a high affinity for carbon dioxide, which is consumed for the creation of inorganic carbon in the photosynthesis process (Huertas, I. E. and Lubián, L. M. 1998). Ammonia-nitrogen can be found in the previous study of *N. gaditana*, *N. gaditana* posseses the capability to digest or secrete ammonia (Sunda, W. G. Graneli, E. et al. 2006). Lipid products were determined in 2010 by Huerlimann, R. et al (Table III-1). Over 10 fatty acid products were found with a variety of yield distributions under three different modified L1 mediums (L-Medium, LL-Medium and S-Medium) (Huerlimann, R. et al. 2010).

Table III‑1 Lipid content of N. gaditana under L1-mudium (Huerlimann, R. et al. 2010)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fatty acid | Common name | L-Medium | LL-Medium | S-Medium |
| 14 | Myristic acid | 4.5 | 5.1 | 5.0 |
| 16 | Palmitic acid | 25.3 | 30.8 | 37.5 |
| 16:1n-7 | Palmitoleic acid | 23.4 | 21.3 | 23.3 |
| 17 | Margaric acid | 0.2 | 0.3 | 0.4 |
| 18 | Stearic acid | 0.9 | 0.8 | 0.9 |
| 18:1n-9 | Oleic acid | 4.8 | 7.5 | 11.6 |
| 18:1n-7 | Vaccenic acid | 0.4 | 0.3 | 0.3 |
| 18:2n-6 | Linoleic acid | 2.2 | 2.2 | 1.5 |
| 20:4n-6 | Arachidonic acid | 6.3 | 5.0 | 3.3 |
| 20:5n-3 | Eicosapentaenoic acid | 30.8 | 26.1 | 15.3 |

## Metabolic Reconstruction

In order to apply metabolic reconstruction, two popular databases (NCBI and KEGG) are employed. Initially, the genome information is downloaded from the NCBI. Then it is uploaded to the KAAS server for automatic annotation. Lists of organisms are chosen to analyse the information against, including all green algae in the KEGG database (Chlamydomonas reinhardtii, Volvox carteri f. nagariensis, Ostreococcus lucimarinus and Ostreococcus tauri), and other related organisms (Cyanothece).

### Procedure of KAAS

The uploaded query sequence is firstly computed against the reference sequence data within the KEGG system for searching homologs. Homologs (found in the reference gene set) above the threshold are selected for the next set calculations including BLAST score and bi-directional hit rate. The BLAST hits with the hit scores under 60 are removed. Then, the selected results are separated into KEGG ortholog (KO) groups and calculated for an assignment score. The KEGG object, in this case enzymes, in a KO group, is chosen to match the query sequence (Moriya, Y. Itoh, M. et al. 2007).



Figure III‑1: KAAS procedure (source from Moriya, Y. Itoh, M. et al. 2007)

### MetNetMaker

MetNetMaker is the software utilised in this project to assemble the metabolic reactions from the KAAS result and other references. The collected data can be saved in the SBML format, so that COBRA, which solves using the FBA method, can read the model directly. (MetNetMaker)

### Network reconstruction of Nannochloropsis gaditana

According to literature, the biological network reconstruction of *Nannochloropsis gaditana* mainly focuses on four aspects: carbon-related metabolism, photosynthesis metabolism, nitrogen metabolism and lipid metabolism.

In a recent publication, a draft genome sequence of *N. gaditana* was published (Radakovits, R. Jinkerson et al 2012). The metabolic pathway was also explored under the KEGG system (Figure III-2). Several metabolisms were found in *N. gaditana* including N-glycan biosynthesis, Carbohydrate metabolism, Nucleditide metabolism, Lipid assembly, Fatty acid synthesis, Fatty acid degradation, Photosynthesis, Terpenoid & sterol metabolism, Carbon fixation, Oxidative phosphorylation and Amino-acid metabolism. The carbon-concentrating mechanisms of *N. gaditana* were detected as well (Figure III-3) (Radakovits, R. Jinkerson et al 2012). The draft reconstruction was obtained by KAAS which contains 248 reaction maps. Over half results were trash, for example type II diabetes mellitus metabolic pathway. This is because some mammalian genes were chosen as references during the annotation process. There is a low probability to get random hits when searching against these mammalian genes.

[An external file that holds a picture, illustration, etc.
Object name is ncomms1688-f4.jpg Object name is ncomms1688-f4.jpg](http://www.ncbi.nlm.nih.gov/core/lw/2.0/html/tileshop_pmc/tileshop_pmc_inline.html?title=Click%20on%20image%20to%20zoom&p=PMC3&id=3293424_ncomms1688-f4.jpg)

Figure III‑2 Assembly of metabolic pathway in N. gaditana

The metabolic pathways existed in *N. gaditana* is represented in colour traces in comparison of the grey tracks of KEGG reference pathways. Respectively, magenta and blue lines response up or down regulated genes during nitrogen deprivation (Radakovits, R. Jinkerson et al 2012).

These pathway maps are represented in the final results with few enzymes (one or two) even these enzymatic reactions are confident inexistence in *N. gaditana*. Thereby, 23 core pathway maps were manually selected as the beginning of the curation step. In the second stage, a metabolic reconstruction of *Chlamydomonas reinhardtii* was utilised as the reference since *C. reinhardtii* is better studied than *N. gaditana* and the model of *C. reinhardtii* was published (BOYLE, N. R. & MORGAN, J. A. 2009). *C. reinhardtii* is a freshwater algae that is highly similar to *N. gaditana*. There are 484 metabolic reactions and 458 metabolites in the metabolic reconstruction of *C. reinhardtii*. Although some mechanisms such as lipid metabolisms are different between the two algae, the crucial carbon metabolism is almost the same (Huerlimann, R. et al. 2010). The missing enzymes during automatic reconstruction are added if the same enzyme is found in *C. reinhardtii*. As a result, 75 enzymes are added out of 100 of total added knowledge. The remaining 25 missing enzymatic reactions were then added with the record of a gap. This was for avoiding errors because running FBA in the final step with gaps would lead faulty in FBA methods.



Figure III‑3 Carbon-concentration mechanisms

Proposed metabolisms of inorganic carbon and C4 like metabolisms are briefly represented in Figure III-3. Black nodes such as Pyr, PEP, MA and OAA are metabolites. Red labels on arrows are enzymes. The pathways are localised in cytosol, chloroplast and mitochondria.

## Metabolic reconstruction

Metabolic pathway maps were obtained from the KAAS server. In the results, enzymes existed in *N. gaditana* as marked in green boxes. The metabolic reconstruction was started at four points, carbon metabolism, photosynthesis, lipid metabolism, and nitrogen metabolism.

Glycolysis/gluconeogenesis (Figure III-4) and the TCA cycle (Figure III-5) were selected as the initial point of topology. In glycolysis/gluconeogenesis, the metabolic flux was blocked by the missing enzymes 5.4.2.1 and 4.2.1.11. These two enzymes were found in the reconstruction of *C. reinhardtii*. Another enzyme 3.1.3.9 is included in *C. reinhardtii,* thus has been added in *N. gaditana* database. This pathway links to carbon fixation in photosynthetic organism’s metabolic pathway, pentose-phosphate pathway and TCA cycle.

The TCA cycle is one of the most common metabolisms in organisms. It provides energy for cell activities, and metabolites for many other metabolic reactions. There were six gaps in the TCA cycle of *N. gaditana*, five of them were filled by finding the same reactions in *C. reinhardtii*. The remaining one reaction is catalysed by enzyme 2.3.1.61 which was added with record of a gap because it plays an important role in TCA cycle. Acetyl-CoA is a crucial metabolite here since it is involved in four metabolisms: glycolysis/gluconeogenesis, the TCA cycle, fatty acid biosynthesis and fatty acid metabolism. The metabolic flux from glycolysis/gluconeogenesis can go into fatty acid biosynthesis and fatty acid metabolism, or join the circulation of the TCA cycle.

The result from the KAAS server does not show enough evidence of fatty acid biosynthesis (Figure III-7) in *N. gaditana*. However, the lipid products of *N. gaditana* have been illustrated (Huerlimann, R. et al. 2010). Thereby, 28 enzymes were added for producing teradecanoic acid, octanoic acid, hexadecanoic acid, hexadecenoic acid, octadecanoic acid and octadecenoic acid. There is a link to pyruvate metabolism from the metabolite malonyl-CoA, but no evidence shows that enzyme 6.4.1.2 existed in *N. gaditana*. The lipid products are valuable sources of synthetic bio-energy.

Only one enzyme, 1.1.1.35, misses in fatty acid metabolism (Figure III-8) of *N. gaditana*. Unfortunately, this enzyme could not be found in *C. reinhardtii* data as well. Therefore, 7 related reactions were included with the note of gap.

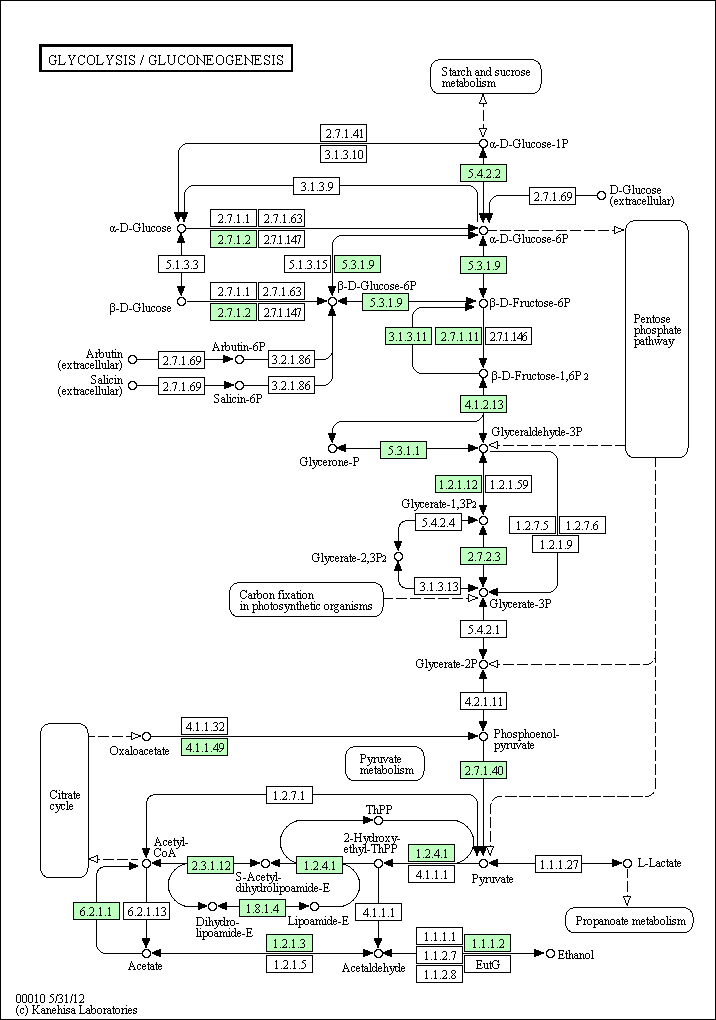


Figure III‑4 Metabolism of glycolysis/gluconeogenesis

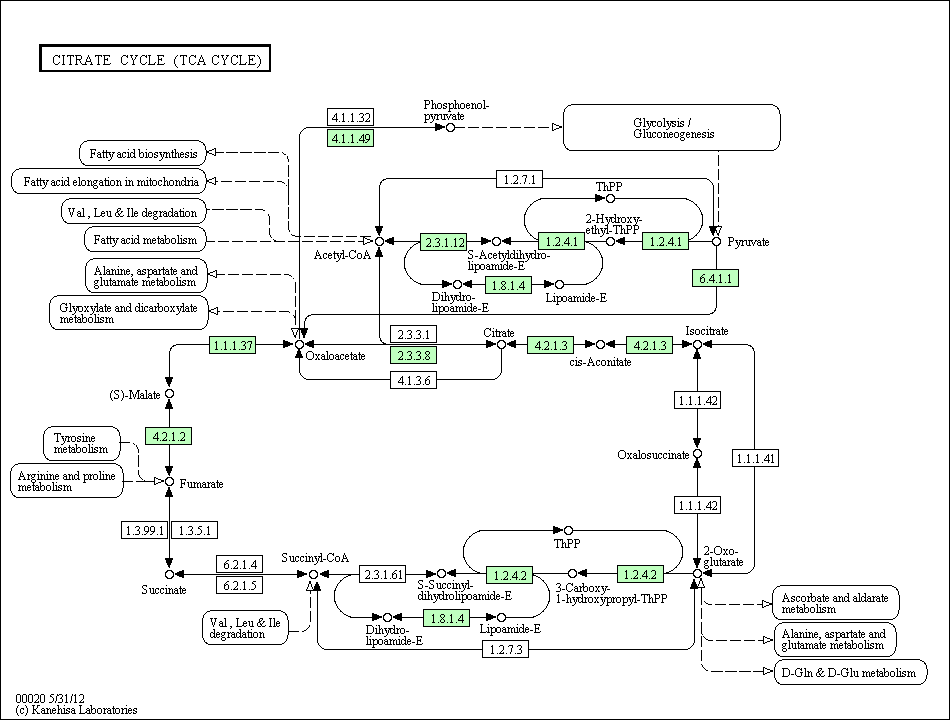


Figure III‑5 Metabolism of citrate cycle (TCA cycle)

Moving to photosynthesis (Figure III-6) in *N. gaditana*, a gap (3.5.3.1) was recorded in manual curation. Photosynthesis exists in *N. gaditana* that is, generating energy from light. A direct link form the photosynthesis pathway is the pathway of carbon fixation in photosynthetic organisms. This metabolic was found in *N. gaditana* according to the KAAS server, connected to the glycolysis/gluconeogenesis metabolic pathway.

Nitrogen is represented in the pathway of glyoxylate and dicarboxylate metabolism (Figure III-9), and nitrogen metabolic pathway (Figure IV‑9). One gap (2.3.3.1) was filled by finding the same enzyme in *C. reinhardtii*. The metabolite can thus follow the flux from the TCA cycle to NH3. This is confirmed by a pervious study that showed ammonia-nitrogen is found in *N. gaditana* (Sunda, W. G. Graneli, E. et al. 2006). Interestingly, the KAAS result demonstrates that nitrite is involved in metabolisms of *N. gaditana* (Figure IV‑9). This might be the reason for the toxicity of *N. gaditana* (Debelius, B. Forja, J. M. et al. 2009). The nitrogen metabolism was not included in the reconstruction work because literature which confirms the nitrogen metabolism in *N. gaditana*, has not been found at the present.

These pathways illustrate the living resources and the biological products of *N. gaditana*. It is possible that generating ammonia to fatty acids according to the connective flux through different pathways. However, not all metabolic fluxes go to the lipid biosynthesis pathway because some reactions are essential for cell life, for instance, amino acid and DNA synthesis. The FBA methods can be then employed to predict the efficiency of manufacturing lipid products and the influence of different environments.

334 enzymes (or reactions) were collected in the reconstruction including the reactions found in *C. reinhardtii* and gap-reactions (not found in *C. reinhardtii* but with a high probability of existence). Assembly of the enzyme list, metabolic reactions and metabolites, as well as pathway maps from KAAS can be found in the supplementary files.

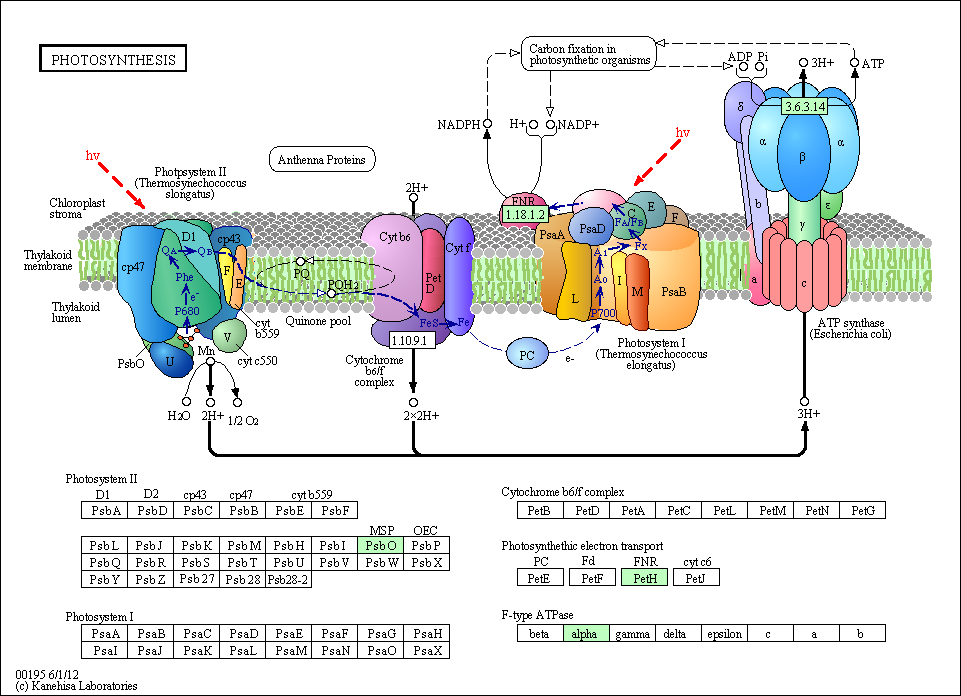


Figure III‑6 Photosynthesis

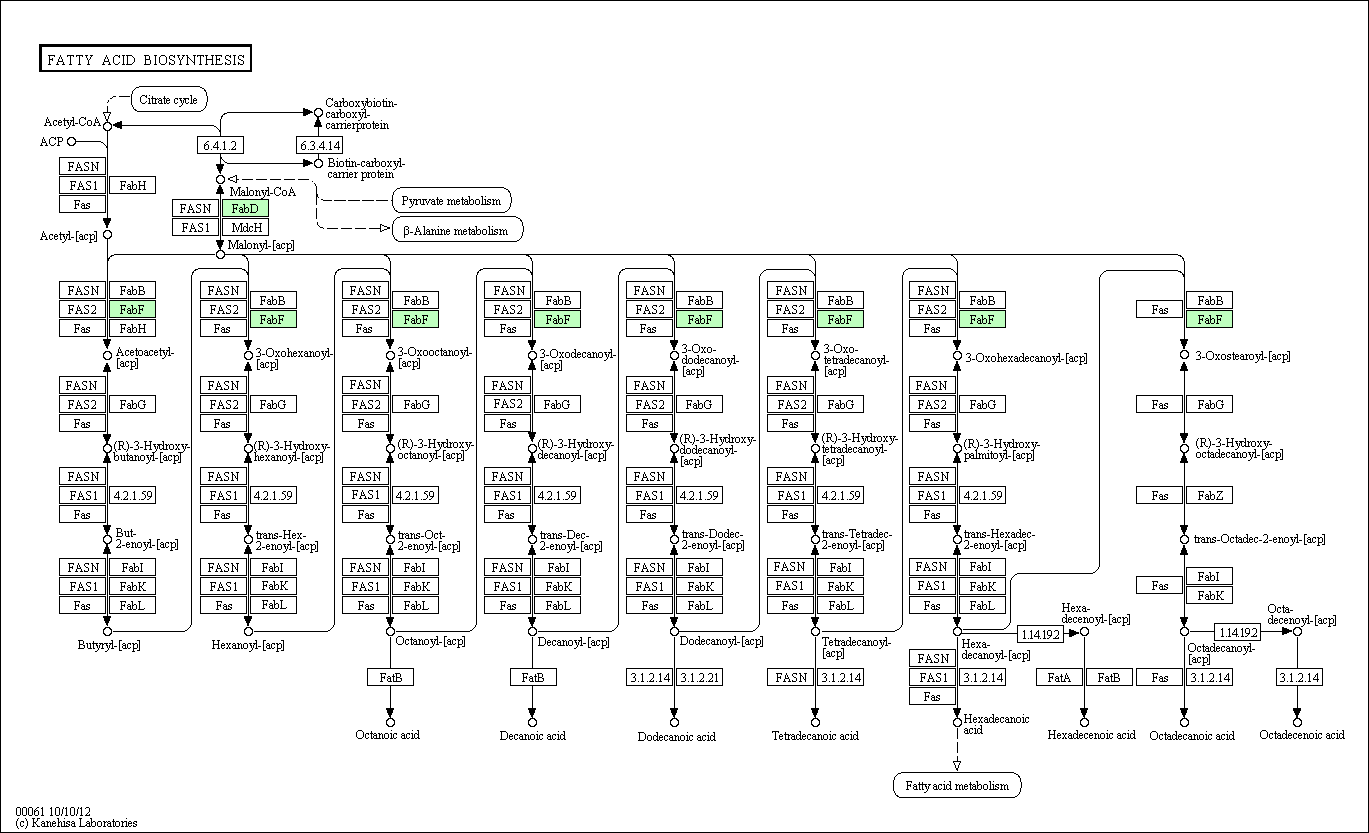


Figure III‑7 Fatty acid biosynthesis

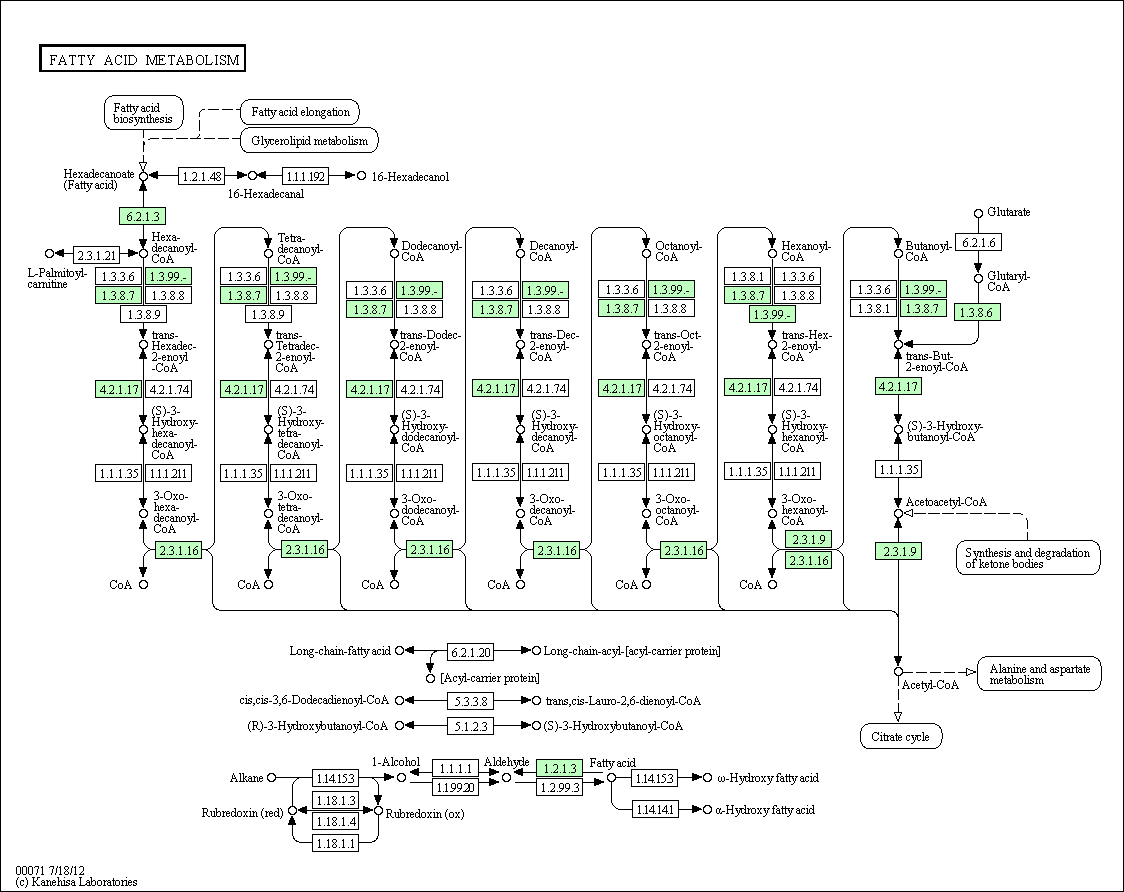


Figure III‑8 Fatty acid metabolism

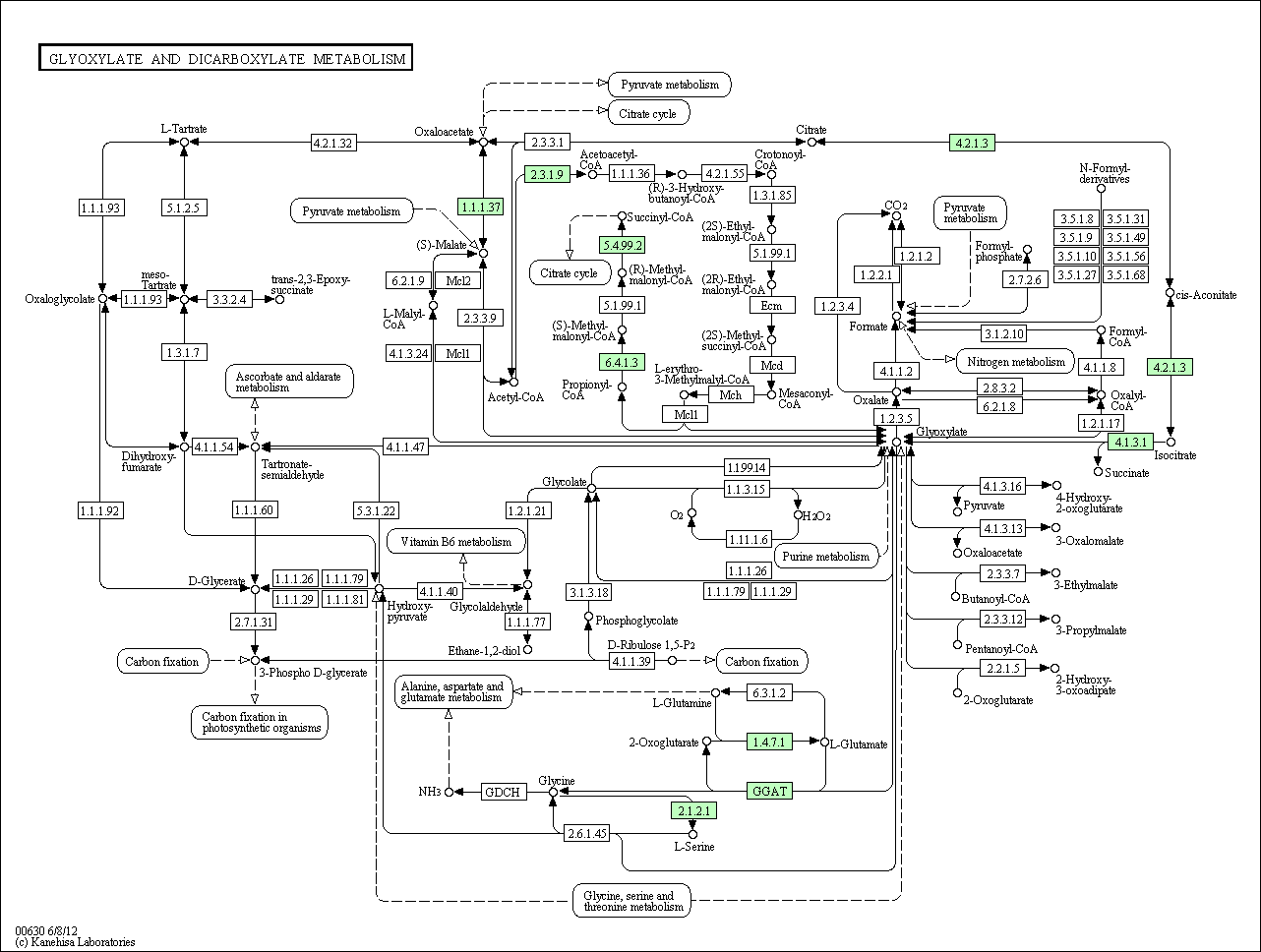


Figure III‑9 Glyoxylate and dicarboxylate metabolism

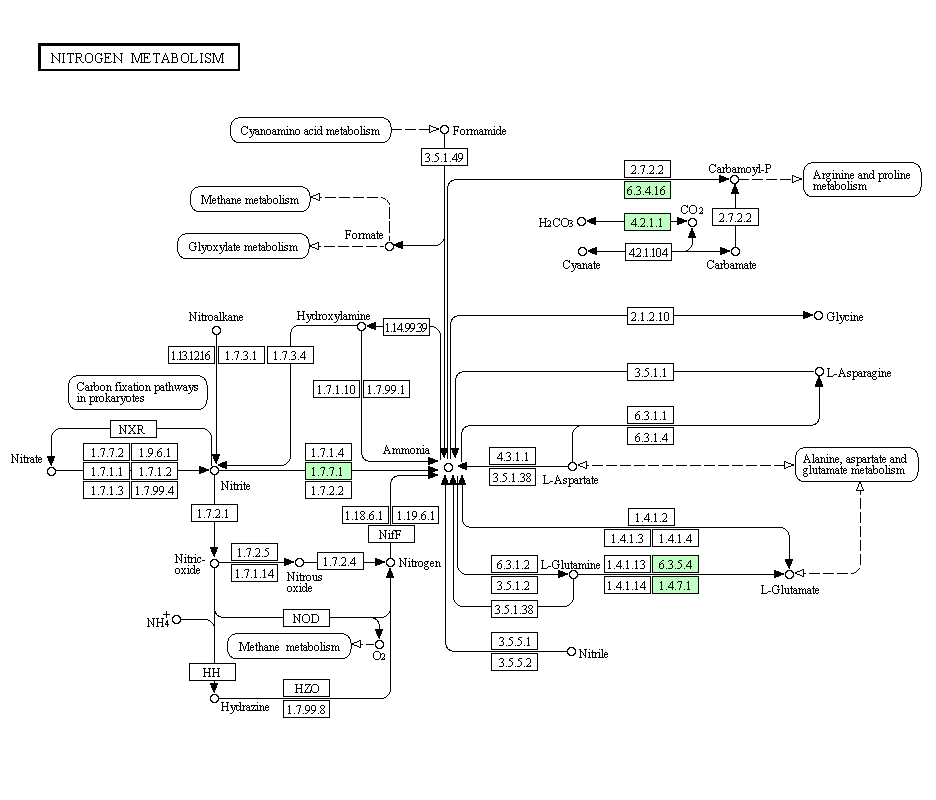


Figure III‑10 Nitrogen metabolism

## Conclusion

The metabolic reconstruction is the initial stage of building a metabolic model. This model can help us understand cellular metabolism, providing an *in silico* vision of the organism. The current reconstruction work is based on the published C. *reinhardtii* model. In their model, over 400 reactions were included. Some of them are in this reconstruction. In this work, 23 core pathways were selected out of 248 metabolic pathway maps from the KAAS server. 75 enzymes out of 100 added knowledge were added from C. *reinhardtii*. The remaining 25 enzymatic reactions were then added with the recorded as gaps. Finally, a draft model containing 334 enzymatic reactions was built.

# Conclusions and Future Work

## Conclusion

In conclusion, this work represents the use of an *in silico* model in the field of systems biology. The model can predict the phenotype of a cellular metabolism and give the general ideas to researchers.

The main methodology used to analyse the model was called flux balance analysis (FBA). With this method, the phenotype of S. o*neidensis* was analysed. It was found that the knockout of two enzymes, NADH dehydrogenase (Ubiquinone-8 & 4 protons) and Succinate dehydrogenase, resulted in an increase of electron productivity by 31% and significantly reduced growth rate simultaneously.

In another use of FBA, the growth of S. o*neidensis* under various carbon sources and oxygen requirements was predicted. The simulation illustrates that larger carbon numbers in the carbon source lead to larger growth rates and higher oxygen requirements.

In 13C labelled experiments, the model performed well under aerobic conditions. This means that the results of aerobic simulation can be used to direct wet-lab experiments in future. However, the prediction under anaerobic conditions has to be improved before use. The comparison between theoretical electron flux and the real electron current suggests that most of the electrons were not transferred to the anode. This might be improved by novel engineering culture tank design or the use of alternative media which can reduce the electrons accepted by the electron acceptor in water.

In the metabolic reconstruction project, metabolism of the algae N. *gaditana* was reconstructed. The reconstruction of this organism has not been published so far, hence the reconstruction work will act as a new model of this organism. In this new model, 234 enzymatic reactions were found in the draft reconstruction. Then, 100 reactions were picked from the similar organism C. *reinhardtii* and reference paper. The reconstruction work start from KEGG rather than the working model of C. *reinhardtii* is because N. *gaditana* uses nitrogen as energy source and has high lipid content, which is different from C. *reinhardtii*. Some redundancy reactions might be included if starting from the model of C. *reinhardtii*. This reconstruction demonstrates the cellular metabolism of N. *gaditana*, giving researchers understanding of the organism. The use of this new model is similar to the model of S. o*neidensis*. The prediction of lipid production under aerobic or anaerobic and light or dark conditions might serve as a good reference for future wet-lab experiments.

## The future plan of FBA simulation of *Shewanella oneidensis* MR-1

### The gene knockout experiment

The gene knockout experiment related to gene knockout simulation could be done in the future. It is possible to knockout NADH dehydrogenase (Ubiquinone-8 & 4 protons), succinate dehydrogenase and lactate dehydrogenase individually to observe electricity production. Then, the double knockout of succinate dehydrogenase and lactate dehydrogenase could be tested to compare electron productivity and growth with the rest of the knockout experiments as well as the anaerobic growth.

### Carbon labelling metabolic flux analysis

The useful data from the 13C tracer experiment provides a view on the distribution of metabolic fluxes. The 13C labelled nutrient (CO2 or lactate) in the medium will be digested by cells step-by-step and 13C will be distributed into different metabolic pathways. Observing the distribution of 13C helps us understand how much carbon is utilised for generating energy and how much carbon is used to produce biomass. This split ratio can be source into the FBA simulation by setting the boundaries of corresponding reactions in specific ratios. The simulation results combined with 13C tracer data are more reliable.

## The future plan of metabolic reconstruction of *Nannochloropsis gaditana*

### The biomass reaction and model prediciton

After all reactions in *N. gaditana* have been assembled, it is possible to create a biomass reaction for stimulation of the growth rate. The biomass reaction includes the precursor metabolites, synthesis of DNA, RNA and macro-molecules as well as cofactor molecules (Teusink, B. Wiersma, A. et al 2006 and EDWARDS, J. S. & PALSSON, B. O. 2000). The biomass reaction describes the growth of the organism. By setting objective functions to the biomass reaction in FBA simulation, a growth rate can be predicted.

### Improvement of FBA method: the Dynamic Flux Balance Analysis

FBA is a powerful method for cell analysis but its weakness is that it can only simulate a cell at steady -state. (Orth, J. D. Thiele, I. et al 2010) This means that a simulation of the model is the cell phenotype at a specific point in cell growth. However, it is very hard to find a motionless example in real context. Dynamic FBA methods, therefore, have been developed.

Qian et al incorporated traditional FBA and thermodynamic constraints. An energy balance analysis was incorporated and shed light on some of the incorrect predictions by FBA (Beard, D. A. Liang, S. et al 2002). In another enhancement of FBA called bi-level optimization, two objective functions were simultaneously optimised (Raman, K. & Chandra, N. 2009). An interesting integrated model of E. *coli* was established by Covert, M. W. Xiao, N. et al in 2008. This model includes metabolic, transcriptional regulatory and signal transduction processes. The simulation result significantly fitted the experimental data (Covert, M. W. Xiao, N. et al 2008). In a dynamic flux balance analysis of *S. oneidensis* MR-1 model, the Monod model was utilised to direct the FBA method. The cell growth was predicted by the Monod model and the concentration of acetate, pyruvate, lactate and biomass were simulated. Based on this simulation, over 400 mini-FBAs were completed, offering an overview of cell growth (Feng, X., et al 2008). In addition, penalty function methods were also integrated with FBA. (Qinghua, Z., Dan, W. & Momiao, X 2007)

The iFBA result is shown in solid lines and experimental data are illustrated by nodes. In A (LACxt, GLCxt and ACxt)) and B (biomass), iFBA result is almost overlapped with experimental data. In D (PtsG and LacZ) and E (EIIA), iFBA shows the similar trend to experimental data.

References

Abdou, R. Rekia (2012). Anaerobic Digestibility of Microalgae Fate and Limitations of Long Chain Fatty Acids in the Biodegradation of Lipids. Papyrus : Digital Institutional Repository Biology - Microbiology (UMI : 0410)

About the Human Genome Project: What is the Human Genome Project. The Human Genome Management Information System (HGMIS). (2011) [Online] Available from: http://www.ornl.gov/sci/techresources/Human\_Genome/project/about.shtml. [Accessed: 24 Jan 2012]

Anderson, N. L. and Anderson, N. G. (1998). Proteome and proteomics: New technologies, new concepts, and new words. ELECTROPHORESIS 19(11): 1853-1861.

Beard, D. A. Liang, S. et al. (2002). Energy Balance for Analysis of Complex Metabolic Networks. Biophysical journal 83(1): 79-86.

Boyle, N. R. and Morgan, J. A. (2009). Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. BMC Systems Biology 3: 4.

Clark, B. C. (2006). The crystal structure of tRNA. Journal of Biosciences 31(4): 453-457.

Covert, M. W. Xiao, N. et al. (2008). Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli. Bioinformatics 24(18): 2044-2050.

Crick, F. H. C. Barnett, L. et al. (1961). General Nature of the Genetic Code for Proteins. Nature 192(4809): 1227-1232.

Debelius, B. Forja, J. M. et al. (2009). Toxicity and bioaccumulation of copper and lead in five marine microalgae. Ecotoxicology and Environmental Safety 72(5): 1503-1513.

Edwards, J. S. & Palsson, B. O. (2000). The Escherichia coli MG1655 in silico metabolic genotype: Its definition, characteristics, and capabilities. Proceedings of the National Academy of Sciences 97, 5528-5533.

Feist, A. M. Herrgard, M. J. et al. (2009). Reconstruction of biochemical networks in microorganisms. Nat Rev Micro 7(2): 129-143.

Fitzgerald, L. A. et al. 2012. Shewanella oneidensis MR-1 Msh pilin proteins are involved in extracellular electron transfer in microbial fuel cells. Process Biochemistry 47, 170-174.

François Jacob and Jacques Monod (1961) Genetic regulatory mechanisms in the synthesis of proteins Journal of Molecular Biology, Volume 3, Issue 3, Pages 318-356

Geer, L. Y. Marchler-Bauer, A. et al. (2010). The NCBI BioSystems database. Nucleic Acids Research 38(suppl 1): D492-D496.

Hucka, M. Finney, A. et al. (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19(4): 524-531.

Huerlimann, R. et al. (2010). Growth, lipid content, productivity, and fatty acid composition of tropical microalgae for scale-up production. Biotechnology and Bioengineering 107(2): 245-257.

Huertas, I. E. and Lubián, L. M. (1998). Comparative study of dissolved inorganic carbon utilization and photosynthetic responses in Nannochloris (Chlorophyceae) and Nannochloropsis (Eustigmatophyceae) species. Canadian Journal of Botany 76(6): 1104-1108.

Hunt, K. A. Flynn, J. M. et al. (2010). Substrate-Level Phosphorylation Is the Primary Source of Energy Conservation during Anaerobic Respiration of Shewanella oneidensis Strain MR-1. Journal of Bacteriology 192, 3345-3351.

Joyce, A. R. and Palsson, B. O. (2006). The model organism as a system: integrating 'omics' data sets. Nat Rev Mol Cell Biol 7(3): 198-210.

Kandpal, R. Saviola, B. et al. (2009). BioTechniques Special Issue 46(5): 351-355.

Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research 28(1): 27-30.

KGML (KEGG Markup Language) (2013) [Online] Available from: http://www.genome.jp/kegg/xml/ [Accessed: 18 Jan 2013]

Kiemer, L. and Cesareni, G. (2007). Comparative interactomics: comparing apples and pears? Trends in Biotechnology 25(10): 448-454.

Lodish, H. Berk, A. et al. (1999). Molecular Cell Biology. ISBN 0-7167-3706-X: Page:405-410

Longfei, M & Wynands, V. (2013). Model-driven elucidation of the inherent capacity of Geobacter sulfurreducens for electricity generation. Journal of Biological Engineering 7, 1-21.

Martin, V. J. J. Pitera, D. J. et al. (2003). Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nat Biotech 21(7): 796-802.

Martino, D. Daniele et al. (2012). A Scalable Algorithm to Explore the Gibbs Energy Landscape of Genome-Scale Metabolic Networks. PLoS Comput Biol 8, e1002562.

Maxam, A. M. and Gilbert, W. (1977). A new method for sequencing DNA. Proceedings of the National Academy of Sciences 74(2): 560-564.

Mcmillan, D. Marritt, S. J. et al. (2012). Menaquinone-7 is specific cofactor in tetraheme quinol dehydrogenase cymA. Journal of Biological Chemistry 287, 14215-14225.

MetNetMaker A free and open-source tool for the creation of novel metabolic networks in SBML format (2010) [Online] Available from: http://www.tomforth.co.uk/metnetmaker/ [Accessed: 18 Jan 2013]

Moriya, Y. Itoh, M. et al. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Research 35(suppl 2): W182-W185.

National Human Genome Research Institute (2010). A Brief Guide to Genomics. Genome.gov. http://www.genome.gov/19016904. [Accessed: 03 Jan 2013]

Newman, M. (2003). The Structure and Function of Complex Networks. SIAM Review 45(2): 167-256.

Noor, E. Even, A. B. et al. (2012). An integrated open framework for thermodynamics of reactions that combines accuracy and coverage. Bioinformatics 28(15), 2037-2044.

Orth, J. D. and Palsson, B. Ø. (2010). Systematizing the generation of missing metabolic knowledge. Biotechnology and Bioengineering 107(3): 403-412.

Orth, J. D. Thiele, I. et al. (2010). What is flux balance analysis? Nat Biotech 28(3): 245-248.

Proulx, S. R. Promislow, D. E. L. et al. (2005). Network thinking in ecology and evolution. Trends in ecology & evolution (Personal edition) 20(6): 345-353.

Radakovits, R. Jinkerson et al (2012). Draft genome sequence and genetic transformation of the oleaginous alga Nannochloropis gaditana. Nat Commun*,* 3, 686.

Raman, K. & Chandra, N. 2009. Flux balance analysis of biological systems: applications and challenges. Briefings in Bioinformatics, 10, 435-449.

Sanger, F. Air, G. M. et al. (1977). Nucleotide sequence of bacteriophage [phi] X174 DNA. Nature 265(5596): 687-695.

Sanger, F. Nicklen, S. et al. (1977). DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences 74(12): 5463-5467.

Schaetzle, O. Barriere, F. Baronian, K. (2008). Bacteria and yeasts as catalysts in microbial fuel cells: electron transfer from micro-organisms to electrodes for green electricity. Energy & Environmental Science, 1, 607-620.

Schellenberger, J. Que, R. et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat. Protocols 6(9): 1290-1307.

Schroder, U. (2007). Anodic electron transfer mechanisms in microbial fuel cells and their energy efficiency. Physical Chemistry Chemical Physics 9, 2619-2629.

Sunda, W. G. Graneli, E. et al. (2006). Positive feedback and the development and persistence of ecosystem disruptive algal blooms1. Journal of Phycology 42(5): 963-974.

Teusink, B. Wiersma, A. et al. (2006). Analysis of Growth of Lactobacillus plantarum WCFS1 on a Complex Medium Using a Genome-scale Metabolic Model. The journal of biological chemistry 281, 40041-40048.

Thomas D. P. and Willam C. E. (2002). Cell Biology. Chapter 10: 163-174..

Transcription Factors in Plants, [Online] Available from: http://www.monsanto.com/products/pages/transcription-factors-in-plants.aspx [access date: 14/04/2014]

Trinh, C. Wlaschin, A. & Srienc, F. (2009). Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. Applied Microbiology and Biotechnology 81, 813-826.

Varma, A. & Palsson, B. O. (1993). Metabolic Capabilities of Escherichia coli: I. Synthesis of Biosynthetic Precursors and Cofactors. Journal of Theoretical Biology 165, 477-502.

Wang, Z. Gerstein, M. et al. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10(1): 57-63.

Watson, J. D. and Crick, F. H. C. (1953) Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. Nature 171, 737–738 doi:10.1038/171737a0

Xamplified Free Online Education Resource (2010) [Online] Available from: http://www.chemistrylearning.com/krebs-cycle/ [Accessed: 18 Jan 2013]

Yinjie, T. Hwang, J. S. et al. (2007). Shewanella oneidensis MR-1 Fluxome under Various Oxygen Conditions. Applied and environmental microbiology 73, 718-729.

Zhang, W. Li, F. et al. (2010). Integrating multiple ‘omics’ analysis for microbial biology: application and methodologies. Microbiology 156(2): 287-301.

Zhuwei, D. Haoran, L. & Tingyue, G. (2007). A state of the art review on microbial fuel cells: A promising technology for wastewater treatment and bioenergy. Biotechnology Advances 25, 464-482.

Supplemental data I

Modified M1 Media Composition

Chemical FW g/L Formula P/N final mM

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| M1 Media |  |  |  |  |  |
| PIPES buffer | 302.4 | 0.91 | C8H18N2O6S2 | Sigma P-1851 | 3 |
| Sodium hydroxide | 40 | 0.3 | NaOH | SigmaS-5881 | 7.5 |
| Ammonium chloride | 53.49 | 1.5 | NH4Cl | Sigma A-5666 | 28.04 |
| Potassium chloride | 74.55 | 0.1 | KCl | Sigma P-4504 | 1.34 |
| Sodium phosphate monobasic | 138 | 0.6 | NaH2PO4 H2O | Sigma S-9638 | 4.35 |
| Fumaric acid (If Anaerobic) | 160.1 | 4.8 | C4H2O4Na2 | Sigma F-1506 | 30 |
| Sodium chloride (If Aerobic) | 58.44 | 1.75 | NaCl | Fisher S271-3 | 30 |
| Mineral Supplement |  | 10 mL |  |  | see below |
| Vitamin Supplement |  | 10 mL |  |  | see below |
| Amino Supplement |  | 10 mL |  |  | see below |
| Sodium lactate (syrup)a | 112.1 | 1.1228 | C3H5O3Na | Sigma L-1375 | 6 |
| Vitamin Supplement (100X) | | | | | |
| biotin (d-biotin) | 244.3 | 0.002 | C10H16N2O3S | Sigma B-4639 | 81.87 |
| folic acid | 441.1 | 0.002 | C19H19N7O6 | Sigma F-7876 | 45.34 |
| pyridoxine HCl | 205.6 | 0.01 | C8H12ClNO3 | Sigma P-9755 | 486.38 |
| riboflavin | 376.4 | 0.005 | C17H20N4O6 | Sigma R-4500 | 132.84 |
| thiamine HCl 1.0 H2O | 355.3 | 0.005 | C18H18Cl2N4OS | Sigma T-4625 | 140.73 |
| nicotinic acid | 123.1 | 0.005 | C6H5NO2 | Sigma N-4126 | 406.17 |
| d-pantothenic acid, hemicalcium salt | 238.3 | 0.005 | C9H16NO5. 1/2Ca | Sigma P-2250 | 209.82 |
| B12 | 1355.4 | 0.0001 | C63H88CoN14O14P | Sigma V-2876 | 0.74 |
| p-aminobenzoic acid | 137.13 | 0.005 | C7H7NO2 | Sigma A-9878 | 364.62 |
| thioctic acid | 206.3 | 0.005 | C8H14O2S2 | Sigma T-5625 | 242.37 |
| Mineral Supplement (100X) |  |  |  |  |  |
| nitrilotriacetic acid(a)  (dissolve with NaOH to pH 8) | 191.1 | 1.5 | C6H9NO3 | Sigma N-9877 | 78.49 |
| magnesium sulfate heptahydrate | 246.48 | 3 | MgSO4 7H2O | Aldrich 23,039-1 | 121.71 |
| manganese sulfate monohydrate | 169.02 | 0.5 | MnSO4 H2O | Aldrich 22,128-7 | 29.58 |
| sodium chloride | 58.44 | 1 | NaCl | Sigma S-3014 | 171.12 |
| ferrous sulfate heptahydrate | 277.91 | 0.1 | FeSO4 7H2O | Sigma F-8633 | 3.6 |
| calcium chloride dihydrate | 146.99 | 0.1 | CaCl2 2H2O | Sigma C-3881 | 6.8 |
| cobalt chloride hexahydrate | 237.93 | 0.1 | CoCl2 6H2O | Sigma C-3169 | 4.2 |
| zinc chloride | 136.28 | 0.13 | ZnCl2 | Sigma Z-3500 | 9.54 |
| cupric sulfate pentahydrate | 249.68 | 0.01 | CuSO4 5H2O | Sigma C-6283 | 0.4 |
| aluminum potassium disulfate | 474.38 | 0.01 | AlK(SO4)2 12H2O | Sigma A-7167 | 0.21 |
| boric acid | 61.83 | 0.01 | H3BO3 | Sigma B-6768 | 1.62 |
| sodium molybdate dihydrate | 241.95 | 0.025 | Na2MoO4 2H2O | Aldrich 22,184-8 | 1.03 |
| nickel chloride hexahydrate | 237.6 | 0.024 | NiCl2 6H2O | Sigma N-6136 | 1.01 |
| sodium tungstate | 329.86 | 0.025 | Na2WO4 2H2O | Sigma S-0765 | 0.76 |
| Amino Acid Supplement (100X) |  |  |  |  |  |
| L-glutamic acid |  | 2 |  | Sigma G-1251 | 2 |
| L-arginine |  | 2 |  | Sigma A-3909 | 2 |
| DL-Serine |  | 2 |  | Sigma S-4375 | 2 |

a or another carbon source of choice

Supplemental data II

The model and coeffiecients used in this research (see attached file)