Storage Metabolism during Feast-Famine Cycles of Activated Sludge Wastewater Treatment and during a Sudden Nutrient Limitation Studied in *Rhodococcus jostii* RHA1 Using a Genome-Scale Metabolic Modeling Approach

Mohammad Tajparast

Department of Civil Engineering and Applied Mechanics
McGill University, Montreal

November 2014

A Thesis Submitted to McGill University in Partial Fulfillment of the Requirements of the Degree of Doctor of Philosophy

Copyright © Mohammad Tajparast, 2014
# Table of Contents

Table of Contents .................................................................................................................. i  
List of Tables ......................................................................................................................... v  
List of Figures ......................................................................................................................... vi  
Abstract ................................................................................................................................. viii  
Résumé ................................................................................................................................. xi  
Dedication ............................................................................................................................... xv  
Acknowledgments .................................................................................................................. xvi  
Preface and Contribution of Authors ....................................................................................... xviii  
Chapter 1 . Introduction ........................................................................................................ 1  
1.1 Introduction ...................................................................................................................... 1  
1.2 Motivation for the Research ............................................................................................ 3  
1.3 Genome-scale Metabolic Modeling for Environmental Conditions ............................. 4  
1.4 Development of FBA Objective Functions ..................................................................... 6  
1.5 Objectives of Thesis ......................................................................................................... 8  
1.6 Scope and Structure of Thesis ......................................................................................... 10  
1.7 References ....................................................................................................................... 12  
Chapter 2 . Literature Review ............................................................................................... 16  
2.1 Genome-scale Metabolic Network Reconstruction ......................................................... 16  
2.2 Flux Balance Analysis ..................................................................................................... 17  
2.3 Selection of Objective Functions .................................................................................... 19  
2.4 $^{13}$C-Metabolic Flux Analysis .......................................................................................... 21  
2.5 Metabolic Modeling in Environmental Engineering and Sciences .............................. 23  
2.6 Storage Compounds ........................................................................................................ 25  
2.6.1 Polyhydroxyalkanoates .............................................................................................. 26  
2.6.2 Triacylglycerols ......................................................................................................... 27  
2.6.3 Glycogen .................................................................................................................... 29  
2.7 *Rhodococcus jostii* RHA1 ............................................................................................... 31  
2.8 References ....................................................................................................................... 33  
Connecting Text: Chapter 2 and Chapter 3 .......................................................................... 38  
Chapter 3 . Genome-scale Metabolic Model of *Rhodococcus jostii* RHA1 (iMT1174) to Study the Accumulation of Storage Compounds during Nitrogen Limited Condition .................................................. 39  
Abstract ................................................................................................................................. 40
3.1 Background .................................................................................................................. 42
3.2 Materials and Methods ............................................................................................... 46
  3.2.1 Highlights of the Model Reconstruction Procedure ............................................ 46
  3.2.2 Flux Balance Analysis .......................................................................................... 47
  3.2.3 $^{13}$C-Metabolic Flux Analysis ............................................................................ 50
  3.2.4 Experimental Procedures ...................................................................................... 50
  3.2.5 Bacterial Strain and Growth Conditions ............................................................. 51
  3.2.6 Biochemical Analysis ......................................................................................... 51
  3.2.7 Analysis of Fatty Acids ....................................................................................... 52
  3.2.8 Analysis of Labelled Amino Acids ....................................................................... 53
  3.2.9 Statistical Analyses ............................................................................................... 54
3.3 Results and Discussion ............................................................................................... 55
  3.3.1 Features of the RHA1 Metabolic Network and Connectivity ............................. 55
  3.3.2 Experimentally Measured Storage Accumulation Rates ..................................... 59
  3.3.3 Metabolic Model of Storage Accumulation ......................................................... 63
  3.3.4 Sensitivity Analysis ............................................................................................... 68
  3.3.5 Application of MoMA in the Environmental Context ......................................... 70
3.4 Conclusions ................................................................................................................ 72
Additional Files .................................................................................................................. 73
Authors' Contributions ...................................................................................................... 75
Acknowledgments ............................................................................................................. 76
3.5 References .................................................................................................................. 76
Connecting Text: Chapter 3 and Chapter 4 ....................................................................... 80
Chapter 4 . Accumulation of Storage Compounds by *Rhodococcus jostii* RHA1 during Feast-Famine Cycles ................................................................. 81
Abstract .......................................................................................................................... 82
Keywords .......................................................................................................................... 82
4.1 Introduction .................................................................................................................. 83
4.2 Materials and Methods ............................................................................................... 85
  4.2.1 Bacterial Strain and Growth Medium ................................................................. 85
  4.2.2 Reactor Condition ............................................................................................... 86
  4.2.3 Analytical Procedures ....................................................................................... 87
  4.2.4 Statistical Analyses ............................................................................................ 90
4.3 Results ......................................................................................................................... 92
4.4 Discussion .................................................................................................................... 98
Acknowledgments........................................................................................................... 100
4.5 References ................................................................................................................ 101
Connecting Text: Chapter 4 and Chapter 5........................................................................ 103
Chapter 5 . Predicting the Accumulation of Storage Compounds by *Rhodococcus jostii* RHA1 in the Feast-famine Growth Cycles Using Genome-scale Flux Balance Analysis ............... 104
Abstract .......................................................................................................................... 105
Keywords ......................................................................................................................... 106
5.1 Introduction ................................................................................................................ 107
5.2 Materials and Methods ............................................................................................ 110
  5.2.1 Metabolic Network Reconstruction and Flux Balance Analysis ...................... 110
  5.2.2 ¹³C-Metabolic Flux Analysis ................................................................................. 113
  5.2.3 Bacterial Strain and Growth Medium ................................................................. 114
  5.2.4 Reactor Condition ............................................................................................... 114
  5.2.5 Analysis of Labelled Amino Acids ...................................................................... 115
  5.2.6 Statistical Analyses ............................................................................................. 116
5.3 Results and Discussion ............................................................................................. 117
  5.3.1 Central Metabolism Reaction Fluxes during Feast-Famine Cycles ................. 117
  5.3.2. FBA to Predict Storage Compounds Fluxes during Feast-Famine Cycles ...... 121
Acknowledgments ............................................................................................................. 126
Supporting Information Available .................................................................................. 127
5.4 References ................................................................................................................ 127
Appendix .......................................................................................................................... 130
Connecting Text: Chapter 5 and Chapter 6 ..................................................................... 137
Chapter 6 . Genome-Scale Metabolic Modeling to Provide Insight into the Production of Storage Compounds during Feast-Famine Cycles of Activated Sludge ......................................................... 138
Abstract .......................................................................................................................... 139
Keywords ......................................................................................................................... 140
6.1 Introduction ................................................................................................................ 141
6.2 Materials and Methods ............................................................................................ 142
  6.2.1 Genome-scale Metabolic Models and Flux Balance Analysis ...................... 142
  6.2.2 Lab-Scale Activated Sludge Reactor Fed Succinate ......................................... 144
6.3 Results and Discussion ............................................................................................. 146
  6.3.1 Genome-Scale Flux Balance Analysis .............................................................. 146
  6.3.2 Lab-Scale Activated Sludge Reactor Fed Succinate ......................................... 153
6.3.3 Comparison between Flux Balance Analysis Predictions and Experimental Observations ................................................................................................................................. 154

6.4 Conclusions .......................................................................................................................................................... 156

Acknowledgments ......................................................................................................................................................... 156

6.5 References ......................................................................................................................................................... 156

Chapter 7. Summary ..................................................................................................................................................... 159

7.1 Intellectual Contributions .................................................................................................................................... 159

7.2 Future Work .......................................................................................................................................................... 162
### List of Tables

**Table 3.1** – Macrobal elemental composition matrix in the non-limited growth condition
**Table 3.2** – Properties of the *in silico* metabolic network of *R. jostii* RHA1 (iMT1174)
**Table 3.3** – Matrix of the conversion rates on glucose as the sole carbon source in the non-limited batch growth condition
**Table 3.4** – Matrix of the conversion rates on glucose as the sole carbon source in the N-limited batch growth condition
**Table 3.5** – Matrix of the conversion rates on acetate as the sole carbon source in the non-limited batch growth condition
**Table 3.6** – Matrix of the conversion rates on acetate as the sole carbon source in the N-limited batch growth condition
**Table 3.7** – Posterior probability share of each objective function, listing in descending order of probability

**Table 4.1** – Macrobal elemental composition matrix
**Table 4.2** – Fatty acid content of *R. jostii* RHA1 at the beginning and end of the feast phase on glucose and acetate
**Table 4.3** – Matrix containing the converted amounts on glucose
**Table 4.4** – Matrix containing the converted amounts on acetate
**Table 4.5** – The estimated specific substrate uptake rates, the growth rates, the storage production rates, storage and biomass yields of *R. jostii* RHA1 during the feast-famine growth cycle

**Table 5.1** – Comparison of the goodness of fit of the non-zero central metabolism reactions of *R. jostii* RHA1 in the famine phase versus the feast phase estimated using $^{13}$C-MFA on glucose and acetate
**Table 5.2** – Comparison of the experimentally observed and simulated storage fluxes (in Cmmol/(g-DW·h)) at the optimum levels of the three objective functions examined in the feast-famine growth of *R. jostii* RHA1 on glucose and acetate
**Table 5.3** – Comparison of the $T^2$ test between the minFluxes and MoMA objective functions for glucose, acetate, and their combination in the feast-famine growth of *R. jostii* RHA1 on glucose and acetate

**Table 6.1** – Comparison between predictions by the three objective functions and experimental observations of single storage compounds accumulated during the feast-famine cycle growth on single substrates
List of Figures

Figure 2.1 – Metabolic reactions of key enzymes involved in the biosynthesis of PHB......... 27
Figure 2.2 – Metabolic reactions of key enzymes associated with the biosynthesis of triacylglycerols and their acylglycerols precursors...................................................... 29
Figure 2.3 – Metabolic reactions of key enzymes associated with the biosynthesis of glycogen 31
Figure 3.1 – Simplified possible pathways of storage metabolism in R. jostii RHA1 for the three main storage compounds: glycogen, PHA, and TAG.......................... 43
Figure 3.2 – The frequency plot of the 20 most frequent metabolites in the biochemical networks.................................................................................................................. 58
Figure 3.3 – The oxygen uptake rate time profile of RHA1 growth on acetate in the non-limited and N-limited conditions.............................................................................................. 60
Figure 3.4 – Total storage flux as a function of the non-growth-associated maintenance energy on a) glucose and b) acetate. ...................................................................................... 64
Figure 3.5 – Measured and calculated storage yields of three different storage compounds glycogen, PHA and TAG on glucose (a-f) and acetate (g-l)............................................. 66
Figure 3.6 – Calculated storage yields of four different storage compounds glycogen, PHB, PHV, and TAG accumulated by the R. jostii RHA1 model on glucose (a) and acetate (b) .......... 70
Figure 3.7 – Comparison of the fluxes of 22 central metabolic reactions of R. jostii RHA1 estimated using the FBA and 13C-MFA in N- and non-limited conditions, respectively, on a) glucose and b) acetate. ......................................................... 72
Figure 4.1 – Analysis of the dynamics of R. jostii RHA1 during a steady-state 6-hour cycle of the SBR fed glucose............................................................................................................. 93
Figure 4.2 – Analysis of the dynamics of R. jostii RHA1 during a steady-state 6-hour cycle of the SBR fed acetate........................................................................................................... 94
Figure 5.1 – Fluxes of the central metabolic reactions of R. jostii RHA1 on unlabeled glucose and 13C-bicarbonate in the feast phase (left panel) and the famine phase (right panel). .......... 118
Figure 5.2 – Fluxes of the central metabolic reactions of R. jostii RHA1 on unlabeled acetate and 13C-bicarbonate in the feast phase (left panel) and the famine phase (right panel). .......... 119
Figure 5.3 – Correlation plots of the central metabolism reactions of R. jostii RHA1 in the feast and famine phases estimated using 13C-MFA on glucose (a) and acetate (b). Note that 11 reactions were zero in both cases............................................................ 120
Figure 5.4 – Total storage flux as a function of the growth-associated maintenance energy on (a) glucose and (b) acetate. Objective functions are: MoMA, minimization of fluxes, and maximization of the growth rate. ........................................................................................................ 122
Figure 5.5 – Contour plots of the weighted average maximum growth rate of R. jostii RHA1 vs. a pair of storage fluxes (glycogen vs PHB) at glucose uptake rate of 156.01 Cmmol/(g-DW·h) (a, b, and c) and acetate uptake rate of 79.51 Cmmol/(g-DW·h) (d, e, and f) at the biosynthetic flux of TAG as 8.81 Cmmol/(g-DW·h) (a, b, and c) and 3.71 Cmmol/(g-DW·h) (d, e, and f) .......... 123
Figure 6.1 – Comparison of the overall biomass yield of R. jostii RHA1 (a and b), the sum of the metabolic fluxes, |V_{Total}| (c and d), and the total metabolic difference between the feast and famine phases, |V_{Feast} – V_{Famine}| (e and f), vs. the specific storage synthesis flux of three storage compounds (simulated one at a time): glycogen, PHB, and TAG at constant growth rate. ........ 148
Figure 6.2 – Comparison of the overall biomass yield of R. jostii RHA1 and E. coli (a and b), the sum of the metabolic fluxes, |V_{Total}| (c and d), and the total metabolic difference between the feast and famine phases, |V_{Feast} – V_{Famine}| (e and f), vs. the specific storage synthesis flux of
three storage compounds (simulated one at a time): glycogen, PHB, and TAG at constant growth rate of 0.7600 d$^{-1}$.

Figure 6.3 – Comparison of the overall biomass yield of *R. jostii* RHA1 and *E. coli* (a and b) vs. the specific storage synthesis flux of three storage compounds (simulated one at a time): glycogen, PHB, and TAG at constant growth rate of 0.7600 d$^{-1}$.

Figure 6.4 – Analysis of the dynamics of mixed liquor constituents during a steady-state 8-hour cycle of the activated sludge reactor fed succinate (a). (b) Weight percentage of glycogen and PHB in $X_{OHO+X_{U,E}}$. 
Abstract

Accumulations of carbon storage are metabolic adaptations to natural conditions where the supply of nutrients vary greatly over time, and where the relative proportions of nutrients does not always meet the growth requirement of the organisms. The current thesis is trying to build on advances in genomics to gain greater insights in the ecophysiology of this adaptation with a long term goal to be able to predict microbial interactions from genomic information. The model organism chosen is a member of the abundant soil genus *Rhodococcus*. These bacteria have been recognized as important for bioremediation because of their capacity to degrade a large number of xenobiotics, and have been found in other biotechnologies like activated sludge wastewater treatment systems. Recent work has shown that *Rhodococcus jostii* RHA1 (the model organism) growing on different substrates is capable of accumulating three types of carbon storage compounds: glycogen, polyhydroxyalkanoates (PHA), and fatty-acid compounds including triacylglycerols (TAG) or wax-esters (WE). During storage accumulation conditions (e.g., nitrogen limitation and carbon excess, or feast-famine cycles in activated sludge wastewater treatment system), the proportion of each storage compound changes with substrate, but it remains unclear how the organism regulates the synthesis of the mixture of the storage compounds. The goal of this thesis is to study this problem using a genome-scale metabolic modeling approach.

The work started by reconstructing *in silico* the genome-scale metabolic network of *Rhodococcus jostii* RHA1. The genome-scale metabolic model (named iMT1174) was then used in flux balance analysis to simulate the reaction fluxes in two experimental systems: (i) the accumulation of storage upon a sudden nitrogen limitation following non-limited growth, and (ii) cyclical feast/famine growth conditions. For the different systems, the performance of different objective functions used with flux balance analysis to predict the proportions of storage
compounds was compared. A hierarchical objective function definition was adopted with the maximization of the growth rate as the main objective function. Then, the following sub-objective functions already defined in the literature were applied and compared: (i) minimization of network fluxes, (ii) environmental minimization of metabolic adjustment (MoMA) between successive environmental conditions, (iii) minimization of ATP production rate, (iv) maximization of ATP production rate, (v) minimization of the production rate of redox potential (NADH production rate), and (vi) maximization of the storage fluxes alone or in conjunction with environmental MoMA.

Accumulation of storage upon a sudden nitrogen limitation following non-limited growth was studied by growing *R. jostii* RHA1 on glucose and acetate as sole carbon sources. The relative fluxes of storage compounds during accumulation in both cases were PHA > TAG > glycogen. Comparison of the sub-objective functions to predict these fluxes by flux balance analysis found that maximization of the storage fluxes in conjunction with environmental MoMA provided the best predictions. This conclusion was supported by the determination of the reaction fluxes in central metabolism during non-limited growth by experiments with $^{13}$C-labeled substrates and comparison with the possible variability of the same fluxes during N-limited conditions.

Cyclical feast/famine growth conditions were also studied by growing *R. jostii* RHA1 on glucose and acetate as sole carbon sources. When glucose was the substrate, the COD accumulation during the feast phase was similar for the three storage compounds: PHB, TAG, and glycogen. However, when acetate was the substrate, essentially no glycogen was accumulated and the COD yield towards PHB accumulation was 3 times higher than toward TAG accumulation. Experiments with $^{13}$C-labeled bicarbonate ($\text{HCO}_3^-$) suggested that the fluxes
through the central metabolism reactions during the feast phases were similar to the one during the famine phase. Analysis of two of the sub-objective functions showed similar results for minimization of metabolic fluxes and *environmental* MoMA, with slightly better prediction for the former sub-objective function.

What was learned through the study of *R. jostii* RHA1 was then applied to the storage metabolisms selected in mixed culture activated sludge reactors fed a single carbon source. The sub-objective functions minimization of metabolic fluxes and *environmental* MoMA between the feast and famine phases were used to predict single storage accumulations for the substrates glucose, acetate, and succinate. In addition to the *Rhodococcus jostii* RHA1 (*iMT1174*) model, the *Escherichia coli* K-12 (*iAF1260*) model was used to test the generality of the conclusions. All sub-objective functions predicted identical substrate-storage associations for the feast-famine growth of the above-mentioned metabolic models on a given substrate when glucose and acetate were set as sole carbon sources (i.e., glucose-glycogen and acetate-PHB), in agreement with experimental observations. However, the *E. coli* model predicted glycogen accumulation and the *R. jostii* model predicted PHB accumulation on succinate which was network dependent. While the accumulation of both PHB and glycogen was observed experimentally, PHB showed higher dynamics during an activated sludge feast-famine growth cycle with succinate as substrate. Nonetheless, we believe that the development of this approach will help guide the optimization of the production of storage compounds as valuable by-products of wastewater treatment.
Résumé

Les accumulations de stockage en carbone sont des adaptations métaboliques à des conditions naturelles où l'apport de nutriments varie considérablement au fil du temps, et où les proportions relatives des éléments nutritifs ne répondent pas toujours aux exigences de croissance des organismes. La thèse actuelle tente de construire sur les progrès en génomique afin d’obtenir une meilleure vue sur l'écophysiologie de cette adaptation avec un objectif à long terme pour être en mesure de prédire les interactions microbiennes basant sur l'information génomique. L'organisme modèle choisi est un membre abondant du sol du genre *Rhodococcus*. Ces bactéries ont été reconnues comme importantes pour la bioremédiation en raison de leur capacité à dégrader un grand nombre de xénobiotiques, et ont été retrouvés dans d'autres biotechnologies comme les systèmes de traitement des eaux usées à boues activées. Des travaux récents ont démontré que *Rhodococcus* RHA1 (l'organisme modèle) croissant sur des substrats différents est capable d'accumuler trois types de composés de stockage en carbone: le glycogène, les polyhydroxyalkanoates (PHA), et les composés d'acides gras dont les triacylglycérols (TAG) ou les esters de cire (WE). Dans des conditions d'accumulation de stockage (par exemple, limitation en azote et excès de carbone, ou des cycles de ‘feast-famine’ dans les systèmes d’épuration des eaux usées à boue activée, la proportion de chaque composé de stockage change avec le substrat, mais il n’est pas clair comment l'organisme régule la synthèse du mélange des composés de stockage. L'objectif de cette thèse est d'étudier ce problème en utilisant une approche de modélisation métabolique à l'échelle du génome.

Les travaux ont commencé par la reconstruction en silico du réseau métabolique à l'échelle du génome de *Rhodococcus* RHA1. Le modèle métabolique à l'échelle du génome (appelé iMT1174) a ensuite été utilisé dans l'analyse de la balance des flux afin de simuler les flux de réaction dans deux systèmes expérimentaux: (i) l'accumulation de stockage durant une
limitation soudaine en azote après une croissance non-limitée, et (ii) en conditions de croissance cycliques de ‘feast-famine’. Pour les différents systèmes, la performance de différentes fonctions objectives utilisées conjointement avec l'analyse de balance de flux afin de prédire les proportions de composés de stockage a été comparée. Une définition hiérarchique de la fonction objective a été adoptée avec la maximisation du taux de croissance comme la fonction objective principale. Ensuite, les fonctions sous-objectives suivantes, déjà définies dans la littérature, ont été appliquées et comparées: (i) la réduction des réseaux de flux, (ii) la minimisation environnementale de l'ajustement métabolique (MoMA) entre des conditions environnementales successives, (iii) la réduction du taux de production d'ATP, (iv) la maximisation de la vitesse de la production d'ATP, (c) la réduction du taux de potentiel redox (le taux de production de NADH), et (vi) la maximisation des flux de stockage uniquement ou en conjonction avec la MoMA environnementale.

L’accumulation de stockage durant une limitation soudaine en azote après une croissance non limitée a été étudiée par la croissance de R. jostii RHA1 sur le glucose et l'acétate comme seule source de carbone. Les flux relatifs des composés de stockage au cours de l'accumulation dans les deux cas étaient PHA > TAG > glycogène. Une comparaison des fonctions sous-objectives par l’analyse de la balance des flux a révélé que la maximisation des flux de stockage en conjonction avec la MoMA a fourni les meilleures prévisions. Cette conclusion a été appuyée par la détermination des flux de réaction dans le métabolisme central pendant la croissance non-limitée par des expériences avec des substrats marqués aux $^{13}$C et la comparaison avec la variabilité possible des mêmes flux dans des conditions limitées en azote.

Des conditions de croissance cycliques de ‘feast-famine’ ont également été étudiées par la croissance de R. jostii RHA1 sur le glucose et l'acétate comme seule source de carbone.
Lorsque le glucose est le substrat, l'accumulation en DCO pendant la phase de ‘feast’ était similaire pour les trois composés de stockage: le PHB, les TAG, et le glycogène. Toutefois, lorsque l'acétate était le substrat, essentiellement pas de glycogène a été accumulé et le rendement en DCO vers l'accumulation en PHB était 3 fois plus élevé que vers l'accumulation des TAG. Les expériences avec du bicarbonate marqué au $^{13}$C ($\text{HCO}_3^-$) ont suggéré que les flux à travers les réactions du métabolisme central pendant les phases de ‘feast’ étaient semblables à celle au cours de la phase de famine. Une analyse de deux des fonctions sous-objectives a démontré des résultats similaires pour la minimisation des flux métaboliques et la MoMA environnementale, avec une meilleure prédiction de la fonction sous-objective précédente.

Qu'est-ce qui a été appris à travers l'étude de R. jostii RHA1 a ensuite été appliqué aux métabolismes de stockage sélectionnés dans des réacteurs à boue activée au culture mixte alimenté avec une seule source de carbone. Les fonctions sous-objectives de minimisation des flux métaboliques et la MoMA environnementale entre les phases de ‘feast-famine’ ont été utilisés pour prédire les accumulations singulières de stockage pour les substrats de glucose, d'acétate et de succinate. En plus du modèle Rhodococcus jostii RHA1 (iMT1174), le modèle d’Escherichia coli K-12 (iAF1260) a été utilisé pour tester la généralité des conclusions. Toutes les fonctions sous-objectives ont prédit des associations identiques de substrat-stockage pour la croissance en ‘feast-famine’ des modèles métaboliques mentionnées ci-dessus sur un substrat donné lorsque le glucose et l'acétate ont été définis comme seules sources de carbone (c'est-à-dire glucose-glycogène et l'acétate-PHB), en accord avec les observations expérimentales. Cependant, le modèle d’E. coli prédit une accumulation de glycogène et le modèle de R. jostii prédit une accumulation de PHB sur le succinate, qui était dépendant du réseau. Bien que l'accumulation de PHB et de glycogène à la fois aient été observée expérimentalement, le PHB a
démontré une dynamique plus élevées pendant un cycle de croissance en ‘feast-famine’ en boue activée avec le succinate comme substrat. Néanmoins, nous pensons que le développement de cette approche aidera à guider l'optimisation de la production de composés de stockage de précieux sous-produits de traitement des eaux usées.
Dedication

I would like to dedicate my thesis to

My Beloved Wife: Marjan

and

My Late Parents: Farhad and Tahereh
Acknowledgments

First and foremost, praises and thanks to the God, the Almighty, for His showers of blessings throughout my research work to complete the research successfully.

I express my deepest gratitude to my thesis advisor Prof. Dominic Frigon for his guidance, constant encouragement, and patience during the course of dissertation. It was a great experience to work under his supervision. I really appreciate his support throughout the course of my PhD studies and I enjoyed all discussions with him which were encouraging and helped me to gain more insight on my research problem. He was a very inspiring mentor and above all a very good friend.

Within Environmental Engineering research group, I would like to thank Prof. Gehr and Prof. Liu for participation in Water Resources and Environmental Engineering seminar series from which we all get good feedback on our research at different stages. I appreciate Prof. Ghoshal’s help regarding GC analysis. I would also like to thank the technical staffs in the department specially Dr. William Cook, Daina Brumelis, John Bartczak, Marek Przykorski, and Jorge Sayat for their valuable helps throughout the course of my PhD studies. I would like to thank all the Environmental Engineering graduate students who help each other and make an enjoyable workplace on the fifth floor of the department. Special thank to Ali, Basanta, Bing, Emilie, Sai, Shameem, Sharmin, Siavash, Sourjya, Stoyan, Tala, Theresa, Usman, and Zeinab for their helps and supports.

Amino-acid measurements were performed at the Rosalind and Morris Goodman Cancer Research Centre Metabolomics Core Facility supported by the Canada Foundation of Innovation, The Dr. John R. and Clara M. Fraser Memorial Trust, the Terry Fox Foundation (TFF
Oncometabolism Team Grant 116128), and McGill University. I specially thank Dr. Daina Avizonis for letting me work in the centre to do my GC/MS analysis.

I would like to acknowledge McGill Engineering Doctoral Award (MEDA) and the Natural Sciences and Engineering Research Council of Canada (NSERC) for their financial supports.

My late parents deserve more thanks than I could ever give for their constant support, encouragement, and generosity throughout my academic studies. I also appreciate the constant support and encouragement of my family throughout the course of my PhD studies.

The last but not the least, I highly appreciate my wife Marjan’s love, support and constant encouragement during this time in doctoral school. Without her support I was not able to finish my PhD successfully.
Preface and Contribution of Authors

The contribution of the authors of the manuscripts presented in Chapters 3 to 6 are stated as follows.

M. Tajparast carried out the reconstruction of \textit{R. jostii} RHA1 (iMT1174), manually curated the metabolic network, developed various computer codes to perform the flux balance analysis and $^{13}$C-metabolic flux analysis during the feast-famine cycles and N- and non-limited conditions, the post processing of data, analyzed and validated the models in different conditions, interpreted the data, performed the experiments in different conditions, and drafted the manuscripts.

Prof. D. Frigon conceived of the study, participated in its design and coordination of the experiments, analyzed and validated the reconstruction of \textit{R. jostii} RHA1, interpreted data, and contributed to the writing of the manuscripts.

Chapter 3

was submitted to BMC Systems Biology, Genome-scale metabolic model of \textit{Rhodococcus jostii} RHA1 (iMT1174) to study the accumulation of storage compounds during nitrogen limited condition, \textbf{Mohammad Tajparast and Dominic Frigon}, November 2014.

Chapter 4

was submitted to Journal of Applied Microbiology, Accumulation of storage compounds by \textit{Rhodococcus jostii} RHA1 during feast-famine cycles, \textbf{Mohammad Tajparast and Dominic Frigon}, November 2014.

Chapter 5
To be submitted to Biotechnology and Bioengineering, Predicting the accumulation of storage compounds by *Rhodococcus jostii* RHA1 in the feast-famine growth cycles using genome-scale flux balance analysis, *Mohammad Tajparast and Dominic Frigon*, November 2014.

Chapter 6

Chapter 1. Introduction

1.1 Introduction

Biological wastewater treatment systems are often highly dynamic with regard to availability of external electron donors. In plug-flow and sequencing batch reactors, the biomass is submitted to cycles of feast and famine with the feast phase being the period of availability of external electron donors and the famine phase being the starvation period (Van Loosdrecht et al. 1997). The presence of such cycles profoundly affects the microbial ecology and the metabolism of the bacterial species present in the biomass. One of the main metabolisms found under this cycling condition is the synthesis (during feast) and consumption (during famine) by activated sludge bacteria of storage compounds such as glycogen and poly-β-hydroxybutyrate (Dawes and Senior 1973). Bacterial species accumulating storage compounds during the feast phase outperform the species lacking this ability probably by maintaining the capacity to rapidly respond to a substrate addition such that the metabolic cost of storage in terms of yield is lower than the competition cost to rapidly gain access to the substrate (Frigon et al. 2006). Therefore, understanding and properly modeling the metabolism of the storage compounds is central to predict the population dynamics in activated sludge wastewater treatment systems.

In addition to the feast-famine cycles, nutrient limited conditions such as nitrogen limitation occur in nature. Microorganisms develop metabolic strategies to cope with environments where nutrient-limitation is common such as arid sites. One of these mechanisms may be the accumulation of storage compounds that can be used by cells as endogenous carbon sources and electron donors during periods of nutritional shortage.
For activated sludge wastewater treatment systems, the prediction of population dynamics is probably most important for modeling the behaviour of bacterial species involved in solids separation problems such as foaming and bulking. The problems are caused by specific bacterial species, each associated with given operational and ecological conditions. Thus, to shed light on a specific solids separation problem, one needs to study the specific species causing it. However, only 1-15% of bacterial species present in the complex activated sludge community are culturable (Amann et al. 1995; Wagner et al. 1993). This limitation can be alleviated in part by molecular biology techniques, which offer culture-free tools to investigate the microbial ecology. For instance, we can assess the presence in the activated sludge of filamentous bacterial species that are often the cause of foaming and bulking using the 16S-rRNA-targeted technique (Levantesi et al. 2004; Wagner et al. 1994a; Wagner et al. 1994b; Wagner et al. 1994c). Although these techniques are useful to identify and monitor the abundance of specific species, the description of their metabolism remains difficult. Novel protocols combining fluorescence in situ hybridization (FISH) and metabolic labelling with radioactive carbon substrates are continuously developed (e.g., (Behrens et al. 2008)). At the same time, there is a need to provide the proper mathematical modeling framework to accurately describe the observations. Furthermore, in the current post-genomic era, genomics (sequence of the entire genetic material of a strain) and meta-genomics (genomics of entire community) will provide useful information to understand the metabolic activity of these microorganisms. Yet, the list of genes that an organism possesses (the information contained in the annotated genome sequence of an organism) is insufficient to predict its metabolic behaviour. The study aims at adapting mathematical modeling approaches developed in biochemical engineering to environmental
biotechnology and engineering in order to relate the list of genes found in a genome to the metabolic yields and rates of a bacterial cell found in different environmental conditions.

Beyond solving operational problems, studying storage metabolism during feast-famine cycles and nutrient limited conditions of activated sludge bacteria provides profound insight in terms of process optimization for the production of value added by-products (e.g., bioplastics and biodiesel) from wastewater. For example, using genome-scale metabolic modeling, one can predict storage compound(s) accumulated by bacteria on a specific type of substrate. This can further help optimize operating conditions of wastewater treatment plants in order to recover carbon and energy from wastewater.

1.2 Motivation for the Research

There is limited knowledge on the genome-scale metabolic modeling of bacteria with environmental relevance (Resendis-Antonio et al. 2007; Zhuang et al. 2011; Zhuang et al. 2012). In addition, to the best of our knowledge there is no information on the genome-scale metabolic modeling of activated sludge bacteria. One way to approach this problem is to select microorganisms involved in activated sludge wastewater treatment systems and closely study them. We selected *Rhodococcus jostii* RHA1 as our model organism of heterotrophic growth, which is an environmental isolate prevalent in the activated sludge system and was recently sequenced at the University of British Columbia (McLeod et al. 2006). This bacterium can accumulate several carbon storage compounds including poly-β-hydroxybutyrate (PHB), poly(3-hydroxyvalerate) (PHV) (Hernandez et al. 2008; Madison and Huisman 1999), triacylglycerols (TAG), wax esters (WE) (Hernandez et al. 2008; Ishige et al. 2002; Waltermann et al. 2007), α,α-trehalose, glycogen, and polyphosphate (PolyP).
Studying storage metabolism during the feast-famine growth cycle is an emerging area of research. The metabolic models developed to study storage metabolism during the feast-famine cycle are generally lumped models and there is no genome-scale metabolic model to give rise to the in-depth insight into the storage metabolism of an activated sludge bacterium (Beun et al. 2002; Dias et al. 2008; Dircks et al. 2001; Johnson et al. 2009). Therefore, there is a need of a genome-scale metabolic model to study the bacterial storage metabolism.

To provide better insight into the storage metabolism of *R. jostii* RHA1 during the feast-famine cycles and the nitrogen-limited (N-limited) and non-limited conditions, we have developed a genome-scale metabolic model of the bacterium. We also carried out $^{13}$C-labeling experiment and reconstructed a small metabolic model comprising of the central metabolic pathways of *R. jostii* RHA1 and performed $^{13}$C-metabolic flux analysis to justify the use of the MoMA sub-objective function in the environmental conditions namely feast-famine growth cycles and N-limited and non-limited batch conditions.

### 1.3 Genome-scale Metabolic Modeling for Environmental Conditions

Flux balance analysis (FBA) is a high-throughput system level mathematical framework that is independent of individual reaction kinetics data that are scarcely available. Using the annotated genome data of the species, one can list the corresponding gene product(s) (mainly protein(s)) and related biochemical reaction(s), and construct a stoichiometric matrix describing the metabolic network of the organism. This leads to a quantitative description of the cellular biochemical network called fluxomics (profile of reaction fluxes), a tool part of the high-throughput systems biology toolbox along with genomics (DNA sequence), transcriptomics (messenger RNA expression profile), proteomics (protein expression profile), and metabolomics.
(metabolite concentration profile) (Palsson 2006). FBA is a constraint-based optimization method that solves a steady-state mass balance model in which the stoichiometry of the underlying biochemical network (i.e., physicochemical constraint), along with equality and inequality constraints, bound the solution space. Typically, the mass balances around each postulated metabolite of the cell compose an underdetermined set of equations that can be solved through optimization of a biologically relevant linear objective function implemented in a linear-programming approach (Kauffman et al. 2003; Reed and Palsson 2003).

The genome-scale FBA approach has successfully been used in various applications, such as the optimization of industrial bio-processes (Kim et al. 2007; Oliveira et al. 2005), the identification of minimal media for the growth of specific microorganisms (Baart et al. 2007; Schilling et al. 2002), the discovery of biotechnologically significant microbial mutants (Puchalka et al. 2008), the identification of drug targets (Becker and Palsson 2005; Beste et al. 2007; Jamshidi and Palsson 2007), and the improvement of genome annotations (Feist et al. 2006; Gonzalez et al. 2008; Oberhardt et al. 2008; Oh et al. 2007). As opposed to the single growth phase analyzed in most of these biotechnology applications, the dynamic feast-famine growth cycle of the activated sludge treatment system was integrated to the analysis in order to investigate the cellular storage compounds during the feast and famine cycles; further, we investigated storage production by bacteria in the nitrogen-limited batch conditions that frequently occur in harsh environments such as soil. So far, our approach has been to perform a steady-state FBA for the two phases in implementing a series of objective functions. The result of this study is to propose the proper objective function(s) in order to predict storage metabolism during the feast-famine cycles and nitrogen-limited conditions.
The interest of studying the metabolic importance of various storage compounds using *Rhodococcus jostii* RHA1, which serves as our model organism of heterotrophs, is that this organism can accumulate most of the storage compounds known: PHB, PHV (Hernandez et al. 2008; Madison and Huisman 1999), TAG, WE (Hernandez et al. 2008; Ishige et al. 2002; Waltermann et al. 2007), α,α-trehalose, glycogen, and polyphosphate (PolyP). Although considerable research has provided numerous basic information on the dynamic microbial metabolism of glycogen and PHB in activated sludge and pure cultures (Beun et al. 2002; Beun et al. 2000; Carta et al. 2001; Frigon et al. 2006; Goel et al. 1998; Van Aalst-Van Leeuwen et al. 1997), little information is available on the utilization of the cellular TAG and WE pools during the feast and famine periods. In fact, the data describing the metabolic role of TAG and WE over the past decade were focused on their synthesis in potential biotechnological applications (e.g., production of oils, cosmetics, candles, printing inks, lubricants, and coating agents (Welson 2006)), or were related to their role in virulence of some pathogenic bacteria such as *Mycobacterium tuberculosis* (Kalscheuer 2010). Therefore, we aim at developing a genome-scale metabolic modeling framework to evaluate the importance of these storage compounds in the feast-famine and carbon excess context of microorganisms and to predict their utilization. This framework will foster a better description of storage phenomena found in activated sludge microbial species and held guide technological developments aiming at maximizing the production of storage compounds for their value as by-products or energy carriers.

### 1.4 Development of FBA Objective Functions

One of the approaches used in conjunction with FBA is to analyze the metabolic model using a suitable objective function. The concept of an objective originates from the assumption
that microbial metabolism results from a near-optimal expression of the genome given the evolutionary and growth history. Thus, the questions that ensue is: what is the optimality principle for a given set of environmental conditions? This question should guide us in defining potential objective functions.

For microorganisms growing in natural habitats (as opposed to axenic laboratory conditions) and competing with others, survival depends on the potential to respond to environmental changes in proper ways. For instance, the goals need to be to minimize or neutralize harmful conditions, and simultaneously to benefit from useful ones. According to the evolutionary principle, microorganisms constantly optimize their metabolic activities for survival over evolutionary time span. Moreover, if two competing species have evolved to behave exactly the same in a given environmental condition, the one harnessing the higher growth rate (or yield) will survive. Thus, “the survival of the fittest” infers that the fastest one showing the optimal responses for adaptation to the environment survives (Ponce De Leon et al. 2008). In light of this principle, we investigated the maximal cellular growth rate (or yield) as the principal objective function. It is important to notice that the maximization of the growth rate (amount of biomass produced per time and per cell weight) does not necessarily mean the maximization of the biomass yield (amount of biomass produced per substrate consumed) because rate maximization may be achieved by using lower yield metabolic pathways (e.g., complete glucose oxidation vs. fermentation) (Schuster et al. 2008). However, throughout this thesis, it was assumed that the metabolism was not wasting carbon and energy. Consequently, maximizing growth rate is equivalent to maximizing yield.

Maximization of growth rate or yield may not be sufficient to accurately predict metabolism, and a changing environment (e.g., sudden N-limitation or feast-famine conditions)
may impose other metabolic constraints on the responsiveness of microorganisms. For instance, microorganisms may tend to use optimally their enzymatic activities under certain conditions by minimizing the overall metabolic fluxes (i.e., the sum of all reactions rates). Therefore, we investigated sub-objective functions like minimizing the total metabolic fluxes as an objective function for maximum cellular efficiency. Alternatively, rapidly changing environments may force the organism to adopt a constant enzymatic profile that would provide an optimal metabolic response in more than one condition (e.g., feast vs. famine growth or non-limited vs. N-limited condition) without having to completely turnover the protein expression profile. This would maximize responsiveness and minimize energy costs due to protein turnover. Mathematically, this could be translated by minimizing the flux differences between successive environmental conditions. This function was inspired by the minimization of metabolic adjustment (MoMA) used for mutational studies.

1.5 Objectives of Thesis

We apply the genome-scale flux balance analysis of *R. jostii* RHA1 as a fluxomics tool in order to elucidate the metabolic activity of an individual species or activated sludge heterotrophs in general. Therefore, we aim to accomplish the following objectives:

I) To reconstruct *in silico* the genome-scale metabolic network of *R. jostii* RHA1.

II) To define objective functions likely capable of predicting the response of the metabolic network of heterotrophic bacteria in different growth conditions promoting the accumulation of storage compounds using flux balance analysis applied to a genome-scale metabolic model.

   a. Definition of plausible objective functions for genome-scale metabolic model simulation.
b. Determination of differences between the defined objective functions for predicting the accumulation of various storage compounds from specific substrates and in specific environmental conditions (sudden N-limitation/C-excess and feast-famine cycles).

III) To experimentally determine the suitability of defined objective functions to model observations by predictions of unobserved systems predicting the response of the metabolic network of heterotrophic bacteria in different growth conditions promoting the accumulation of storage compounds.

a. Experimental assessment of the accuracy of defined objective functions to describe the mixture of storage compounds accumulated by *R. jostii* RHA1 during nitrogen-limited substrate uptake (glucose and acetate).

b. Experimental assessment of the accuracy of defined objective functions to describe the mixture of storage compounds accumulated by *R. jostii* RHA1 during feast-famine growth cycle on glucose and acetate.

c. Experimental assessment of the accuracy of defined objective functions used in conjunction with flux balance analysis of specific genome-scale metabolic models (*R. jostii* RHA1 and *Escherichia coli*) to predict the accumulation of storage compounds by mixed activated sludge heterotrophs growing on succinate.

IV) To experimentally ascertain the fluxes through central metabolism reactions of *R. jostii* RHA1 during growth and storage accumulation phases and substantiate the choice of objective functions.

a. Measurements using $^{13}$C-lableing of the amino acid pools of the central metabolism reaction fluxes during non-limited growth period before a sudden N-limitation/C-excess period for glucose and acetate as substrate.
b. Measurements using $^{13}$C-labeling of the amino acid pools of the central metabolism reaction fluxes during feast and famine phases.

1.6 Scope and Structure of Thesis

In this work, we focused on the elucidation of the role of storage compounds during the feast and famine cycle and the N-limited condition using genome-scale flux balance analysis and $^{13}$C-metabolic flux analysis ($^{13}$C-MFA) of a model bacterial species commonly found in the environment and in activated sludge wastewater treatment systems: *Rhodococcus jostii* strain RHA1. Several studies were undertaken to address the above objectives. Genome-scale metabolic network of *R. jostii* RHA1 was *in silico* reconstructed forming the basic model used throughout this thesis. This model was used to study the storage metabolism of *R. jostii* RHA1 in the N- and non-limited conditions and the model was validated using experimentally observed phenotypic behaviour of *R. jostii* RHA1 on glucose and acetate as the sole carbon sources in the non- and N-limited conditions that were presented in Chapter 3. In this chapter, the predictions of the genome-scale flux balance analysis approach for the central metabolic reactions were also compared with those predicted by the $^{13}$C-metabolic flux analysis technique in the N- and non-limited conditions in order to justify the use of the MoMA sub-objective function to compare the N- and non-limited conditions. Chapter 4 presented the experimental data on the feast-famine growth of *R. jostii* RHA1 on glucose and acetate as the sole carbon sources. Chapter 5 displayed the predictions of storage metabolism of *R. jostii* RHA1 during the feast-famine cycles on glucose and acetate and compared the predicted storage fluxes with the experimentally observed ones; here, we justified the use of the MoMA sub-objective function by comparing the predicted metabolic fluxes of the central metabolic pathways using the $^{13}$C-metabolic flux analysis of the
central metabolic pathways of *R. jostii* RHA1 in the feast and famine phases. In addition, the genome-scale metabolic model of *R. jostii* RHA1 was utilized as the model organism of heterotrophs to study the storage metabolism of activated sludge bacteria during the feast-famine growth cycle on succinate and the model was validated using the feast-famine growth data of activated sludge on succinate which was presented in Chapter 6.

The following remarks serve as a guide in reading this manuscript based thesis.

Chapter 1 provides the introduction, motivation, objectives, scope, and structure of the thesis.

Chapter 2 provides a detailed literature review on the following subjects: genome-scale metabolic network reconstruction and flux balance analysis, selection of objective functions for genome-scale metabolic modeling, the use of metabolic modeling in the environmental context, *Rhodococcus jostii* RHA1, storage compounds, and $^{13}$C-metabolic flux analysis.

Chapters 3 to 6 are four journal manuscripts that address the specific objectives mentioned above. The manuscript detailed in Chapter 3 “Genome-scale metabolic model of *Rhodococcus jostii* RHA1 (iMT1174) to study the accumulation of storage compounds during nitrogen limited condition” has been submitted to the journal of BMC Systems Biology.

The manuscript presented in Chapter 4 “Accumulation of storage compounds by *Rhodococcus jostii* RHA1 during feast-famine cycles” has been submitted to the Journal of Applied Microbiology.
The manuscript contained in Chapter 5 “Predicting the accumulation of storage compounds by *Rhodococcus jostii* RHA1 in the feast-famine growth cycles using genome-scale flux balance analysis” will be submitted to the journal of Biotechnology and Bioengineering.

The manuscript in Chapter 6 “Genome-scale metabolic modeling to provide insight into the production of storage compounds during feast-famine cycles of activated sludge” has been published in the journal of Water Science and Technology (Tajparast and Frigon 2013).

Chapter 7 summarizes the conclusion and intellectual contributions of the research performed.

1.7 References


Chapter 2 . Literature Review

2.1 Genome-scale Metabolic Network Reconstruction

A metabolic model of *Haemophilus influenza* was the first genome-scale metabolic reconstruction to be published over a decade ago (Edwards and Palsson 1999). The field of genome-scale metabolic network reconstruction and analysis has advanced fast since then, and today more than 50 manually curated genome-scale metabolic models have been reconstructed (Oberhardt et al. 2009). Furthermore, recently developed automated model building tools have advanced the range of available models to more than 100 organisms, and have the potential to generate a draft model for any finished genome (Klitgord and Segre 2011).

A genome-scale metabolic reconstruction can be carried out in four steps: First, an initial reconstruction is made from gene-annotation data in combination with information from online databases such as Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al. 2006) and EXPASY (Gasteiger et al. 2003), which links known genes to functional categories and help connect the genotype-phenotype gap. Second, the initial reconstruction is manually curated by an investigation of the primary literature; then, the reconstruction as a knowledge base is transformed into a mathematical model that can be analyzed through constraint-based methods. Third, the reconstruction is validated with comparison of model predictions to phenotypic traits. Finally, a metabolic reconstruction is subjected to continued validation against wet-lab and *in silico* experiments, which improve accuracy and allow to examine key hypotheses (Oberhardt et al. 2009).

As the first step in reconstructing the cellular metabolic network of organism, the list of genes in its annotated genome sequence is converted to a list of associated balanced biochemical
reactions, which form the basic framework of the model. Subsequently, the model is analyzed in light of literature data to determine if some reactions may have been missing from the list of annotated genes. The reactions are added along with transport reactions to make a physiologically meaningful cellular model (Palsson 2006).

In microorganisms with environmental relevance (e.g., *Rhodococcus jostii* RHA1) although annotated, several degradative pathways of various xenobiotics are missing from KEGG (the primary reaction database used (Kanehisa et al. 2006)) and they should be added to the model to account for their catabolic abilities. Moreover, the metabolic network should be completed by the pathways of storage compounds that can also be considered and manually curated in the model including PHB, PHV (Hernandez et al. 2008; Madison and Huisman 1999), TAG, WE (Hernandez et al. 2008; Ishige et al. 2002; Waltermann et al. 2007), α,α-trehalose, glycogen, and PolyP. Finally, macromolecules are assigned to make up the biomass (e.g., protein, DNA, RNA, phospholipids, small molecules such as electron carriers and coenzymes, peptidoglycan, carbohydrates, and corynomycolic acids). This model is contained in the stoichiometric matrix (S) of the biochemical network with rows and columns corresponding to the metabolites and reactions, respectively.

2.2 Flux Balance Analysis

The gene-protein-reaction (GPR) association, along with the stoichiometric matrix, is a qualitative representation of the cellular network. This model is studied quantitatively by setting up a series of mass balances around each intracellular metabolite that can be expressed in matrix notation as:
\[
\frac{dX}{dt} = S \times v
\]  

(1)

where \( X \) is the \((m \times 1)\) vector of the concentrations of the balanced metabolites (i.e., intracellular and macromolecular compounds), \( v \) denotes the \((n \times 1)\) vector of the entire metabolic fluxes, and \( S \) stands for the \(m \times n\) stoichiometric matrix.

The model is then solved by assuming that the cellular network is at steady-state, which simplifies the previous equation to:

\[
S \times v = 0
\]  

(2)

This assumption is plausible due to the fast equilibration of intracellular metabolite pools (time-scale of seconds) compared to the time-scale of genetic regulation (minutes) (Segre et al. 2002) and due to high reaction rates of intracellular reactions compared to exchange rates such as substrate uptake, cell growth, and by-product secretion rates (Lee et al. 2006; Roels 1982; Varma and Palsson 1994). Consequently, the output of the model is a distribution of metabolic fluxes through the various chemical reactions (note that the kinetics is not modeled). Since the number of fluxes normally exceeds the number of metabolites \((n > m)\), the problem is said to be underdetermined. In this case, there is a solution space for the fluxes that can be studied by optimizing proper objective functions subject to defined constraints (mass balance and inequalities) (Becker et al. 2007).

One way to simulate metabolic behaviour using the stoichiometric network is to apply objective functions (optimization criteria). For instance, one could experimentally examine the phenotypic responses when activated sludge (or pure culture) bacteria are grown in nutrient-restricted conditions. Then, one can find the objective function(s) that properly calculate the
observed phenotypic responses using the metabolic model of the model organism under study (here, *R. jostii* RHA1). In this context, the following is the type of question that is aimed to be answered: is there an objective function that enables us to predict the type of storage compound(s) and their associated metabolic fluxes exhibited by the bacteria?

### 2.3 Selection of Objective Functions

Constraint-based modeling includes metabolic flux analysis by which flux distributions can be predicted using flux measurements in addition to network stoichiometry. The limitations in measurements led to the inclusion of additional constraints stemming from the “Darwinian theory of optimization in evolution”. Particularly, optimality principles such as maximizing biomass or given reaction fluxes at normalized input rate are broadly implemented, therefore enabling one to predict fluxes by linear programming. This is the heart of Flux Balance Analysis (FBA), by which phenotypically relevant flux distributions in metabolic networks can be predicted (Oberhardt et al. 2009; Ruppin et al. 2010).

Although various objective functions exist (Palsson 2006), the maximization of the growth rate was reported to be successful at predicting experimental observations for bacteria (Edwards et al. 2001; Ibarra et al. 2002; Segre et al. 2002). Based on the hypotheses developed earlier, we investigated the value of maximizing the cellular growth rate in order to calculate the unknown metabolic fluxes in the nitrogen non-limited and nitrogen-limited (non-limited and N-limited, respectively) conditions.

Most research works were focused on developing objective functions for pure culture such as symbiotic nitrogen fixation in *Rhizobium etli* (Resendis-Antonio et al. 2007) and
sinorhizobium meliloti (Zhao et al. 2012). The principle of flux minimization was also developed to estimate stationary fluxes in metabolic networks (Holzhutter 2004). Moreover, minimization of metabolic adjustment (MoMA) was proposed to compare the wild type and mutant strains of bacteria (Segre et al. 2002). However, the analysis of bacteria in environmental conditions, such as feast-famine growth cycles and non-limited and N-limited conditions, has not been much investigated. This is the field may need new objective functions.

Different tools have been generated to compare various objective functions. A quantitative framework termed “ObjFind” based on a bilevel optimization procedure was developed in order to test, disprove, or fine-tune the consistency of different hypothesized objective functions with experimentally determined flux distributions. In this method, the so-called coefficients of importance (CoIs) were determined computing the additive contribution of a given flux to an objective function with an optimization that describes the experimental flux data. In turn, a large CoI value infers that the experimental flux data are consistent with the hypothesis that the corresponding flux is maximized by the network, while a low value explains the opposite. This method was further applied to identify the CoIs for E. coli flux distributions under aerobic and anaerobic growth conditions (Burgard and Maranas 2003). Schuetz and coworkers systematically assessed the predictive capability of FBA by comparing all objective/constraint permutations to $^{13}$C-detected in vivo flux distributions from six growth conditions, including glucose- and ammonium-limited chemostat cultures and batch cultures with excess nutrient supply. In short, they defined the so-called predictive fidelity which is a measure of how far the predicted fluxes are from experimental measurements and ranked the objective functions accordingly (Schuetz et al. 2007). Knorr and coworkers compared five different objective functions using a Bayesian objective function discrimination technique to
identify which objective function can closely predict the phenotypic behaviour of *E. coli* growing on succinate. Using the Bayesian approach, the so-called posterior probability shares of different objective functions (models) were calculated comparing the selectively calculated fluxes against those of experimental data. The one possessing the highest posterior probability share is the best objective function among the others (Knorr et al. 2007).

2.4 $^{13}$C-Metabolic Flux Analysis

Metabolic flux analysis (MFA) has become an essential method to quantifying metabolic pathways, which is significant for profound understanding of biological systems. In experimentally based MFA, a biological system is provided with labeled substrates and resulting labeling patterns are determined to gain internal flux information. Three techniques are required to perform the $^{13}$C-MFA: (1) a steady state microorganism culture in a defined medium with $^{13}$C substrates; (2) accurate measurements of the labeling pattern of targeted metabolites such as amino acids by nuclear magnetic resonance (NMR) or mass spectrometry (MS); (3) mathematical modeling to design experiments, analyze data, and compute flux (Gomes and Simoes 2012). The labelled amino acids form the basis of the $^{13}$C-MFA, and then one can compute the metabolic fluxes somehow to minimize the differences between the calculated isotopomer distributions against the experimental counterparts.

To increase the consistency and resolution of the stoichiometric MFA (equation 2), a different set of constraints on intracellular flux partitioning can be posed by carbon-labeling experiments. There have been a number of modeling efforts that have led to the advancement of the scope of MFA, partly by formalizing how to build the resulting mathematical equations. One of the earliest such modeling efforts was the introduction of atom mapping matrices (AMM) that
determine the transfer of carbon atoms from educts to products (Zupke and Stephanopoulos 1994). This concept was next generalized in the form of isotopomer mapping matrices (IMM) (Schmidt et al. 1997). The application of IMMs enables the construction of all isotopomer mass balances of a metabolite pool using isotopomer distribution vectors (IDV) to determine the fraction of each metabolite present in a specific isotope pattern. The cumomer concept (Wiechert et al. 1999) was next introduced to first demonstrate that a unique IDV exists that satisfies any given possible flux distribution and consequently develop an IDV determination procedure by solving a set of equations. Note that the artificial word “cumomer fraction” was coined which is an abbreviation for “cumulated isotopomer fraction” meaning a certain sum of isotopomer fractions of a metabolite. For instance, the so-called 0-cumomer fraction of a metabolite is simply the sum of all its isotopomer fractions. Nevertheless, these methods are computationally expensive.

One of the latest modeling efforts is the elementary metabolite unit (EMU) concept (Antoniewicz et al. 2007). This approach investigates the atom transitions and keeps only the appropriate combinations that give rise to the experimentally measured mass distributions. This framework decreases the number of variables significantly, which leads to a reduction in the computations required for flux analysis, confidence intervals and degree of resolution of each flux (Suthers et al. 2010). OpenFlux is a software tool that has recently been developed to ease $^{13}$C-MFA of large-scale metabolic models based on the EMU concept (Quek et al. 2009). In this thesis, we employed this approach to perform the $^{13}$C-MFA of the central metabolism of *R. jostii* RHA1.
2.5 Metabolic Modeling in Environmental Engineering and Sciences

There are well-documented research activities being performed in the field of metabolic modeling; however, the application of metabolic modeling to address environmental issues is limited. This thesis contributes to the existing knowledge by investigating the storage metabolism of activated sludge bacteria using the genome-scale metabolic modeling and flux balance analysis of *R. jostii* RHA1, the model organism of heterotrophs, during the feast-famine growth cycles and non- and N-limited conditions. We further examined different objective functions to better predict storage metabolism in these conditions.

In this context lumped models were introduced to explain the glycogen and poly-β-hydroxybutyrate (PHB) metabolism in mixed cultures. The metabolism of glycogen storage (in the feast period) and consumption (in the famine phase) in mixed cultures under aerobic conditions was described by a small metabolic model that included six internal reactions. The authors defined two stoichiometric variables namely the efficiency of oxidative phosphorylation (δ) and maintenance energy requirement in units of adenosine triphosphate (mATP) which were estimated as 1.8 mol ATP/mol NADH₂ and 0.017 mol ATP/(Cmol·h) based on the average conversion in the feast and famine phases (Dircks et al. 2001). Another metabolic model was developed by the same research group to study production and consumption of PHB during the feast and famine phases, respectively (Beun et al. 2002). In this study, seven internal reactions were considered and values of δ and mATP were defined as 2 mol ATP/mol NADH₂ and 0.02 mol ATP/(Cmol·h), respectively.

Beyond lumped models, the genome-scale metabolic model of *Rhizobium etli* was reconstructed to study symbiotic nitrogen fixation. The reconstructed metabolic network of *R. etli* comprises of 387 reactions, 371 metabolites, and 363 genes. *In silico* analysis was conducted
for deletions in PHB synthase, glycogen synthase, arginine deiminase, myo-inositol dehydrogenase, and pyruvate carboxylase. Simulations for deletions of PHB synthase predicted that symbiotic nitrogen fixation increases, in agreement with the experimental observations in *R. etli*. A similar effect on symbiotic nitrogen fixation was predicted for the glycogen synthase deletion, which also agreed with the observed physiological response reported for *Rhizobium tropici* (Resendis-Antonio et al. 2007).

In the context of polyphosphate and glycogen-accumulating organisms, a metabolic model with thirteen reactions and six intracellular metabolites were built to study the production of polyhydroxyalkanoate copolymers in mixed microbial cultures, using mixtures of acetic and propionic acid as carbon sources (Dias et al. 2008). The efficiency of oxidative phosphorylation was estimated as 2.94 and 0.94 mol ATP/mol NADH$_2$ for cultures enriched on acetate and propionate, respectively. The authors concluded that in PHA producing systems where the microbial culture enrichment and PHA production phases are separated, it is desirable to enrich the microbial culture using acetate, while feeding a combination of acetate and propionate in the PHA production phase. Feeding scenarios were also optimized by FBA targeting maximal productivity of a PHA copolymer with a desired monomeric composition (Dias et al. 2008).

Genome-scale models of *Geobacter sulfurreducens* and *Rhodoferax ferrireducens* were used to evaluate how *Geobacter* and *Rhodoferax* species compete under different conditions found in a uranium-contaminated aquifer (Zhuang et al. 2011; Zhuang et al. 2012). The authors established a dynamic multi-species metabolic modeling framework to study the interaction of *Geobacter* and *Rhodoferax* species. The model predicted that at the low rates of acetate flux expected under natural conditions, *Rhodoferax* will outcompete *Geobacter* as long as sufficient ammonium is available. The model also predicted that when high concentrations of acetate are
added during in situ bioremediation, *Geobacter* species would predominate, which agreed with field-scale observations. Furthermore, the model predicted that under nitrogen fixation, higher carbon and electron fluxes would be channelled toward respiration rather than biomass formation in *Geobacter*, providing a potential explanation for enhanced in situ U(VI) reduction in low-ammonium zones (Zhuang et al. 2011).

According to the above-mentioned research activities, studying storage metabolism of mixed culture (i.e., activated sludge bacteria) using genomics data is a crucial step towards understanding value-added by-products from wastewater, as well as the microbial dynamics of activated sludge.

### 2.6 Storage Compounds

Studying microbial storage pools are crucial to the environmental community, since one can better understand microbial communities within activated sludge wastewater treatment processes, as well as carbon and energy recovery from wastewater. Further, genome-scale metabolic modeling is an emerging topic in environmental engineering and sciences; that can be applied to study storage metabolism in details. In this section, we describe the common storage compounds including polyhydroxyalkanoates, triacylglycerols, and glycogen that are accumulated by most of activated sludge microorganisms in different environmental conditions (e.g., feast-famine cycles and N-limited conditions).
2.6.1 Polyhydroxyalkanoates

Storage compounds are formed in microorganisms as the energy and carbon reserves under unbalanced growth conditions such as nutrient limitation (Dawes and Senior 1973). Lipophilic storage compounds are often present in many eukaryotic and prokaryotic organisms. Bacteria commonly store specialized lipids such as polyhydroxyalkanoates (PHA) as intracellular inclusions. PHA are thermoplastic and elastomeric polyesters that are readily biodegradable and can be utilized for numerous technical applications such as biofuel, polymeric material, medical implant material, and drug delivery carrier (Alvarez et al. 2000; Chen 2009; Madison and Huisman 1999; Wu et al. 2009). There are currently over 100 different kinds of known basic building blocks for PHA polymers reported (Noda et al. 2005). PHA can be divided into three categories according to the number of carbon atoms in the monomer units incorporated into the polymers. PHA containing up to five carbon atom (C5) monomers are categorized as short-chain-length PHA (scl-PHA). PHA possessing C6-C14 and > C14 monomers are classified as medium-chain-length (mcl-PHA) and long-chain-length (lcl-PHA) PHA, respectively. The scl-PHA possess properties similar to conventional plastics such as polyethylene or polypropylene, whereas the mcl-PHA are close to elastomers and rubbers (Wu et al. 2009).

The first and the most studied PHA is the homopolymer poly-β-hydroxybutyrate (PHB). In their metabolism, bacteria generate acetyl-coenzyme-A (acetyl-CoA), which is transformed into PHB by three enzymatic steps (Figure 2.1). First, 3-ketothiolase (PhaA) combines two molecules of acetyl-CoA to produce acetoacetyl-CoA. Second, Acetoacetyl-CoA reductase (PhaB) catalyzes the reduction of acetoacetyl-CoA by NADH to 3-hydroxybutyryl-CoA. Third, PHB synthase (PhaC) polymerizes 3-hydroxybutyryl-CoA to PHB and coenzyme-A is liberated. It should be noted that only (R)-isomers are accepted as substrates for the polymerizing enzyme (Verlinden et al. 2007).
Figure 2.1 – Metabolic reactions of key enzymes involved in the biosynthesis of PHB

2.6.2 Triacylglycerols

Triacylglycerols (TAG) are non-polar, water-insoluble triesters of glycerol with fatty acids. TAG are good energy reserves since they are more reduced and have a higher calorific value than glycogen and PHB. Consequently, they produce considerably more energy upon oxidation (Alvarez et al. 2002). TAG occurred in bacteria as insoluble inclusions surrounded with a thin membrane as was reported for R. opacus PD630 (Alvarez et al. 1996). TAG in cells of Rhodococcus opacus PD630 account for up to 76 or 87% of the cellular dry mass in gluconate- or olive oil-grown cells, respectively (Alvarez et al. 1996). This species is termed as oleaginous bacteria because it produces more than 20% of its biomass as lipids.
Beside \textit{R. opacus} only few unusual bacterial taxa are able to accumulate TAG, and they are more common in eukaryotic organisms. The bacteria generally capable of TAG accumulation are the Gram-negative genera \textit{Acinetobacter} and \textit{Pseudomonas}, and the actinobacterial genera \textit{Dietzia}, \textit{Gordonia}, \textit{Mycobacterium}, \textit{Nocardia}, \textit{Rhodococcus}, and \textit{Streptomyces} (Alvarez et al. 2002; Alvarez et al. 1997).

Biosynthesis of TAG can be divided into three steps: First, production of fatty acyl-compounds; second, formation of glycerol intermediates; and third, sequential esterification of glycerol moiety with fatty acyl-residues. Because fatty acid and glycerol biosynthesis by bacteria have been detailed extensively (Alvarez et al. 2002), these steps are not described more here. The enzymes involved in the esterification of the glycerol moiety probably act via sequential acylation of the \textit{sn}-1,2 and 3 positions of glycerol-3-phosphate, with the liberation of the phosphate group happening before the final acylation step (Figure 2.2). The enzymatic activity of diacylglycerol acyltransferase (DGAT) is depended on the cellular growth phase. That is, DGAT activity is low during the exponential phase, it increases by almost 5 folds in the stationary phase, and it is reduced again in older decaying cultures (Alvarez et al. 2002).

A potential application of TAG is in the production of biodiesel, monoalkyl esters of long-chain fatty acids with short-chain alcohols (Kurosawa et al. 2010). The use of bacterial oils for nutritional purposes is still uncertain because of economical reasons, and the lack of information on their digestion and absorption. However, bacterial oil can be used to produce soaps and detergents, plastics, personal care products, resins and lubricants, oleochemicals or ingredients for pharmaceutical products. Furthermore, these lipids can be utilized as protective coating for probiotic systems (Welson 2006).
2.6.3 Glycogen

Glycogen metabolism has mostly been the focus of enhanced biological phosphorus removal literature (Erdal et al. 2003; Fang et al. 2013). However, glycogen metabolism in activated sludge microorganisms during feast-famine cycles and N-limited conditions is not well understood. Furthermore, the biosynthesis and accumulation of storage lipids, such as PHA and
TAG, is a well-documented feature of *Rhodococcus* species. However, only recently the occurrence of glycogen in a *Rhodococcus* strain, *Rhodococcus jostii* RHA1, has been reported (Hernandez et al. 2008).

Glycogen is a glucose polymer with $\alpha$-1,4 and $\alpha$-1,6 linkages, which is formed by various bacteria. It is generally considered to be a storage compound providing both carbon and energy for the microorganisms that accumulate it (Preiss 1984). The storage of glycogen has been reported previously for other related *Actinobacteria*, including *Mycobacterium* (Belanger and Hatfull 1999) and *Corynebacterium* (Seibold et al. 2007). Recently, the accumulation of glycogen by cells of several species of the genus *Rhodococcus* was analyzed. In general, the glycogen content accounted to approximately up to 5% of cellular dry weight (CDW) in the strains studied in this work (Hernandez and Alvarez 2010).

The metabolic pathway of glycogen biosynthesis in bacteria is performed by the action of three enzymes (Figure 2.3), namely ADP-glucose pyrophosphorylase (ATP : $\alpha$-D-glucose-1-phosphate adenylyltransferase; EC 2.7.7.27), glycogen synthase (EC 2.4.1.21) and branching enzyme (EC 2.4.1.18), that are encoded by the *glgC*, *glgA* and *glgB* genes, respectively (Seibold et al. 2007).
To investigate storage metabolism in activated sludge bacteria using genome-scale flux balance analysis, we selected *Rhodococcus jostii* RHA1 as the model organism of heterotrophs. This microorganism can accumulate various storage compounds making it a good candidate to study storage metabolism in different environmental conditions (i.e., feast-famine cycles and non- and N-limited conditions) (Hernandez et al. 2008). Furthermore, its genome has already been sequenced and annotated (McLeod et al. 2006).
Rhodococcus is a genus of catabolically versatile soil bacteria that belong to Corynobacterineae, a mycolic acid-containing suborder of the GC-rich Actinobacteria phylum. Rhodococci are described as aerobic, Gram-positive, non-motile, mycolate-containing, nocardioform actinomycetes (Bell et al. 1998; Finnerty 1992). The ability of rhodococci to convert a broad range of organic compounds and pollutants, coupled with their robust growth and exceptional stress tolerance, has led to their use in a wide range of environmental and biotechnological applications from biodegradation of pollutants to the biocatalytic production of drugs and hormones and biodesulphurization of petroleum products (Chen et al. 2012; Van Der Geize and Dijkhuizen 2004; Yam et al. 2011).

Rhodococci are also commonly found in wastewater as the remarkable PCB degraders (Krzmarzick and Novak 2014). Among the mycolic acid-producing Actinobacteria, the mycolata, Rhodococci are also responsible for foam formation in activated sludge wastewater treatment systems (Petrovski et al. 2011).

Rhodococcus jostii RHA1 was isolated from lindane contaminated soil (Seto et al. 1995) and is best known for its superior ability to degrade a wide variety of organic compounds such as polychlorinated biphenyls (PCBs). The genome of R. jostii RHA1 comprises of 9.7 Mbp arranged in four linear replicons: one chromosome and three plasmids (McLeod et al. 2006). Its genome encodes a surprisingly large number of oxidoreductases, enzymes that are often involved in the hydroxylation and cleavage of aromatic compounds. This is in line with the degradation capability of a wide range of xenobiotic compounds, sterols and steroids of R. jostii RHA1 (Larkin et al. 2005; Van Oosterwijk et al. 2011).
R. jostii RHA1 can accumulate most of the storage compounds including poly-β-hydroxybutyrate (PHB), poly(3-hydroxyvalerate) (PHV) (Hernandez et al. 2008; Madison and Huisman 1999), triacylglycerols (TAG), wax esters (WE) (Hernandez et al. 2008; Ishige et al. 2002; Waltermann et al. 2007), α,α-trehalose, glycogen, and polyphosphate (PolyP). In this study, we focus only on glycogen, PHB, and TAG, since α,α-trehalose, PHV, and WE have similar characteristics as compared to glycogen, PHB, and TAG. Moreover, PolyP was not studied in this work, since it is not a carbon reserve.

2.8 References


Erdal UG, Erdal ZK, Randall CW. 2003. The competition between PAOs (phosphorus accumulating organisms) and GAOs (glycogen accumulating organisms) in EBPR (enhanced biological phosphorus removal) systems at different temperatures and the effects on system performance. Water Science and Technology. p 1-8.


In chapter 2 a literature review was provided covering topics including genome-scale metabolic network reconstruction and flux balance analysis, selection of objective functions, \(^{13}\text{C}\)-metabolic flux analysis, application of metabolic modeling in environmental engineering and sciences, storage compounds, and *Rhodococcus jostii* RHA1. In this first research chapter, the genome-scale metabolic model of *R. jostii* RHA1 was reconstructed and some of the basic network connectivity properties were presented. The modeling of a sudden change in environmental conditions was then investigated. Experimentally, the cultures of *R. jostii* RHA1 were allowed to deplete the nitrogen in the growth medium while the carbon source was in excess. This triggered the accumulation of the three storage compounds (PHA, TAG, and glycogen) in different fractions depending on the carbon substrate (acetate and glucose). The capacities of different objective functions to reproduce *in silico* the growth of *R. jostii* RHA1 in the non-limited conditions and the storage accumulation in the N-limited conditions were compared using a Baysian approach. It was found that maximizing storage during the N-limited conditions coupled with the *environmental* MoMA sub-objective function (which minimized the biochemical flux differences between the non-limited and N-limited conditions) was the most effective set of objective functions. This conclusion was substantiated by comparing the fluxes of the central metabolism of *R. jostii* RHA1 measured by \(^{13}\text{C}\)-metabolic flux analysis during non-limited growth and simulated by flux balance analysis for N-limited storage accumulation.
Chapter 3. Genome-scale Metabolic Model of *Rhodococcus jostii* RHA1 (iMT1174) to Study the Accumulation of Storage Compounds during Nitrogen Limited Condition

Mohammad Tajparast and Dominic Frigon

Submitted to

BMC Systems Biology
Abstract

*Rhodococcus jostii* RHA1 growing on different substrates is capable of accumulating three types of carbon storage compounds: glycogen, polyhydroxyalkanoates (PHA), and triacylglycerols (TAG). Under nitrogen-limited (N-limited) condition, the level of storage increases as is commonly observed for other bacteria. The proportion of each storage compound changes with substrate, but it remains unclear how the organism regulates the synthesis of the mixture of the storage compounds. We analysed the growth of *R. jostii* RHA1 under N-limited conditions using a genome-scale metabolic modeling approach to determine the global metabolic objective function pursued by the organism for storage metabolism.

The *R. jostii* RHA1 model (iMT1174) contains 1,243 balanced metabolites, 1,935 unique reactions, and 1,174 open reading frames (ORFs). NAD, NADH, and oxygen were the most frequent metabolites due to the presence of a large number of oxygenase and oxidoreductase reactions involved in xenobiotics degradative pathways in the *R. jostii* RHA1 model.

Seven objective functions were compared in this work applying the Bayesian approach using their posterior probability as relative likelihood criterion. The flux balance analysis of the model for the N-limited condition on glucose and acetate found that the maximization of the storage fluxes (in the N-limited condition)+MoMA objective function applied between the non-limited and N-limited conditions could explain most of the differences in the proportions of the three storage compounds (glycogen, PHA, and TAG) observed between different substrates. The posterior probability share of this objective function was the highest among the six investigated. Through a sensitivity analysis, we found that the max storage+MoMA predictions for the proportion of the storage compounds accumulated were most sensitive to the biomass composition, as expected.
$^{13}$C-labelling experiments, in conjunction with the $^{13}$C-metabolic flux analysis, showed that the metabolic behaviours of the central metabolism of *R. jostii* RHA1 are closely similar in the non-limited and N-limited conditions. In turn, it substantiated the application of the MoMA sub-objective function to compare two environmental conditions such as the non- and N-limited conditions. Since we compared the fluxes of central metabolism, thus this finding should happen for other experiments regardless of microorganisms being used.

The genome-scale metabolic model analysis suggested that *R. jostii* RHA1 tries to keep its carbon substrate uptake rate and its intracellular reaction flux profile as similar as possible during the N-limited condition compared to the immediately preceding non-limited growth. PHA as the main storage pool of the mixture in *R. jostii* RHA1 seems to be a result of the high demands in lipids for the biomass synthesis of *R. jostii* RHA1.
3.1 Background

The number of genomes and meta-genomes is increasing, but the prediction of metabolism in environmentally relevant conditions remain to be determined. Rhodococci are widely spread in many habitats such as soil, fresh water, seawater, activated sludge wastewater treatment systems (Bell et al. 1998; Finnerty 1992; Larkin et al. 2006; Warhurst and Fewson 1994). Their versatile metabolic activities make them good candidates for environmental bioremediation, and for a number of industrial applications including biodesulfurization of fossil fuels (Kilbane Li and Jackowski 1992; McFarland 1999; McFarland et al. 1998), production of biosurfactants (Bell et al. 1998), and production of acrylic acid (Finnerty 1992), just to name a few. Specifically, *Rhodococcus jostii* RHA1 (RHA1) was isolated from lindane-contaminated soil and is best known for its superior ability to biodegrade polychlorinated biphenyls (PCBs) (Seto et al. 1995). The genome of RHA1 has been completely sequenced and annotated (McLeod et al. 2006). It holds one of the largest bacterial genomes sequenced to date with 9.7 Mbp and 9,221 predicted open reading frames (ORFs).

To face rapidly changing environmental conditions in its natural habitat, the genome of *R. jostii* RHA1 also contains many genes for the metabolism of various storage compounds including polyphosphate (PolyP), glycogen, wax esters (WE), triacylglycerols (TAG), and polyhydroxyalkanoates (PHA) (Hernandez et al. 2008). Figure 3.1 shows the simplified possible metabolic pathways for the substrates and storage compounds included in this study. One may predict glycogen accumulation on glucose and PHB accumulation on acetate, since the metabolic pathways from glucose to glycogen and acetate to PHB are short. PHV can be produced via two different pathways in this model: one from succinyl-coA and the other from L-valine. It was experimentally observed that, under N-limited condition on different substrates, *R. jostii* RHA1 accumulates these various storage compounds at the same time, but in different proportions.
(Hernandez et al. 2008). It remains unclear how the cells regulate the carbon fluxes to the various storage compounds.

**Figure 3.1** – Simplified possible pathways of storage metabolism in *R. jostii* RHA1 for the three main storage compounds: glycogen, PHA, and TAG.

- **glc-D**: alpha-D-Glucose
- **adpglc**: ADP-glucose
- **g1p**: D-Glucose 1-phosphate
- **g6p**: alpha-D-Glucose 6-phosphate
- **g3p**: D-Glyceraldehyde 3-phosphate
- **pyr**: Pyruvate
- **2acl**: 2-Acetolactate
- **23dh3mb**: 2,3-Dihydroxy-3-methylbutanoate
- **3mob**: 3-Methyl-2-oxobutanoic acid
- **val-L**: L-Valine
- **smms**: (S)-Methylmalonate semialdehyde
- **accoa**: Acetyl-CoA
- **oaa**: Oxaloacetate
- **TCA**: Tricarboxylic acid cycle
- **succoa**: Succinyl-CoA
- **mmcoa-R**: (R)-Methylmalonyl-CoA
- **mmcoa-S**: (S)-Methylmalonyl-CoA
- **ppcoa**: Propanoyl-CoA
- **3kvcoa**: 3-Ketovaleryl-CoA
- **r3hvcoa**: (R)-3-Hydroxyvaleryl-CoA
- **PHB**: Poly-β-hydroxybutyrate
- **PHV**: Poly-3-hydroxyvalerate
- **PHA**: Poly-β-hydroxyalkanoate
- **3hbcoa**: (S)-3-Hydroxybutanoyl-CoA
- **aacoa**: Acetoacetyl-CoA
- **ac**: Acetate
- **acac**: Acetoacetate
- **TAG**: Triacylglycerol
- **glyc3p**: Glycerol 3-phosphate
- **acylACP**:

In order to investigate possible metabolic principles “guiding” the cell in determining the regulation of the combined synthesis of storage compounds, we reconstructed *in silico* the genome-scale metabolic network of *R. jostii* RHA1 (i.e., iMT1174). For this reconstruction, we defined the storage compounds independently of the biomass composition as pseudo-secretory by-products; that is, although storage compounds are not secreted, they are set so that they are secreted by the metabolic model for analysis of the storage metabolism. This model definition allows us to study the variation of storage metabolism.

Genome-scale metabolic models are typically underdetermined (i.e., no unique solution exist) because the number of metabolites (mass balance equations) are lower than the number of reaction fluxes (variables). Consequently, they can be studied by means of linear programming to find an optimal metabolic flux profile as defined by a linear objective function (optimization criterion). These objective functions could be understood as a possible global regulatory principle implemented by the cell. In the context of this study, here are the two main questions we aim to answer. (1) Is there an objective function that enables us to predict the proportion of storage compound(s) and their associated metabolic fluxes observed under N-limited condition? (2) What are the main factors influencing the proportion of the storage compounds accumulated?

Several objective functions have been defined and examined in the past to study a large number of metabolic situations (Knorr et al. 2007; Palsson 2006), we give here some examples. For a number of bacterial species, the maximization of the growth rate (or yield as they are not always independent) was reported to be successful at predicting experimentally observed exponential growth phenotypes (Edwards et al. 2001; Ibarra et al. 2002; Segre et al. 2002). For
minimization of cellular “effort” in utilizing available energy and external resources, it was found that the minimization of the sum of reaction fluxes (Manhattan norm of the flux vector) was a good way to predict the outcome of the metabolic network (Holzhutter 2004), which was interpreted as a maximally efficient use of the available biochemical reactions. Finally, for gene deletion mutants, a successful objective function was to find the reaction flux distribution profiles most similar to the ones of the wild-type growing in the same conditions, an objective function known as minimization of metabolic adjustments (MoMA) (Segre et al. 2002).

In the current study, maximization of the growth rate was adopted as the main objective function for the growth of RHA1 during non-limited growth condition. For the N-limited storage accumulation, seven objective functions were examined including: 1) minimization of the metabolic adjustment between the N-limited and non-limited conditions (environmental MoMA), 2) minimization of the metabolic fluxes (i.e., minFluxes or flux minimization) (Holzhutter 2004), 3) minimization of ATP production rate (minATP), 4) maximization of ATP production rate (maxATP), 5) minimization of the production rate of redox potential (minNADH), 6) maximization of the total storage fluxes in the N-limited condition (maxStorage), and 7) maximization of the total storage fluxes in conjunction with environmental MoMA (maxStorage+MoMA). The implementation of these objective functions consisted in developing a nested optimization algorithm by first maximizing the growth rate under nutrient-limited and non-limited conditions, and then searching the solution space thus delimited using the above mentioned objective functions.

The performance of the objective functions in predicting the accumulation of the mixture of storage compounds on two different carbon sources (glucose and acetate) under N-limited conditions were compared with our experimental observations using a Bayesian approach. $^{13}$C-
labelling experiments, along with $^{13}$C-metabolic flux analysis, were conducted to substantiate the application of the *environmental* MoMA sub-objective function comparing the two environmental conditions (i.e., non-limited vs. N-limited conditions).

### 3.2 Materials and Methods

#### 3.2.1 Highlights of the Model Reconstruction Procedure

As the first step in reconstructing the cellular metabolic network of *R. jostii* RHA1, the list of genes in its annotated genome sequence was converted to a list of associated balanced biochemical reactions, which form the basic framework of the model. Subsequently, the model was analyzed in light of literature data to determine if some reactions may have been missing from the list of annotated genes. The reactions were added along with transport reactions to make a physiologically meaningful cellular model. In the case of RHA1, although annotated, several degradation pathways of various xenobiotics were missing from the Kyoto Encyclopedia of Genes and Genomes (KEGG) (the primary biochemical database used (Kanehisa et al. 2004)) and they were added to the model using literature data and other online databases (Gao et al. 2010; Karp et al. 2005) to account for its catabolic ability. Then, the metabolic network was completed by the pathways of seven storage compounds that have also been considered and manually curated in the model: glycogen, α,α-trehalose, PHA such as PHB and PHV (Hernandez et al. 2008; Madison and Huisman 1999), TAG, WE (Hernandez et al. 2008; Ishige et al. 2002; Waltermann et al. 2007), and PolyP. Finally, eight macromolecules were assigned to make up the biomass (e.g., protein, DNA, RNA, phospholipids, small molecules such as electron carriers and coenzymes, peptidoglycan, carbohydrates, and corynomycolic acids). We also figured out the directionality and reversibility of most of the reactions in the *iMT1174* model using online databases (Caspi et al. 2010; Schellenberger et al. 2010). The non-growth-associated
maintenance energy (NGAME) was set differently for different objective functions to adjust the total flux of storage compounds (g-COD/(g-COD-Xa·h)) to values comparable to those of the experimental data. Finally, the growth-associated maintenance energy (GAME) was set to 30 mmol-ATP/g-DW according to the literature data (Borodina et al. 2005).

Additional File 1 details the biomass and storage lipid compositions. Additional Files 2 and 3 also provide the complete lists of genes, enzymes, reactions, and metabolites appear in the network of RHA1. Additional File 2 shows all reactions involved in specified biochemical pathways, along with added reactions; it contains reaction properties such as name, ID, definition, equation, directionality, lower and upper bounds, enzyme commission (EC) numbers, ORFs, replaced reactions, and references. Additional File 3 summarizes unique reactions and metabolites, along with their properties, that were simulated in this work. This biochemical network was converted into a stoichiometric matrix (S) with rows and columns corresponding to the metabolites and reactions, respectively, in order to be quantified by means of the FBA technique.

3.2.2 Flux Balance Analysis

The gene-protein-reaction (GPR) association, along with the stoichiometric matrix, is a qualitative representation of the cellular network. The metabolic model of RHA1 was studied quantitatively by setting up a series of mass balances around each intracellular metabolite that can be expressed in matrix notation as:  \( \frac{dx}{dt} = Sv \); where \( X \) is the (m×1) vector of the concentrations of the balanced metabolites (i.e., intracellular and macromolecular compounds), \( v \) denotes the (n×1) vector of the entire metabolic fluxes, and \( S \) stands for the m×n stoichiometric matrix; note that in the case of the iMT1174 model with the unique reactions, the dimension of the stoichiometric matrix is 1243×1935.
The model was basically solved by assuming that the cellular network is at steady-state, which simplifies the previous equation to \( \mathbf{S} \cdot \mathbf{v} = \mathbf{0} \). This assumption is plausible due to the fast equilibration of intracellular metabolite pools (time-scale of seconds) compared to the time-scale of genetic regulation (minutes) (Segre et al. 2002) and due to high reaction rates of intracellular reactions compared to phenotypic exchange rates such as substrate uptake, cell growth, and by-product secretion rates (Lee et al. 2006; Roels 1982; Varma and Palsson 1994). Consequently, the output of the model is a distribution of metabolic fluxes through the various chemical reactions (note that the kinetics is not modeled). Since the number of fluxes normally exceeds the number of metabolites \( n > m \), the problem is said to be underdetermined. In this case, there is a solution space for the fluxes that can be studied by optimizing proper objective functions subject to defined constraints (physicochemical constraints as mass balance and enzymatic capacity constraints as inequalities) (Becker et al. 2007).

Here for the first time, we have implemented seven objective functions namely: 1) minimization of the metabolic adjustment between N-limited and non-limited conditions (environmental MoMA), 2) minimization of the metabolic fluxes (i.e., minFluxes) (Holzhutter 2004), 3) minimization of ATP production rate (minATP), 4) maximization of ATP production rate (maxATP), 5) minimization of the production rate of redox potential (NADH) (e.g., minNADH), 6) maximization of the total storage fluxes in the N-limited condition (maxStorage), and 7) maximization of the total storage fluxes in conjunction with environmental MoMA (maxStorage+MoMA) in order to investigate storage metabolism of the iMT1174 metabolic network in both the N- non-limited and limited conditions. We solved the linear programming problem using the COBRA toolbox (Becker et al. 2007) within *The Language of Technical Computing* MATLAB environment (Mathworks Inc, Massachusetts).
With seven sub-objective functions, we maximized the growth rates in both non-limited and N-limited conditions, and then applied the sub-objective functions. The model was balanced for elemental mass and charges and a number of reactions were considered irreversible (physicochemical constraints), but the enzymatic capacity of reactions remained unbounded. The substrate uptake rate (in the non-limited and N-limited conditions), phosphorus and ammonium uptake rates in the non-limited condition were constrained according to our experimental observations. N-limited conditions were then simulated by reducing the uptake flux of NH$_3$ into the cellular network. All the simulation results are presented in yields.

Minimization of redox potential was chosen as a plausible objective function since it can be referred as maximization of energy efficiency by microorganisms. A cell minimizes redox potential to reduce the number of oxidizing reactions that happen and in turn conserving its energy or utilizing its energy in the most efficient way possible (Knorr et al. 2007). Minimization of ATP production was investigated, since this involves efficient energy use. The aim of this objective function is for the cell to grow while utilizing the minimum amount of energy required, thus conserving ATP. In turn, the cell uses ATP as efficiently as possible. Note that this objective function is concerned with efficient metabolism, it is different from minimizing redox potential, as ATP is the compound in question instead of NDAH (Savinell and Palsson 1992). We also proposed that the ATP production rate should be as high as possible, increasing availability of ATP as much as possible. However, this objective function would naturally make less sense, as suggested in the description of minimizing the rate of ATP production. Although it was often used as an objective function in modeling mitochondria (Ramakrishna et al. 2001), it may not be the best interest of bacteria to generate high amount of ATP, particularly if resources of the bacterial cell could be channeled towards other processes and the ATP made might be
excessively wasted. Nevertheless, maximization of ATP production rate was also analyzed in this study. It should be noted that all the above-mentioned objective functions were subsets of the minimization of fluxes objective function and in this work we showed that their predictions were identical. Therefore, maximization of NADH will also show the same predictions as the others.

3.2.3 $^{13}$C-Metabolic Flux Analysis

We aim at estimating the reactions involved in central metabolism in the non-limited condition using $^{13}$C-metabolic flux analysis ($^{13}$C-MFA). The model represents the central metabolism which includes glycolysis and gluconeogenesis, Entner-Doudoroff pathway, tricarboxylic acid (TCA) cycle, pentose phosphate pathway, anaplerotic carboxylation and decarboxylation, storage metabolic reactions, amino acid biosynthetic reactions, and anabolic routes into biomass. Additional File 4 shows the central metabolic pathways of $R$. jostii RHA1 for which $^{13}$C-MFA was implemented. We used the openFLUX software application under MATLAB environment (Mathworks Inc, Massachusetts) to solve for the fluxes (Quek et al. 2009). The application is based on the Elementary Metabolite Unit (EMU) framework. Stoichiometric data on growth, substrate uptake rate, storage formation, and on the cellular composition of $R$. jostii RHA1 together with mass isotopomer distribution data of the labelled amino acids (Additional File 5) that were produced using the iMS2Flux software (Poskar et al. 2012) were used as model input.

3.2.4 Experimental Procedures

To validate the model we performed two different sets of culture conditions namely: N- and non-limited conditions of $R$. jostii RHA1 in batch; labelled and unlabelled glucose and acetate were employed as the sole carbon sources. The labeling tests were performed using a μ-
24 bioreactor (Applikon Biotechnology), while the unlabelled batch tests were accomplished with a respirometer model AER-200 (Challenge Technology).

3.2.5 Bacterial Strain and Growth Conditions

*R. jostii* strain RHA1 was cultivated aerobically at 28 °C and pH 7 in mineral salts medium (MSM) according to Schlegel et al. (Schlegel et al. 1961). The MSM medium contains: NH₄Cl (1.00 g/L), KH₂PO₄ (1.50 g/L), Na₂HPO₄·12H₂O (9 g/L), FeNH₄-Citrate (1.20 mg/L), MgSO₄·7H₂O (0.20 g/L), CaCl₂ (0.02 g/L), Hoagland solution (2.00 mL/L), and carbon source (10 g-COD/L). Glucose and sodium acetate were used as the sole carbon sources. Cells were harvested during mid-exponential (non-limited condition) and end-exponential (both non- and N-limited conditions), and late stationary phases (N-limited condition), centrifuged and stored at −80 °C for further analyses. When N-limiting conditions were applied, the concentration of ammonium chloride in MSM was reduced to 0.1 g/L (N-Lim) to allow lipid accumulation. When labeling conditions were specified, 1-¹³C-Acetate and 2-¹³C-Acetate Sodium salt, 1-¹³C-Glucose, and U-¹³C-Glucose were utilized as the sole carbon sources and cells were harvested during the exponential growth phase.

3.2.6 Biochemical Analysis

Total and volatile suspended solids (TSS/VSS) were measured according to the standard method 2540 D/2540 E (Clesceri et al. 1998); chemical oxygen demand (COD) was quantified according to the standard method 5220 D (Clesceri et al. 1998); phosphorus was measured by the Ascorbic Acid method (standard methods 4500-P E) (Clesceri et al. 1998); ammonium concentration was measured by colorimetry (Rhine et al. 1998) (spectrophotometric measurements at 630 nm); cellular glycogen was quantified by a hexokinase enzymatic kit and
colorimetry (Hexokinase protocol, measurement of NADH concentration at 340 nm, Sigma, St. Louis, MO) (Maurer et al. 1997); cellular PHA was measured by colorimetry (Law and Slepecky 1961; Paganelli et al. 2011); (spectrophotometric measurements at 235 nm). Spectrophotometric measurements of the last four components were performed in microplates using a SpectraMax5 reader (Molecular Devices, LLC, USA). Cellular TAG was measured by gas chromatography (GC) according to (Brandl et al. 1988) with some modifications.

3.2.7 Analysis of Fatty Acids
For the identification of lipids, cultures were harvested by centrifugation at 6000g for 15 min and lyophilisation overnight. 5-8 mg of lyophilized cells was resuspended in 0.6 mL of chloroform and 0.6 mL of methanol containing 15% (v/v) H₂SO₄; 1 µL Trinonadecanoin (Nu-Check Prep Inc., Elysian, MN) was used as internal standard for quantification (Frigon 2005). Methanolysis was carried out at 100 °C for 2.5 h. After cooling to room temperature and then on ice, 0.3 mL of deionized water was added to the solution, which was then vigorously vortexed for 1 min. After phase separation, 0.45 mL of the organic phase (bottom layer) was removed and transferred to a small screw-cap glass vial (Brandl et al. 1988). The organic phase containing fatty acid methyl esters (FAMEs) was analyzed by using an Agilent 6890N GC system equipped with an Agilent HP-88 column (60 m by 0.25 mm, 0.2 µm thick film) with helium as the carrier gas at flow rate of 2 mL/min. A 1 µL portion of the organic phase was injected with a 50:1 split ratio using the auto-sampler. The inlet was maintained at 250 °C. The oven was held at 175 °C for 15 min, heated to 220 °C at 3 °C/min, and then held at 220 °C for 5 min. Peak detection was performed by a flame ionization detector, which was maintained at 280 °C. The fatty acids were identified and quantified by comparison to standard FAMEs (Sigma). Fatty acid content was defined as the percentage of the ratio of fatty acids to cell dry weight (% CDW).
3.2.8 Analysis of Labelled Amino Acids

Proteinogenic amino acids of the biomass were analyzed using gas chromatography mass spectrometry (GC-MS) technique according to Nanchen et al. with some modifications (Nanchen et al. 2007). The biomass samples which had already been stored at −80 °C were thawed and resuspended in 1.5 mL sterile deionized water, vortexed, and homogenized using ultrasonic treatment (60 watts) for 15 min. After resuspension, 100 µL of sample was transferred into a 2 mL microcentrifuge tube, further spun down the cell pellets (at 15800g at room temperature for 15 min). Cell pellets were washed twice by resuspension in 1 mL 0.9% NaCl, and centrifuged at 15800g at room temperature for 15 min. The washed pellets were resuspended in 1 mL of 6 M HCl, and hydrolyzed for 24 h at 110 °C in a well-sealed screw-capped tube to prevent evaporation. The hydrolyzate was dried overnight in a heating block at 70 °C and under a constant air stream in a fume hood. The dry hydrolyzate was dissolved in 30 µL of a reagent containing 10 mg of Methoxyamine Hydrochloride per 1 mL of Anhydrous Pyridine. 1 µL of the internal standard was added into each sample tube, note that the internal standard was 750 ng/µL of deuterated Myristic acid (so-called D$_{27}$-myristic acid), the samples were mixed by vortex and sonication for several times, each time taking 10-20 sec. Furthermore, the samples were centrifuged at 15000 rpm for 10 min at room temperature. The samples were transferred into GC-MS vials and cooked at 70 °C for 30 min in a hot block. 70 µL of N-tert-butyldimethysilyl-N-methyltrifluoroacetamide (TBDMS) was added into the vials and cooked at 70 °C for 1 hr in a hot block to derivatize the samples. GC-MS analyses were performed with an Agilent 5975C mass selective detector coupled to a 7890A gas chromatograph (Agilent Technologies, Santa Clara, CA, USA) fitted with a 7693 autosampler and a DB-5MS+DG capillary column (30 m plus 10 m Duraguard®), diameter 0.25 mm, film thickness 0.25 mm (Agilent J & W, Santa Clara,
CA, USA). The GC temperature program started with a 1 min hold at 60 °C followed by a 10 °C/min ramp to 300 °C. Bake-out was at 320 °C for 10 min. The injector and interface to the MS were held at 285 °C. The helium carrier flow rate was held constant at 1.5 mL/min (or a flow rate such that the TBDMS derivative of the D27-myristic acid has a retention time of 18 min). When operated in full scan mode, the scan range was 50 - 700 Da. 1 μL of the sample was injected in splitless mode.

3.2.9 Statistical Analyses
Elemental mass balances on the measured conversions of substrate (either glucose or acetate), chemical oxygen demand (COD), NH$_4^+$, PO$_4^{3-}$, biomass, glycogen, PHB, PHV, TAG, O$_2$, and CO$_2$ were performed to check the consistency of the data using the procedure of Van Der Heijen et al. (1994). There were more conversions measured than needed to define the whole system with elemental balances. Macrobal software (Hellinga, 1992) was employed to reconcile the balances of all elements and compounds and calculating errors using the elemental composition matrix (Table 3.1). The $\chi^2$ test was used to evaluate the correlation between the measured and estimated rates at 99% confidence level (corresponding to $\alpha$ level of 0.01). $h$ is the measured statistics that is $\chi^2$-distributed with a number of degrees of freedom equal to the rank of the covariance matrix of the residuals. An error is not significant if $h < \chi^2$. This also accepts the null hypothesis stating that there is no systematic error for the appropriate number of degrees of freedom and the desired confidence level in the model and there is no difference between the measured and predicted conversion rates (Van Der Heijen et al. 1994).
Table 3.1 – Macrobal elemental composition matrix in the non-limited growth condition

<table>
<thead>
<tr>
<th>Conversion</th>
<th>C&lt;sup&gt;g&lt;/sup&gt;</th>
<th>COD</th>
<th>N&lt;sup&gt;h&lt;/sup&gt;</th>
<th>P&lt;sup&gt;i&lt;/sup&gt;</th>
<th>COD&lt;sub&gt;total&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substrate</td>
<td>0.38</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>COD&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>NH&lt;sub&gt;4&lt;/sub&gt;+</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PO&lt;sub&gt;4&lt;/sub&gt;³⁻</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>µ&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.32</td>
<td>1</td>
<td>0.07</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>Glycogen</td>
<td>0.38</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PHB&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.33</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PHV&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.31</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TAG&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.27</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OUR&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<sup>a</sup>: Chemical oxygen demand; <sup>b</sup>: growth rate; <sup>c</sup>: poly-β-hydroxybutyrate; <sup>d</sup>: poly(3-hydroxyvalerate); <sup>e</sup>: triacylglycerol; <sup>f</sup>: oxygen uptake rate; <sup>g</sup>: carbon; <sup>h</sup>: nitrogen; <sup>i</sup>: phosphorus.

Note that, in the N-limited condition the columns corresponding to N and P, along with the rows corresponding to NH<sub>4</sub>⁺, PO<sub>4</sub>³⁻, and µ are absent.

A Bayesian-based objective function discrimination method was also applied to compare the performance of the objective functions and find the best objective function (Knorr et al. 2007). Briefly, this method calculates the normalized posterior probability of the models (objective functions), given a set of experimental observations, that is a function of some variables, parameters and errors. The model with the largest posterior probability is considered the most probable one predicting the experimental data (Knorr et al. 2007).

### 3.3 Results and Discussion

#### 3.3.1 Features of the RHA1 Metabolic Network and Connectivity

The reconstructed in silico metabolic network of *R. jostii* RHA1 (iMT1174) contains 1,243 balanced (intracellular) compounds, 1,935 unique reactions, and 1,174 ORFs (Table 3.2). Totally, it also includes 330 extracellular compounds that are associated with 518 exchange reactions. Through manual curation, 495 reactions were added to fill metabolic gaps in the network, which was justified by biochemical literature data from either *R. jostii* RHA1 or related
species. Beside gap fillers, we made a special effort to include a number of xenobiotics degradation pathways (consisting of 131 transport reactions and 326 conversion reactions) in order to make the model most useful for the research community interested in *R. jostii* RHA1.

Since we would like to model bacteria growth in a minimal medium, we set most exchange reactions (out of 518 exchange reactions) to zero. However, we fixed substrate uptake rate and we set oxygen and nutrients (such as phosphorus, nitrogen, and minerals) uptake rates unbounded to be calculated with the flux balance analysis. There was also no secretion of carbon except for CO₂ production rate that was set unbounded to be calculated with the flux balance analysis.

Table 3.2 – Properties of the *in silico* metabolic network of *R. jostii* RHA1 (iMT1174)

<table>
<thead>
<tr>
<th>Property</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathway</td>
<td>114</td>
</tr>
<tr>
<td>Total reactions (unique)</td>
<td>3,007 (1,935)</td>
</tr>
<tr>
<td>Biochemical conversion (unique)</td>
<td>2,489 (1,417)</td>
</tr>
<tr>
<td>Transport</td>
<td>518</td>
</tr>
<tr>
<td>Reactions with ORFᵃ ( % of total reactions)</td>
<td>1,876 (62.39%)</td>
</tr>
<tr>
<td>Number of ORFs</td>
<td>1,174</td>
</tr>
<tr>
<td>Total metabolites</td>
<td>1,573</td>
</tr>
<tr>
<td>Intracellular</td>
<td>1,243</td>
</tr>
<tr>
<td>Extracellular</td>
<td>330</td>
</tr>
</tbody>
</table>

ᵃ: Open reading frame.

Seven storage compounds were defined independently of the biomass composition as extracellular compounds connected to 11 fictional transport reactions to study the variation of storage compound accumulation namely: glycogen, a,a-trehalose, poly-β-hydroxybutyrate (PHB), poly(3-hydroxyvalerate) (PHV), TAG, WE, and the non-carbon-based PolyP. As these compounds were observed experimentally or detailed genomic investigation revealed the presence of specific genes (Hernandez et al. 2008), pathways for the synthesis and the degradation of these compounds were included in the biochemical model. Biomass composition of *R. jostii* RHA1 is detailed in Additional File 1, all reactions involved in genome-scale
metabolic model of *R. jostii* RHA1 are detailed in Additional File 2, and unique reactions and metabolites involved in the genome-scale metabolic model of *R. jostii* RHA1 are detailed in Additional File 3.

Frequency of all the metabolites involved in the *i*MT1174 model was calculated and compared to the genome-scale models of *Escherichia coli* K-12 (*i*JR904) (Reed et al. 2003) and *Mycobacterium tuberculosis* (GSMN-TB) (Beste et al. 2007) through the frequency plot of the 20 most frequent metabolites (Figure 3.2). Proton and water are the most frequent metabolites in the *i*JR904 and *i*MT1174 models, respectively; however, proton is the 82nd metabolite in the GSMN-TB model that is due to the use of neutral forms of the metabolites in this model; moreover, water was excluded from the GSMN-TB model. The most connected metabolites are associated with the energy metabolism (i.e., ATP, orthophosphate, and ADP) in the *i*JR904 model and in the GSMN-TB model. However, NAD, NADH, and oxygen are more frequent than the energy-associated metabolites in the *i*MT1174 model. This is because of the large number of oxygenase and oxidoreductase reactions involved in xenobiotics degradation pathways and included in the current *i*MT1174 model. Finally, NADPH has similar connectivity in the three metabolic models, which is expected as it is the main reducing power in anabolism and cell synthesis.
Figure 3.2 – The frequency plot of the 20 most frequent metabolites in the biochemical networks.

Microorganisms are: Escherichia coli K-12 (iJR904) (Reed et al. 2003), Mycobacterium tuberculosis (GSMN-TB) (Beste et al. 2007), and R. jostii RHA1 (iMT1174) [this study]. H: proton, H₂O: water, ATP: Adenosine 5’-triphosphate, PI: Orthophosphate, ADP: Adenosine 5’-diphosphate, NAD: Nicotinamide adenine dinucleotide, PPI: Diphosphate, NADH: Reduced nicotinamide adenine dinucleotide phosphate, CO₂: Carbon dioxide, NADP: Nicotinamide adenine dinucleotide phosphate, PYR: Pyruvate, NADPH: Reduced nicotinamide adenine dinucleotide phosphate, L-GLU: L-Glutamate, COA: Coenzyme A, NH₄: Ammonium, AMP: Adenosine 5’-

3.3.2 Experimentally Measured Storage Accumulation Rates

The conversion rates of various storage compounds on glucose and acetate as the sole carbon sources during non-limited and N-limited conditions were measured in order to evaluate objective functions. Oxygen uptake rate (OUR) time profiles of RHA1 in culture media containing acetate and either non-limiting or N-limiting concentrations of ammonia demonstrated a clear change in the respiration behaviour of the culture when the nitrogen source was exhausted (Figure 3.3); OUR profiles obtained with glucose as the carbon source exhibited similar trends (data not shown). Samples to determine the composition of the non-limited cultures were obtained at the middle and the end of the exponential phase. Initial ammonia concentrations in the N-limited cultures were adjusted such that the N-source would be completely consumed at approximately the equivalent point as the mid-exponential of the non-limited cultures. Thus, the composition of the N-limited cultures was based on samples obtained immediately before the N-Source was completely consumed and after the carbon source was depleted in the media (Figure 3.3).
Figure 3.3 – The oxygen uptake rate time profile of RHA1 growth on acetate in the non-limited and N-limited conditions.

Three samples were collected at the middle and end of the exponential growth phase in the non-limited condition and at the end of feed in the N-limited condition.

As mentioned earlier, elemental mass balances on the measured conversions of substrate (either glucose or acetate), chemical oxygen demand (COD), \( \text{NH}_4^+ \), \( \text{PO}_4^{3-} \), biomass, glycogen, PHB, PHV, TAG, \( \text{O}_2 \), and \( \text{CO}_2 \) were performed to check the consistency of the data. The balanced conversion rates for the acetate and glucose cultures under non-limited and N-limited
conditions are presented in Tables 3.3-3.6. The calculated $h$ statistics compared to the critical $\chi^2$ values suggest that the measured rates were consistent with respect to the various elemental and degree of reduction balances (Tables 3.3-3.6). These conversion rates were used to constrain the genome-scale metabolic model of *R. jostii* RHA1, $^{13}$C-metabolic flux models, and also compare the experimentally observed storage production rates with those obtained via the flux balance analysis technique.

**Table 3.3 – Matrix of the conversion rates on glucose as the sole carbon source in the non-limited batch growth condition**

<table>
<thead>
<tr>
<th>Conversion</th>
<th>Units</th>
<th>Calculation</th>
<th>Measurements</th>
<th>Balance Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Substrate</strong></td>
<td>g-COD/(g-COD-Xa⋅h)</td>
<td>B</td>
<td>$-2.94 \times 10^{-1}$</td>
<td>$-2.99 \times 10^{-1}$</td>
</tr>
<tr>
<td><strong>COD</strong></td>
<td>g-COD/(g-COD-Xa⋅h)</td>
<td>C</td>
<td>NA</td>
<td>$-2.99 \times 10^{-1}$</td>
</tr>
<tr>
<td><strong>NH$_4^+$</strong></td>
<td>g-N/(g-COD-Xa⋅h)</td>
<td>C</td>
<td>NA</td>
<td>$-2.99 \times 10^{-1}$</td>
</tr>
<tr>
<td><strong>PO$_4^{3-}$</strong></td>
<td>g-P/(g-COD-Xa⋅h)</td>
<td>B</td>
<td>$-1.00 \times 10^{-2}$</td>
<td>$-2.61 \times 10^{-3}$</td>
</tr>
<tr>
<td><strong>µ</strong></td>
<td>1/h</td>
<td>C</td>
<td>NA</td>
<td>$-2.61 \times 10^{-3}$</td>
</tr>
<tr>
<td><strong>Glycogen</strong></td>
<td>g-COD/(g-COD-Xa⋅h)</td>
<td>B</td>
<td>$1.53 \times 10^{-3}$</td>
<td>$1.50 \times 10^{-3}$</td>
</tr>
<tr>
<td><strong>PHB</strong></td>
<td>g-COD/(g-COD-Xa⋅h)</td>
<td>B</td>
<td>$1.13 \times 10^{-2}$</td>
<td>$1.13 \times 10^{-2}$</td>
</tr>
<tr>
<td><strong>PHV</strong></td>
<td>g-COD/(g-COD-Xa⋅h)</td>
<td>B</td>
<td>$2.64 \times 10^{-2}$</td>
<td>$2.64 \times 10^{-2}$</td>
</tr>
<tr>
<td><strong>TAG</strong></td>
<td>g-COD/(g-COD-Xa⋅h)</td>
<td>B</td>
<td>$3.70 \times 10^{-3}$</td>
<td>$3.70 \times 10^{-3}$</td>
</tr>
<tr>
<td><strong>OUR</strong></td>
<td>g/(g-COD-Xa⋅h)</td>
<td>B</td>
<td>$2.53 \times 10^{-2}$</td>
<td>$2.53 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

$^{a}$: Chemical oxygen demand; $^{b}$: growth rate; $^{c}$: poly-β-hydroxybutyrate; $^{d}$: poly(3-hydroxyvalerate); $^{e}$: triacylglycerol; $^{f}$: oxygen uptake rate; $^{g}$: this column indicates whether the compound will be balanced (B), which means that a better estimate will be found and be calculated (C) using the elemental composition matrix in Table 3.2; $^{h}$: not applicable. Note that $h = 0.07$ against $\chi^2 = 6.63$ at 99% confidence level.
Table 3.4 – Matrix of the conversion rates on glucose as the sole carbon source in the N-limited batch condition

<table>
<thead>
<tr>
<th>Conversion</th>
<th>Units</th>
<th>Calculation</th>
<th>Measurements</th>
<th>Balance Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conversions</td>
<td>Error</td>
</tr>
<tr>
<td>Substrate</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>-2.80×10⁻²</td>
<td>1.23×10⁻²</td>
</tr>
<tr>
<td>COD</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>-4.42×10⁻²</td>
<td>1.36×10⁻²</td>
</tr>
<tr>
<td>Glycogen</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>6.00×10⁻⁴</td>
<td>2.00×10⁻⁴</td>
</tr>
<tr>
<td>PHB</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>7.80×10⁻³</td>
<td>4.90×10⁻³</td>
</tr>
<tr>
<td>PHV</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>1.83×10⁻²</td>
<td>1.15×10⁻²</td>
</tr>
<tr>
<td>TAG</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>5.60×10⁻³</td>
<td>1.60×10⁻³</td>
</tr>
<tr>
<td>OUR</td>
<td>g/(g-COD-Xa·h)</td>
<td>B</td>
<td>-1.86×10⁻²</td>
<td>6.00×10⁻³</td>
</tr>
<tr>
<td>CO₂</td>
<td>g/(g-COD-Xa·h)</td>
<td>C</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

^a: Chemical oxygen demand; ^b: poly-β-hydroxybutyrate; ^c: poly(3-hydroxyvalerate); ^d: triacylglycerol; ^e: oxygen uptake rate; ^f: this column indicates whether the compound will be balanced (B), which means that a better estimate will be found and be calculated (C) using the elemental composition matrix in Table 3.2; ^g: not applicable. Note that h = 1.66 against $\chi^2 = 9.21$ at 99% confidence level.

Table 3.5 – Matrix of the conversion rates on acetate as the sole carbon source in the non-limited batch growth condition

<table>
<thead>
<tr>
<th>Conversion</th>
<th>Units</th>
<th>Calculation</th>
<th>Measurements</th>
<th>Balance Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conversions</td>
<td>Error</td>
</tr>
<tr>
<td>Substrate</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>-1.14</td>
<td>3.68×10⁻¹</td>
</tr>
<tr>
<td>COD</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>C</td>
<td>NA ^h</td>
<td>NA</td>
</tr>
<tr>
<td>NH₄⁺</td>
<td>g-N/(g-COD-Xa·h)</td>
<td>C</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>PO₄⁻³</td>
<td>g-P/(g-COD-Xa·h)</td>
<td>B</td>
<td>-5.5×10⁻³</td>
<td>3.50×10⁻³</td>
</tr>
<tr>
<td>µ</td>
<td>1/h</td>
<td>C</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Glycogen</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>1.70×10⁻³</td>
<td>8.00×10⁻⁴</td>
</tr>
<tr>
<td>PHB</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>5.50×10⁻³</td>
<td>1.80×10⁻³</td>
</tr>
<tr>
<td>PHV</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>1.28×10⁻²</td>
<td>4.30×10⁻³</td>
</tr>
<tr>
<td>TAG</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>7.60×10⁻³</td>
<td>5.60×10⁻³</td>
</tr>
<tr>
<td>OUR</td>
<td>g/(g-COD-Xa·h)</td>
<td>B</td>
<td>-7.60×10⁻²</td>
<td>1.53×10⁻²</td>
</tr>
<tr>
<td>CO₂</td>
<td>g/(g-COD-Xa·h)</td>
<td>C</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

^a: Chemical oxygen demand; ^b: growth rate; ^c: poly-β-hydroxybutyrate; ^d: poly(3-hydroxyvalerate); ^e: triacylglycerol; ^f: oxygen uptake rate; ^g: this column indicates whether the compound will be balanced (B), which means that a better estimate will be found and be calculated (C) using the elemental composition matrix in Table 3.2; ^h: not applicable. Note that h = 0.00 against $\chi^2 = 6.63$ at 99% confidence level.
Table 3.6 – Matrix of the conversion rates on acetate as the sole carbon source in the N-limited batch condition

<table>
<thead>
<tr>
<th>Conversion</th>
<th>Units</th>
<th>Calculation</th>
<th>Measurements</th>
<th>Balance Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substrate</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>-6.84×10⁻²</td>
<td>-8.23×10⁻²</td>
</tr>
<tr>
<td>CODa</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>C</td>
<td>NAg</td>
<td>6.05×10⁻³</td>
</tr>
<tr>
<td>Glycogen</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>8.00×10⁻⁴</td>
<td>7.98×10⁻⁴</td>
</tr>
<tr>
<td>PHBb</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>9.50×10⁻³</td>
<td>9.34×10⁻³</td>
</tr>
<tr>
<td>PHVc</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>2.22×10⁻²</td>
<td>2.13×10⁻²</td>
</tr>
<tr>
<td>TAGd</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>B</td>
<td>7.80×10⁻³</td>
<td>7.79×10⁻³</td>
</tr>
<tr>
<td>OURe</td>
<td>g/(g-COD-Xa·h)</td>
<td>B</td>
<td>-4.44×10⁻²</td>
<td>-4.31×10⁻²</td>
</tr>
<tr>
<td>CO₂</td>
<td>g/(g-COD-Xa·h)</td>
<td>C</td>
<td>NA</td>
<td>1.87×10⁻²</td>
</tr>
</tbody>
</table>

a: Chemical oxygen demand; b: poly-β-hydroxybutyrate; c: poly(3-hydroxyvalerate); d: triacylglycerol; e: oxygen uptake rate; f: this column indicates whether the compound will be balanced (B), which means that a better estimate will be found and be calculated (C) using the elemental composition matrix in Table 3.2; g: not applicable. Note that h = 0.91 against χ² = 6.63 at 99% confidence level.

3.3.3 Metabolic Model of Storage Accumulation

Accumulations of storage compounds by RHA1 were examined using the genome-scale metabolic model (iMT1174) and comparing seven objective functions. To do so, a list of unique reactions was generated by eliminating the redundancy in the network due to the large number of isoenzymes (i.e., the model used 1,935 unique reactions, see Additional File 3). The goal was to determine the most suitable objective function to predict the biosynthesis rate of glycogen, PHA, and TAG after the exhaustion of ammonia (NH₃). For the purpose of this paper, PHA was assumed to be a co-polymer of PHB and PHV in the same proportion as experimentally observed (Hernandez et al. 2008).

The total storage production rate (defined as total g-COD/(g-COD-Xa·h)) was found to be very sensitive to the non-growth associated maintenance energy (mATP) for most of the objective functions examined as was expected (Figure 3.4). However, variations in mATP did not affect the observed orders in the yields of storage compounds obtained for all objective
functions. To obtain the proper mATP for each objective function, we interpolated mATP such that the total storage fluxes are the same as those obtained experimentally on glucose and acetate 0.023 and 0.039 g-COD/(g-COD-Xa·h), respectively (Table 3.7). The total storage fluxes for the environmental MoMA sub-objective function were insensitive to mATP, and they remained much lower than the experimentally measured ones. This was due to the generation of futile cycles in the simulation, which were dissipating the carbon and energy. Nonetheless, an mATP value of 0.5 mmol/(g-DW·h) was adopted for the environmental MoMA sub-objective function for comparison purposes. However, this had no impact on the prediction of the environmental MoMA objective function, as it was examined for different mATP.

**Figure 3.4** – Total storage flux as a function of the non-growth-associated maintenance energy on a) glucose and b) acetate.

Objective functions are: MoMA, minimization of fluxes, minimization and maximization of ATP production, minimization of NADH production, maximization of the storage pools, and maximization of the storage pools along with MoMA.
Table 3.7 – Posterior probability share of each objective function, listing in descending order of probability

<table>
<thead>
<tr>
<th>Objective Function</th>
<th>Posterior Probability Share (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximization of storage and MoMA</td>
<td>76.20</td>
</tr>
<tr>
<td>Maximization of storage</td>
<td>21.30</td>
</tr>
<tr>
<td>MoMA</td>
<td>2.37</td>
</tr>
<tr>
<td>Flux minimization</td>
<td>$6.57 \times 10^{-2}$</td>
</tr>
<tr>
<td>Minimization of ATP production</td>
<td>$2.13 \times 10^{-2}$</td>
</tr>
<tr>
<td>Maximization of ATP production</td>
<td>$2.13 \times 10^{-2}$</td>
</tr>
<tr>
<td>Minimization of NADH production</td>
<td>$2.13 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

Note that the experimentally observed mATP were 0.023 and 0.039 g-COD/(g-COD-Xa·h) on glucose and acetate, respectively. The calculated mATP values (in mmol/(g-DW·h)) at the measured total storage flux for different objective functions on glucose are as follows: minFluxes: 2.236, maxATP, minATP, minNADH: 1.200, maxStorage: 3.502, and maxStorage+MoMA: 3.44. The calculated mATP values (in mmol/(g-DW·h)) at the measured total storage flux for different objective functions on acetate are as follows: minFluxes: 3.385, maxATP, minATP, minNADH: 3.409, maxStorage: 4.027, and maxStorage+MoMA: 3.823.

For the seven objective functions compared, two patterns of storage accumulation were simulated during the period of N-limitation and carbon excess; most of the objective functions simulated a single storage compound accumulation, while the environmental MoMA sub-objective functions (MoMA and MaxStorage+MoMA) simulated a mixture of storage compounds accumulation (Figure 3.5b, c, h, i).

For the objective functions that simulated a single storage compound accumulation, only the MaxStorage objective function simulated PHA accumulation when the culture contained glucose and acetate as C-source (Figure 3.5d, j). The other objective function in this group simulated the accumulation of glycogen and PHA, respectively, when glucose and acetate were the C-sources (Figure 3.5e, f, k, l). These objective functions are subsets of the family of objective functions aiming at the minimization of the reaction fluxes. Thus, it seems that this principle promotes early diversion of carbon from the substrate to the storage pools (Figure 3.1), a principle that does not seem to be the one operating for RHA1 when a sudden N-limitation occurs.
Figure 3.5 – Measured and calculated storage yields of three different storage compounds glycogen, PHA and TAG on glucose (a-f) and acetate (g-l).
(a) measured storage yields on glucose, (b) maximization of the storage fluxes + MoMA on glucose, (c) MoMA on glucose, (d) maximization of the storage fluxes on glucose, (e) minimization of the metabolic fluxes on glucose, (f) minimization and maximization of ATP production, and minimization of NADH production, (g) measured storage yield on acetate, (h) maximization of the storage fluxes + MoMA on acetate, (i) MoMA on acetate, (j) maximization of the storage fluxes on acetate, (k) minimization of the metabolic fluxes on acetate, (l) minimization and maximization of ATP production, and minimization of NADH production. Note that the storage yields were identical for minimization and maximization of ATP production and minimization of NADH production; therefore, they were grouped in one graph. 

Using a Bayesian-based objective function discrimination method (Knorr et al. 2007), all the information in Figure 3.5 were combined to yield the most probable objective function. Interestingly, maximization of the storage compounds in the N-limited condition, in conjunction with MoMA was the most probable (with the posterior probability share of 76.2%), with the maximization of storage objective function the second most probable (with the posterior probability share of 21.3%), as shown in Table 3.7. The other objective functions showed posterior probability share in the range of 0.02-2%.

Explicit in our hypotheses, the simulation results presented here could be understood as possible outcomes of global regulatory principles applied by R. jostii RHA1 when faced with sudden nutrient-limited condition. Based on our results, it appears that, on glucose and acetate, the cells try to maintain the reaction flux profiles under carbon-excess/nutrient-limited conditions as similar as possible to the one that they can develop in non-limited growth conditions, yet maximizing their storage accumulation. This seems logical from a global regulation perspective as this solution would simply follow the regulatory program for growth and diverting the excess
carbon found locally in the biochemical network to the closest storage compound. It would also maximize the responsiveness of the organism. In order to adopt the *early carbon diversion* strategy, the cells would probably need to change several points of regulation, which may be evolutionary more complex.

### 3.3.4 Sensitivity Analysis

There are uncertainties associated with the model presented here such as the values of the non-growth- and growth-associated maintenance energies, and the biomass composition, which were derived from different experiments with organisms related to *R. jostii* RHA1 (see Additional File 1). Therefore, we performed a series of sensitivity analyses to investigate the effect of these parameters on the storage metabolism. See section “Metabolic Model of Storage Accumulation” for more details on the effect of maintenance energy on simulation results.

We tested the sensitivity of the accumulation of storage compounds with different biomass composition, by substituting the biomass of *R. jostii* RHA1 used in our model by the *E. coli* biomass. The *E. coli* biomass did not influence the order of the storage compounds for the maximization of storage in conjunction with the *environmental* MoMA objective function on both glucose and acetate. That is, PHA was predicted as the main storage compound, followed by TAG and glycogen. However, the amount of TAG yield on glucose decreased (8.0×10⁻³ as compared to 6.5×10⁻²), which can be attributed to the lower lipid content in the *E. coli* biomass. Moreover, the amount of glycogen yield on glucose decreased by substituting the *E. coli* biomass into the RHA1 model (6.8×10⁻⁴ as compared to 5.6×10⁻³). Nevertheless, the amount of PHA yield on glucose increased by substituting the *E. coli* biomass into the RHA1 model (0.6 as compared to 0.5). The amount of TAG yield on acetate decreased (1.7×10⁻⁴ as compared to
8.8×10^{-2}), which can be attributed to the lower lipid content in the *E. coli* biomass. Moreover, the amount of glycogen yield on acetate decreased by substituting the *E. coli* biomass into the RHA1 model (1.4×10^{-5} as compared to 7.4×10^{-3}). Nevertheless, the amount of PHA yield on acetate increased by substituting the *E. coli* biomass into the RHA1 model (0.5 as compared to 0.4). It seems that the prediction depends on the biomass composition as it was found in other studies (Borodina et al. 2005).

Storage accumulation did not occur when the *E. coli* biomass was applied in the RHA1 model on acetate for flux minimization, while the order of storage did not change and glycogen turned out to be the main storage compound on glucose. Note that the predicted storage compounds for the *E. coli* biomass on glucose were glycogen and PHA with the yields of 0.6 g-COD/g-COD and 0.1 g-COD/g-COD, respectively. However, the *E. coli* biomass increased the yields of glycogen and PHA on glucose for flux minimization as compared to the RHA1 biomass (Figure 3.5e).

It should be noted that the simulation results reported here were obtained by defining PHA as the specific copolymer mixture of PHB and PHV observed experimentally on each substrate, with PHV as the most abundant polymer (Hernandez et al. 2008). To see what would be the natural mixture of PHB and PHV for our model, a number of simulations were conducted in which these two compounds were accumulated independently. PHB was typically the most abundant polymer of the two on glucose and acetate when using the maxStorage+MoMA sub-objective function. Two pathways are possible for the synthesis of PHV: a long pathway from acetyl-coA via the valine degradation pathway, and a short pathway from succinyl-coA via methylmalonyl-coA (Figure 3.1). These simulation results were closely identical regardless of the pathways selected (see Figure 3.6 for the active short biosynthetic pathway of PHV). These
results further strengthen the conclusion that \textit{R. jostii} RHA1 tries to maintain its reaction flux profile unchanged as simulated by the maxStorage+MoMA sub-objective function.

![Graph](image)

**Figure 3.6** – Calculated storage yields of four different storage compounds glycogen, PHB, PHV, and TAG accumulated by the \textit{R. jostii} RHA1 model on glucose (a) and acetate (b).

In this simulation, the objective function was maxStorage+MoMA; note that PHB and PHV were accumulated independently. Here, the short biosynthetic pathway of PHV was activated and the long one was deactivated.

### 3.3.5 Application of MoMA in the Environmental Context

The MoMA sub-objective function was introduced to the flux balance analysis to compare the wild-type and mutant organisms (Segre et al. 2002); however, in this study we extended the application of this objective function to compare balanced fluxes of metabolisms operating two different but successive environmental conditions: non-limited and N-limited conditions. One may wonder if this is an appropriate comparison. We applied $^{13}$C-metabolic flux analysis ($^{13}$C-MFA) to simulate the metabolic behaviour of the cell in the non-limited condition.
in order to justify our hypothesis; and then we compared the reactions in the central metabolic pathways predicted by two objective functions including maximization of storage in conjunction with MoMA and flux minimization with those obtained with $^{13}$C-MFA on glucose and acetate. Additional Files 6-1, 6-2, and 6-3 show 100 solutions of the central metabolic fluxes, as well as their corresponding errors, calculated using $^{13}$C-MFA on 1-$^{13}$C-glucose, 1-$^{13}$C-acetate, and 2-$^{13}$C-acetate, along with average, standard deviation, minimum, maximum, and coefficient of correlation.

Figure 3.7 compares the central metabolic fluxes obtained from FBA in the N-limited condition (at zero nitrogen uptake rate) for different objective functions with those obtained from $^{13}$C-MFA in the non-limited condition on different substrates. It includes the minimum and maximum fluxes of 22 central metabolic reactions that were calculated in the N-limited condition at zero nitrogen uptake rate using FBA on glucose and acetate, along with the average fluxes and a set of fluxes possessing the minimum difference between the measured and calculated mass isotopomer distribution data, as well as their respective minimum and maximum values, obtained from $^{13}$C-MFA on 1-$^{13}$C-glucose and 1-$^{13}$C-acetate. Here, the range of the fluxes in the N-limited condition is within the range of the fluxes in the non-limited condition on both substrates. Note that the calculated data on 2-$^{13}$C-acetate showed similar trends as those of 1-$^{13}$C-acetate (data not shown). In addition, the results are fairly identical regardless of the choice of objective functions used (i.e., maxStorage+MoMA vs. minFluxes). Therefore, it appears that *R. jostii* RHA1 essentially attempts to maintain its reaction flux profile during N-limited storage accumulating condition most similar to an immediately preceding period when its growth is not limited, while maximizing its storage production to efficiently save the carbon source; this substantiated the use of MoMA in the environmental context.
Figure 3.7 – Comparison of the fluxes of 22 central metabolic reactions of *R. jostii* RHA1 estimated using the FBA and $^{13}$C-MFA in N- and non-limited conditions, respectively, on a) glucose and b) acetate.

In the non-limited condition fluxes were calculated using the $^{13}$C-MFA approach, while in the N-limited conditions the fluxes were computed using the genome-scale FBA at zero nitrogen uptake rate. Note that the best estimated fluxes were that solutions cluster possessing the lowest sum of the absolute errors between the measured and calculated mass isotopomer distributions (MIDs) of the labelled amino acids among 100 solutions. Note also that the average flux in non-limited condition is the average over 100 solutions estimated by the $^{13}$C-MFA. The reactions in parentheses are the backward reactions of the reversible reactions. Note that the range of the fluxes in the N-limited condition is within that of the non-limited condition.

3.4 Conclusions

*R. jostii* RHA1 can accumulate several storage compounds simultaneously, but the global cellular regulation of the synthesis of these compounds during nutrient-limited condition remain to be clarified. A genome-scale metabolic model of *R. jostii* RHA1 (iMT1174) was built, and
used in flux balance analysis with different objective functions to predict storage compound fluxes during a period of N-limitation following a period of balanced maximal growth. Among the seven objective functions examined, the objective function maximization of storage in the N-limited condition combined with MoMA was able to reproduce best the storage accumulation fluxes observed experimentally for glucose and acetate as C-source. To further substantiate the use of MoMA sub-objective function, the variability of fluxes through 22 central metabolic reactions during N-limited storage-accumulating conditions were compared to fluxes through the same reactions during non-limited growth obtained by $^{13}$C-MFA. The reaction fluxes during storage accumulation were well within the possible range determined by $^{13}$C-MFA for both substrates tested. Therefore, it appears that *R. jostii* RHA1 mainly attempts to maintain its reaction flux profile during N-limited storage accumulating condition most similar to an immediately preceding period when its growth is not limited, while maximizing its storage production to efficiently save the C-source. Finally, the data presented demonstrate the applicability of the MoMA sub-objective function to analyze rapid changes in successive different environmental conditions such as non-limited growth followed by N-limited storage-accumulating substrate consumption.

**Additional Files**

**Additional File 1 – Biomass and storage lipid compositions (Ch3_AF1.xls):**

It includes the compositions of the RHA1 biomass, macromolecules, PHA, TAG, and WE used in this study.

**Additional File 2 – Total reactions and their corresponding biochemical pathways, along with exchange reactions, that are involved in the iMT1174 network (Ch3_AF2.xlsx):**
It contains reaction properties such as name, ID, definition, equation, directionality, lower and upper bounds, enzyme commission (EC) numbers, open reading frames (ORFs), KEGG IDs of reactions that were replaced by added reactions (replaced reactions), and references.

Additional File 3 – Unique reactions and metabolites, along with their properties, that are involved in the iMT1174 network (Ch3_AF3.xlsx):

It includes unique reactions (Additional File 3-1) and unique metabolites (Additional File 3-2). These data were utilized to make the stoichiometric matrix of the iMT1174 metabolic network. Note also that the names, IDs and chemical formulas of the metabolites are reported.

Additional File 4 – Central metabolic pathways of R. jostii RHA1 (Ch3_AF4.pdf).

Additional File 5 – Measured and calculated mass isotopomer fractions of amino acids along with their errors on different substrates (Ch3_AF5.xlsx):

It includes measured and calculated mass isotopomer fractions of amino acids along with their errors and the calculated mass isotopomer of a solution set that had the lowest difference between the measured and calculated mass isotopomer fractions on 1-13C-glucose (Additional File 5-1), 1-13C-acetate (Additional File 5-2), and 2-13C-acetate (Additional File 5-3).

Additional File 6-1 – The central metabolic fluxes, as well as their corresponding errors, calculated using 13C-MFA on 1-13C-glucose along with average, standard deviation, minimum, maximum, and coefficient of correlation (Ch3_AF6-1.xlsx):

Additional File 6-1-1 shows 100 solution sets of the central metabolic fluxes, as well as their corresponding errors, calculated using 13C-MFA on 1-13C-glucose; Additional File 6-1-2 shows average, standard deviation, minimum, and maximum of the central metabolic fluxes calculated using 13C-MFA on 1-13C-glucose; Additional File 6-1-3 shows coefficient of correlation of the central metabolic fluxes calculated using 13C-MFA on 1-13C-glucose.
Additional File 6-2 – The central metabolic fluxes, as well as their corresponding errors, calculated using $^{13}$C-MFA on 1-$^{13}$C-acetate along with average, standard deviation, minimum, maximum, and coefficient of correlation (Ch3_AF6-2.xlsx):

Additional File 6-2-1 shows 100 solution sets of the central metabolic fluxes, as well as their corresponding errors, calculated using $^{13}$C-MFA on 1-$^{13}$C-acetate; Additional File 6-2-2 shows average, standard deviation, minimum, and maximum of the central metabolic fluxes calculated using $^{13}$C-MFA on 1-$^{13}$C-acetate; Additional File 6-2-3 shows coefficient of correlation of the central metabolic fluxes calculated using $^{13}$C-MFA on 1-$^{13}$C-acetate.

Additional File 6-3 – The central metabolic fluxes, as well as their corresponding errors, calculated using $^{13}$C-MFA on 2-$^{13}$C-acetate along with average, standard deviation, minimum, maximum, and coefficient of correlation (Ch3_AF6-3.xlsx):

Additional File 6-3-1 shows 100 solution sets of the central metabolic fluxes, as well as their corresponding errors, calculated using $^{13}$C-MFA on 2-$^{13}$C-acetate; Additional File 6-3-2 shows average, standard deviation, minimum, and maximum of the central metabolic fluxes calculated using $^{13}$C-MFA on 2-$^{13}$C-acetate; Additional File 6-3-3 shows coefficient of correlation of the central metabolic fluxes calculated using $^{13}$C-MFA on 2-$^{13}$C-acetate.

Authors’ Contributions

MT carried out the reconstruction of $R. jostii$ RHA1 (iMT1174), manually curated the metabolic network, developed various computer codes to perform the flux balance analysis and $^{13}$C-MFA and the post processing of data, analyzed and validated the models, interpreted data, and drafted the manuscript. DF conceived of the study, participated in its design and coordination and defined this manuscript, analyzed and validated the reconstruction of $R. jostii$
RHA1, interpreted data, and contributed to the writing of this manuscript. Both authors read and approved the final manuscript.

Acknowledgments
This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant awarded to DF. MT would like to thank the McGill Engineering Doctoral Award (MEDA) for the financial support of his Ph.D. study. Labelled amino acid measurements were performed at the Rosalind and Morris Goodman Cancer Research Centre Metabolomics Core Facility supported by the Canada Foundation of Innovation, The Dr. John R. and Clara M. Fraser Memorial Trust, the Terry Fox Foundation (TFF Oncometabolism Team Grant 116128), and McGill University.

3.5 References


RHA1 provides insights into a catabolic powerhouse. Proceedings of the National Academy of Sciences 103(42):15582-15587.


In chapter 3 the genome-scale metabolic model of *R. jostii* RHA1 was reconstructed. The *in silico* growth of *R. jostii* RHA1 in non- and N-limited conditions were modelled using the different objective functions showing that the maxStorage+MoMA sub-objective function could predict a mixture of the storage compounds (glycogen, PHA, and TAG) with PHA as the main storage pool on both glucose and acetate. The $^{13}$C-MFA results also substantiated the use of the MoMA objective function in the environmental context. In Chapters 4 and 5, the use of similar objective functions is investigated for growth under feast and famine cycles. There are no experimental data on the feast-famine growth of *R. jostii* RHA1 in the literature. Therefore, chapter 4 first reports on experimental observations of *R. jostii* RHA1 growth during the feast-famine cycles on glucose and acetate as the sole carbon sources.
Chapter 4. Accumulation of Storage Compounds by Rhodococcus jostii RHA1 during Feast-Famine Cycles

Mohammad Tajparast\textsuperscript{1} and Dominic Frigon\textsuperscript{1*}

\textsuperscript{1}Microbial Community Engineering Laboratory, Department of Civil Engineering and Applied Mechanics, McGill University, 817 Sherbrooke Street West, Montreal, Quebec, Canada, H3A 2K6

*Corresponding author

Email addresses:

MT: mohammad.tajparast@mail.mcgill.ca

DF: dominic.frigon@mcgill.ca

Submitted to

Journal of Applied Microbiology
Abstract

Feast-famine cycles in conventional activated sludge treatment systems select for bacterial species that accumulate intracellular storage compounds such as glycogen, poly-β-hydroxybutyrate (PHB), and triacylglycerols (TAG). These species survive better the famine phase and resume rapid substrate uptake at the beginning of the feast phase faster than microorganisms unable to accumulate storage. Ecophysiological conditions favouring the selection of species accumulating either storage compounds remain to be clarified. Example questions: Are substrates present in the wastewater the main factor determining this selection? Are there metabolic organizations and strategies deployed by certain species that can explain this selection? These questions were investigated by subjecting Rhodococcus jostii RHA1 to the feast-famine growth cycles on glucose and acetate as the sole carbon sources.

The feast phase lasted 136.7±6.4 and 145.7±0.0 min on glucose and acetate, respectively. When glucose was the substrate, the COD accumulation during the feast phase was similar for the three storage compounds: PHB, TAG, and glycogen. However, when acetate was the substrate, essentially no glycogen was accumulated and the COD flux (and yield) towards PHB accumulation was 3 times higher than toward TAG accumulation. Finally, on both substrates, hexadecanoic acid (Palmitic acid, C16:0) was the predominant fatty acid present in TAG followed by cis-octadecenoic acid (C18:1).

Keywords
Feast-famine cycle, storage metabolism, Rhodococcus jostii RHA1, glycogen, poly-β-hydroxybutyrate, triacylglycerol
4.1 Introduction

Natural environments are often highly dynamic with regard to availability of external electron donors. For example, biomass in biological wastewater treatment systems is submitted to cycles of feast and famine with the feast phase being the period of availability of external electron donors and the famine phase being the starvation period (Van Loosdrecht et al. 1997). Similar situations occur in the soil rhizosphere (Gasser et al. 2009) and in marine environments (Srinivasan and Kjelleberg 1998). Such cycles profoundly affect the microbial ecology and the metabolism of the bacterial species present in these environments. One of the main metabolisms found under feast-famine cycling is the synthesis (during feast) and consumption (during famine) of storage compounds such as glycogen and poly-β-hydroxybutyrate (Dawes and Senior 1973). Bacterial species accumulating storage compounds during the feast phase outperform the species lacking this ability probably by maintaining the capacity to rapidly respond to a substrate addition such that the metabolic cost of storage in terms of yield is lower than the competition cost to rapidly gain access to the substrate (Frigon et al. 2006). Therefore, understanding and properly modeling the metabolism of the storage compounds is central to predict the population dynamics in natural habitats.

In the current study, the storage metabolism of *Rhodococcus jostii* RHA1 during the feast-famine growth cycle was studied. *R. jostii* RHA1 was isolated from a soil contaminated with lindane and assigned to the genus *Rhodococcus* (Seto et al. 1995). The genus *Rhodococcus* is important for both environmental and biotechnological applications because of its extraordinary capacity for metabolizing recalcitrant organic compounds (Larkin et al. 2005). Several representatives of this genus have been isolated from soil, marine environment and activated sludge wastewater treatment systems (Goodfellow et al. 1998), making it an interesting model heterotroph to study. One of the specific interests of the strain *R. jostii* RHA1 is that it is
an environmental isolate that has recently been sequenced (McLeod et al. 2006), and its in silico genome-scale metabolic network has recently been reconstructed in our research group (Chapter 3, (Tajparast and Frigon 2014)).

The frequent occurrence of Rhodococcus species in a variety of dynamic environments around the world may reflect their capacity to adapt. Like other microorganisms in environments where transient nutrient-limitation is common, Rhodococcus spp. can accumulate storage compounds for subsequent utilization by cells as endogenous carbon sources and electron donors during periods of nutritional scarcity (Hernandez et al. 2008). Of special interest for the current study, R. jostii RHA1 is able to accumulate most of the storage compounds known as: poly-β-hydroxybutyrate (PHB), poly(3-hydroxyvalerate) (PHV), triacylglycerols (TAG), wax esters (WE), α,α-trehalose, glycogen, and polyphosphate (PolyP) (Hernandez et al. 2008).

Organic storage polymers are generally divided into two major types of metabolites: polysaccharides and lipids (including polyhydroxyalkanoates and triacylglycerols) (Dawes and Senior 1973). The different storage polymers originate from different substrates. Glycogen formation is often selected for when the primary substrate is glucose or a substrate that can be converted into pyruvate with an increase in reducing power; for example, different carbohydrates, glycerol, or proteins.

Polyhydroxyalkanoates (PHA) are selected for when the substrates available are rapidly transformed to acyl-coenzyme A, such as PHB from acetyl-coA and PHV from propionyl-coA (Verlinden et al. 2007). PHA are polymers of hydroxy fatty acids with the most abundant PHA being PHB. PHA created interest as bioplastics as they can exhibit thermoplastic and elastomeric properties, are enantiomerically pure chemicals (only R-stereoisomer), nontoxic, biocompatible, made from renewable resources, and biodegradable. Moreover, the polymer can be hydrolyzed
and the hydroxy fatty acid monomers could serve as chiral building block chemicals for the production of biochemicals (Johnson et al. 2009).

Only a few genera, including the genus *Rhodococcus*, produce TAG during the consumption of different carbon sources (Hernandez et al. 2008). For example, *Rhodococcus opacus* PD630 accumulates TAG comprising up to 76% of its cellular dry weight after growth on gluconate (Alvarez et al. 1996; Waltermann et al. 2000). Biodiesel (monoalkyl esters of long-chain fatty acids with short-chain alcohols derived from triacylglycerols) can be generated from renewable biomass resources (Kurosawa et al. 2010; Rottig et al. 2010).

From the description above, it is clear that understanding the storage metabolism in organisms like *R. jostii* RHA1 has an interest both to understand its ecophysiology and to advance possible biotechnological applications. Thus, the objective of the current study was to investigate storage metabolism of *R. jostii* RHA1 during feast-famine growth condition on glucose and acetate as the sole carbon sources. The reactor conditions selected were similar to activated sludge wastewater treatment systems with approximately 6 h between nutrient additions, and solids and hydraulic retention times both equal to 1 day.

### 4.2 Materials and Methods

#### 4.2.1 Bacterial Strain and Growth Medium

*Rhodococcus jostii* strain RHA1 was grown on the medium supplemented with a certain amount of carbon source (e.g., α-D-glucose or acetate) as described below (Beun et al. 2000; Vishniac and Santer 1957). The purity check of the cultures was done by plating them on mineral salt agar plates supplemented with biphenyl (provided as crystals on the lid of the Petri dishes) (Atlas and Bartha 1998) or by serial dilution plating on nutrient agar (Frigon et al. 2006; Van Aalst-Van Leeuwen et al. 1997). Incubations were at 28 °C.
The mineral salt medium (MSM) was composed of NH₄Cl (6.00 mM), KH₂PO₄ (1.65 mM), MgSO₄.7H₂O (1.11 mM), KCl (1.50 mM), and trace element solution (1 mL/L). The agar plates were made by adding agar, 1.5 % (w/v), into the media. We also added glucose and acetate (in some plates we added crystals of biphenyl on the lids of the plates) to the final concentration of 10.50 and 31.50 mM, respectively, as the sole carbon sources, which are equivalent to 504 mg-COD/L (Beun et al. 2000). The trace element solution contained: Ethylenediamine tetra-acetic acid (EDTA: C₁₀H₁₆N₂O₈) (50.00 g/L), ZnSO₄.7H₂O (22.00 g/L), CaCl₂ (5.54 g/L), MnCl₂.4H₂O (5.06 g/L), FeSO₄.7H₂O (4.99 g/L), (NH₄)₆Mo₇O₂₄.4H₂O (1.10 g/L), CuSO₄.5H₂O (1.57 g/L), CoCl₂.6H₂O (1.61 g/L) dissolved in 1000 mL dilution water; adjusted to pH 6.0 with KOH (Vishniac and Santer 1957). The dilution water was supplemented with 0.17 M MgSO₄.7H₂O in excess to promote biomass flocculation.

4.2.2 Reactor Condition
A 2-L sequencing batch reactor (SBR) was employed to simulate dynamic conditions as occur during the feast and famine periods. The SBR was inoculated with a pure culture R. jostii RHA1, which was grown to late exponential phase in 250 mL solution bottles containing the above-mentioned MSM solution supplemented with either glucose or acetate (depending on the reactor feed) at 28 °C and 175 rpm. The hydraulic residence time (HRT) and the solids residence time (SRT) of 24 h were maintained in the SBR. Continuous operation of the SBR was based on a 6 h- feast-famine cycle comprising of 11 min start period followed by 2 min dilution water addition (464 mL) and 4 min influent phase adding 36 mL fresh medium, reaction phase (308 min), 30 min settling period, and 5 min effluent withdrawal period during which 500 mL of mixed reactor liquor was wasted. The reactor content needs to be well mixed and aerated to supply enough dissolved oxygen; the dissolved oxygen (DO) concentration during the steady-
state conditions were 80-90%. Few drops of anti-foam were also added to the reactor to prevent foaming and enhance the aeration. Mixing and aeration were only performed for the start, dilution water and influent addition, and reaction phases (Beun et al. 2000; Johnson et al. 2010).

The stirrer speed was maintained at 300 rpm throughout the cycle. The inlet air flow rate was controlled with a mass flow controller for air at the fixed value of 1.5 L/min. The pH was also set at 7.0 using 1 M KOH and 1 M HCl solutions. The stirrer, the pH, the airflow, and the pumps for feeding and effluent withdrawal were automatically controlled. The temperature of the reactor was fixed at 28 °C by means of a water jacket.

The steady-state operation of the SBR was achieved when for a few days the length of the feast periods and the dissolved oxygen (DO) time profiles throughout the cycles remained the same. At steady state, the SBR was observed (for pH and DO) and sampled (for residual carbon source, chemical oxygen demand (COD), NH$_4^+$, phosphorus, glycogen, PHB, TAG, total suspended solids (TSS), volatile suspended solids (VSS), and ash content) for several time steps.

**4.2.3 Analytical Procedures**

Samples were obtained throughout a cycle to monitor changes in the chemical composition of the culture. Samples of 5 mL (triplicate) were rapidly filtered through a 0.22-µm filter in order to measure TSS (filter dried for 24 h at 105 °C) and VSS/ash content (dry filter in an oven for 2 h at 550 °C) according to the standard method 2540 D/2540 E (Clesceri et al. 1998). The filtrates were stored at −20 °C for further analysis of COD according to the standard method 5220 D (Clesceri et al. 1998), substrate (either glucose or acetate), spectrophotometric analysis of NH$_4^+$ at 630 nm (Rhine et al. 1998), and phosphorus (the Ascorbic Acid method, standard methods 4500-P E) (Clesceri et al. 1998). A few drops of formaldehyde were added to the rest of the remaining volume of the samples to stop all biological activities, and biomass
were centrifuged for analysis of glycogen (10 mL subsamples), PHB (20 mL subsamples), and TAG (20 mL subsamples). The harvested biomass samples were stored at −80 °C until analysis.

The acetate concentration in the filtrate was determined enzymatically according to the manufacturer (Megazyme International Ireland Ltd., Bray Business Park, Bray, Co. Wicklow, IRELAND). Glucose was quantified by a hexokinase enzymatic kit and colorimetry (Hexokinase protocol, measurement of NADH concentration at 340 nm, Sigma, St. Louis, MO) (Maurer et al. 1997).

Glycogen was measured according to Maurer et al. (1997) with some modifications. The biomass subsamples were resuspended in 1.5 mL of sterile deionized water, vortexed, and homogenized using ultrasonic treatment for 15 min. Further, 40 µL of 6 M HCl solution was added into 0.9 mL glass vials, that had already been washed with ethanol and further with concentrated HCl solution, dried in a dry-bath, and pre-weighed, and they were weighed. Moreover, 400 µL of each sample was collected into the vials and weighed. The vials were then closed and incubated in a dry-bath at 95 °C for 2 h. To determine the exact volume of the sample (if evaporation occurred), the vials were reweighed. The mixture was further neutralized with 33 µL of 10 M KOH and 57 µL of 0.9 M phosphate buffer (these volumes are varying, depending on the samples; i.e., the vials were weighed after neutralizing pH). After centrifugation of the solution at 6000 rpm for 10 min, the glucose concentration in the supernatant was measured by a hexokinase enzymatic kit (Hexokinase protocol, measurement of NADH concentration at 340 nm (UV range), Sigma, St. Louis, MO) as follows: First, 15 µL of the sample was added into 150 µL of the reagent in a 0.6 mL micro-centrifuge tube (that had already been washed with ethanol and dried) and mixed for 15 min at room temperature (18-35 °C). Then, 40 µL of the mixture
was loaded into the micro-plate (triplicate) and the absorbance was measured against the deionized water at 340 nm.

The amount of PHB was determined, according to (Law and Slepecky 1961; Paganelli et al. 2011), with some modifications. The biomass subsamples were resuspended in 1.5 mL of sterile deionized water and prepared the same way as the above-mentioned procedure for the glycogen analysis. The empty ethanol- and acid-washed 0.9 mL glass vials were weighed, 0.2 mL of 2 N HCl was further added, and the vials were reweighed; further, 0.1 mL of the biomass suspension was added into the vial and reweighed; then, the mixture was incubated at 65 °C for 2 h. The vials were centrifuged at 6000 rpm for 20 min, and then weighed. To obtain a precipitate, 0.3 mL of chloroform was added, and then the samples were agitated at 140g for 2 h at 29 °C, and the vials were weighed. The vials were centrifuged at 6000 rpm for 30 min. After centrifugation, 100 µL of chloroform extracts were transferred into the pre-weighed 0.9 mL glass vials, dried at room temperature (overnight), and the vials were weighed. After drying, 0.3 mL of concentrated sulfuric acid was added, the vials were reweighed, and the samples were incubated at 95 °C in a dry-bath incubator for 20 min. After cooling at room temperature, the vials were weighed, and the amount of PHB was further determined on a spectrophotometer at 235 nm (UV range) by loading 70 µL of the samples into a micro-plate. Concentrated sulfuric acid was used as the blank. All samples were analyzed in triplicate. For the conversion of the absorbance values to milligrams of PHB, a standard curve was constructed relating optical density 235 nm to the amount in mg of a PHB standard (DL-3-Hydroxybutiric acid, SIGMA, St Louis, MO, USA). Note that a microplate reader (model SpectraMax5) was employed to perform the spectrophotometric measurements (Molecular Devices, LLC, USA).
For the identification of lipids, cultures were harvested by centrifugation at 6000g for 15 min and lyophilisation overnight. Further, 5-8 mg of lyophilized cells was resuspended in 0.6 mL of chloroform and 0.6 mL of methanol containing 15% (v/v) H$_2$SO$_4$; note that 1 µL Trinonadecanoin (Nu-Check Prep Inc., Elysia, MN) was used as internal standard for quantification (Frigon 2005). Methanolsysis was carried out at 100 ºC for 2.5 h. After cooling to room temperature and then on ice, 0.3 mL of deionized water was added to the solution, which was then vigorously vortexed for 1 min. After phase separation, 0.45 mL of the organic phase (bottom layer) was removed and transferred to a small screw-cap glass vial (Brandl et al. 1988). The organic phase containing fatty acid methyl esters (FAMEs) was analyzed by using an Agilent 6890N GC system equipped with an Agilent HP-88 column (60 m by 0.25 mm, 0.2 µm thick film) with helium as the carrier gas at flow rate of 2 mL/min. A 1 µL portion of the organic phase was injected with a 50:1 split ratio using the auto-sampler. The inlet was maintained at 250 ºC. The oven was held at 175 ºC for 15 min, heated to 220 ºC at 3 ºC/min, and then held at 220 ºC for 5 min. Peak detection was performed by a flame ionization detector, which was maintained at 280 ºC. The fatty acids were identified and quantified by comparison to standard FAMEs (Sigma). Fatty acid content was defined as the percentage of the ratio of fatty acids to cell dry weight (% CDW).

4.2.4 Statistical Analyses

After measuring the above-mentioned parameters, elemental mass balances on the measured conversions can be performed to check the consistency of the data. For every substrate an elemental composition matrix was defined (Table 4.1). The elemental composition matrix contained the balances over one cycle for the elements carbon (C) and nitrogen (N) for both the feast and famine periods, the balances for the amount of COD for both the feast and famine
periods, and the total balances for glycogen, PHB, TAG, and active biomass (Xa). However, there were more conversions measured than needed to define the whole system with elemental balances. The $\chi^2$ test was used to evaluate the correlation between the measured and estimated rates at 99% confidence level (corresponding to $\alpha$ level of 0.01). $h$ is the measured statistics that is $\chi^2$-distributed with a number of degrees of freedom equal to the rank of the covariance matrix of the residuals. An error is not significant if $h < \chi^2$. This also accepts the null hypothesis stating that there is no systematic error for the appropriate number of degrees of freedom and the desired confidence level in the model and there is no difference between the measured and predicted conversion rates (Van Der Heijen et al. 1994). Macrobal software (Hellinga, 1992) was employed to balance all the converted amounts and calculate errors. Macrobal can find the best estimate for all measured data, based on elemental mass balancing principles. Using Macrobal it was also possible to define the feast and famine period separately in terms of converted compounds. The calculated $\chi^2$ and $h$ values in different conditions were displayed in the footnotes of Tables 4.3 and 4.4.
Table 4.1 – Macroal elemental composition matrix

<table>
<thead>
<tr>
<th>Conversion</th>
<th>Feast Balance</th>
<th>Famine Balance</th>
<th>Feast-Famine Equality</th>
<th>Overall Biomass</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C&lt;sup&gt;d&lt;/sup&gt;</td>
<td>N&lt;sup&gt;e&lt;/sup&gt;</td>
<td>COD&lt;sup&gt;f&lt;/sup&gt;</td>
<td>C</td>
</tr>
<tr>
<td><strong>Feast</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substrate</td>
<td>0.375</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NH&lt;sub&gt;4&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt;</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Glycogen</td>
<td>0.375</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PHB&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.333</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TAG&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.266</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Biomass</td>
<td>0.320</td>
<td>0.073</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>O&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>CO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Famine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH&lt;sub&gt;4&lt;/sub&gt;&lt;sup&gt;+&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Glycogen</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.375</td>
</tr>
<tr>
<td>PHB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.333</td>
</tr>
<tr>
<td>TAG</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.266</td>
</tr>
<tr>
<td>Biomass</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.320</td>
</tr>
<tr>
<td>O&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Overall Biomass</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<sup>a</sup>: poly-β-hydroxybutyrate; <sup>b</sup>: triacylglycerol; <sup>c</sup>: total amount of active biomass produced in one feast-famine cycle; <sup>d</sup>: carbon; <sup>e</sup>: nitrogen; <sup>f</sup>: chemical oxygen demand.

4.3 Results

The cultures generally reached steady-state five days after inoculation, which was defined as a recurrent DO profile in each cycle. The profiles were characterized by a rapid decrease in DO at the beginning of the cycle and rapid increase that corresponded to the exhaustion of the carbon substrate (i.e., the end of the feast phase) (Figures 4.1 and 4.2). The DO remained relatively high throughout the famine phase. Due to differences in biomass concentrations and activities, the feast phase lasted 136.7±6.4 and 145.7±0.0 min on glucose and acetate, respectively, which represented about 38% and 40% of the total cycle.

During the feast phase, linear decreases in glucose or acetate concentrations were observed, indicating a zero-order rate for substrate uptake (Figures 4.1 and 4.2). At the same
time, the concentrations of glycogen, PHB, and TAG increased; however, the accumulation profile was dependent on the carbon substrate. While the three storage compounds were accumulated in similar fashions when glucose was the substrate (Figure 4.1), almost no glycogen accumulation was observed on acetate (Figure 4.2).

**Figure 4.1** – Analysis of the dynamics of *R. jostii* RHA1 during a steady-state 6-hour cycle of the SBR fed glucose.

Chemical oxygen demand (COD), glucose, NH$_4^+$, glycogen, PHB, TAG, and dissolved oxygen concentration time profile during a steady-state feast-famine cycle in the glucose fed SBR process at a SRT of 1 day. The data are average of triplicate runs except for dissolved oxygen.
Figure 4.2 – Analysis of the dynamics of *R. jostii* RHA1 during a steady-state 6-hour cycle of the SBR fed acetate.

Chemical oxygen demand (COD), acetate, NH$_4^+$, glycogen, PHB, TAG, and dissolved oxygen concentration time profile during a steady-state feast-famine cycle in the acetate fed SBR process at a SRT of 1 day. The data are average of duplicate runs except for COD and dissolved oxygen.

The fatty acid content of *R. jostii* RHA1 varied slightly between growth on glucose and acetate (Table 4.2). On both substrates, hexadecanoic acid (Palmitic acid, C16:0) was the predominant fatty acid present in TAG followed by cis-octadecenoic acid (C$_{18:1}$); which is consistent with data available for other members of the genus *Rhodococcus* (Alvarez et al. 1997; Alvarez et al. 2000). However, hexadecanoic acid and octadecanoic acid (C$_{18:0}$) concentrations may have been a little higher on glucose than on acetate, and the cis-octadecenoic acid concentrations were a little lower on glucose.
Table 4.2 – Fatty acid content of *R. jostii* RHA1 at the beginning and end of the feast phase on glucose and acetate

<table>
<thead>
<tr>
<th>Fatty Acid</th>
<th>Units</th>
<th>Glucose</th>
<th>Acetate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Beginning</td>
<td>Feast</td>
</tr>
<tr>
<td><strong>C14:0</strong></td>
<td>%g-COD/g-COD-TAG</td>
<td>2.1±1.1</td>
<td>3.0±0.2</td>
</tr>
<tr>
<td><strong>C15:0</strong></td>
<td>%g-COD/g-COD-TAG</td>
<td>3.9±1.9</td>
<td>5.7±0.2</td>
</tr>
<tr>
<td><strong>C16:0</strong></td>
<td>%g-COD/g-COD-TAG</td>
<td>66.7±16.7</td>
<td>48.4±2.4</td>
</tr>
<tr>
<td><strong>C17:0</strong></td>
<td>%g-COD/g-COD-TAG</td>
<td>5.2±2.6</td>
<td>7.8±0.1</td>
</tr>
<tr>
<td><strong>C18:0</strong></td>
<td>%g-COD/g-COD-TAG</td>
<td>7.5±4.7</td>
<td>13.2±3.4</td>
</tr>
<tr>
<td><strong>C18:1 Trans</strong></td>
<td>%g-COD/g-COD-TAG</td>
<td>0.0±0.0</td>
<td>0.0±0.0</td>
</tr>
<tr>
<td><strong>C18:1 Cis</strong></td>
<td>%g-COD/g-COD-TAG</td>
<td>14.6±7.5</td>
<td>21.5±0.9</td>
</tr>
<tr>
<td><strong>C18:2 Trans</strong></td>
<td>%g-COD/g-COD-TAG</td>
<td>0.0±0.0</td>
<td>0.3±0.3</td>
</tr>
<tr>
<td><strong>Total TAGs</strong></td>
<td>%g-COD/g-COD-Xa</td>
<td>2.6±1.3</td>
<td>4.2±0.6</td>
</tr>
</tbody>
</table>

Note that the data are average of triplicate for glucose and duplicate for acetate.

Given the amount of information obtained on the cultures (substrate [glucose or acetate], biomass, glycogen, PHB, TAG, NH₄⁺, CO₂, and DO), it was possible to do elemental mass balances on the measured conversions to check the consistency of the data (Tables 4.3 and 4.4). In addition to data consistency, the biomass production during the feast and famine phases could be estimated. These data could not be estimated directly from VSS and storage or NH₄⁺ concentrations because of the small changes in concentrations compared to the background. For both glucose and acetate, it was estimated that approximately 60% of the biomass was synthesized during the feast phase and that 40% was synthesized in the famine phase (Tables 4.3 and 4.4).
### Table 4.3 – Matrix containing the converted amounts on glucose

<table>
<thead>
<tr>
<th>Conversion</th>
<th>Units</th>
<th>Calculation&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Measurements</th>
<th>Balance Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conversions</td>
<td>Error</td>
</tr>
<tr>
<td><strong>Feast</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Average Biomass</strong></td>
<td>mg-COD/L</td>
<td>B</td>
<td>2331.7</td>
<td>55.0</td>
</tr>
<tr>
<td><strong>Average ASH</strong></td>
<td>mg/L</td>
<td></td>
<td>636.9</td>
<td>30.6</td>
</tr>
<tr>
<td><strong>Substrate</strong></td>
<td>mg-COD/cycle</td>
<td>B</td>
<td>−956.3</td>
<td>20.8</td>
</tr>
<tr>
<td><strong>NH₄⁺</strong></td>
<td>mg-N/cycle</td>
<td>B</td>
<td>−37.8</td>
<td>13.5</td>
</tr>
<tr>
<td><strong>Glycogen</strong></td>
<td>mg-COD/cycle</td>
<td>B</td>
<td>89.4</td>
<td>24.1</td>
</tr>
<tr>
<td><strong>PHB&lt;sup&gt;a&lt;/sup&gt;</strong></td>
<td>mg-COD/cycle</td>
<td>B</td>
<td>113.7</td>
<td>30.6</td>
</tr>
<tr>
<td><strong>TAG&lt;sup&gt;b&lt;/sup&gt;</strong></td>
<td>mg-COD/cycle</td>
<td>B</td>
<td>99.0</td>
<td>61.4</td>
</tr>
<tr>
<td><strong>Biomass</strong></td>
<td>mg-COD/cycle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>O₂</strong></td>
<td>mg-O₂/cycle</td>
<td>B</td>
<td>−554.9</td>
<td>14.4</td>
</tr>
<tr>
<td><strong>CO₂</strong></td>
<td>mg-C/O/cycle</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Famine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Average Biomass</strong></td>
<td>mg-COD/L</td>
<td>B</td>
<td>2361.8</td>
<td>54.3</td>
</tr>
<tr>
<td><strong>Average ASH</strong></td>
<td>mg/L</td>
<td></td>
<td>618.1</td>
<td>20.3</td>
</tr>
<tr>
<td><strong>NH₄⁺</strong></td>
<td>mg-N/cycle</td>
<td>B</td>
<td>0.0&lt;sup&gt;f&lt;/sup&gt;</td>
<td>27.0&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Glycogen</strong></td>
<td>mg-COD/cycle</td>
<td>B</td>
<td>−89.4</td>
<td>24.1</td>
</tr>
<tr>
<td><strong>PHB</strong></td>
<td>mg-COD/cycle</td>
<td>B</td>
<td>−113.7</td>
<td>30.6</td>
</tr>
<tr>
<td><strong>TAG</strong></td>
<td>mg-COD/cycle</td>
<td>B</td>
<td>−99.0</td>
<td>61.4</td>
</tr>
<tr>
<td><strong>Biomass</strong></td>
<td>mg-COD/cycle</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>O₂</strong></td>
<td>mg-O₂/cycle</td>
<td>B</td>
<td>−182.2</td>
<td>39.1</td>
</tr>
<tr>
<td><strong>CO₂</strong></td>
<td>mg-C/O/cycle</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall Biomass&lt;sup&gt;c&lt;/sup&gt;</strong></td>
<td>mg-COD/cycle</td>
<td>B</td>
<td>253.9</td>
<td>57.0</td>
</tr>
</tbody>
</table>

<sup>a</sup>: poly-β-hydroxybutyrate; <sup>b</sup>: triacylglycerol; <sup>c</sup>: total amount of active biomass produced in one feast-famine cycle; <sup>d</sup>: this column indicates whether the compound will be balanced (B), which means that a better estimate will be found and be calculated (C) using the elemental composition matrix in Table 4.1; <sup>e</sup>: not applicable; <sup>f</sup>: the ammonium uptake rate in the famine phase was set to zero, and its respective error was set twice as much as that of the feast phase. Note that h = 4.87 against $\chi^2 = 16.80$ at 99% confidence level.
Table 4.5 summarizes the estimated specific substrate uptake rates, the growth rates, the storage production rates, biomass and storage yields of *R. jostii* RHA1 during the feast-famine growth cycle. The specific glucose uptake rate was almost twice as much as the acetate uptake rate. The *R. jostii* RHA1 growth rates in the feast and famine phases on glucose were the same to those on acetate, respectively. In turn, the overall growth rates on both substrates were the same.

Moreover, the growth rate in the feast phase on both substrates was almost twice as large as that of the famine phase. The storage production rates and storage yields of glycogen, PHB, and TAG...
were similar on glucose. Furthermore, the storage production rate and yield of PHB on acetate was the highest which was 3 times higher than those of TAG, while glycogen was virtually not accumulated on acetate. Finally, the biomass yield on acetate was twice as large as that on glucose in both feast and famine periods; as a result, the overall biomass yield on acetate was twice as high as that on glucose.

Table 4.5 – The estimated specific substrate uptake rates, the growth rates, the storage production rates, storage and biomass yields of *R. jostii* RHA1 during the feast-famine growth cycle

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Units</th>
<th>Glucose</th>
<th>Acetate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Feast</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$q_{\text{Substrate}}$</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>3.57±0.07</td>
<td>1.82±0.07</td>
</tr>
<tr>
<td>$\mu_{\text{Feast}}$</td>
<td>h$^{-1}$</td>
<td>0.52±0.20</td>
<td>0.51±0.18</td>
</tr>
<tr>
<td>$q_{\text{Glycogen}}$</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>0.32±0.06</td>
<td>0.01±0.00</td>
</tr>
<tr>
<td>$q_{\text{PHB}}^a$</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>0.40±0.08</td>
<td>0.34±0.20</td>
</tr>
<tr>
<td>$q_{\text{TAG}}^b$</td>
<td>g-COD/(g-COD-Xa·h)</td>
<td>0.29±0.16</td>
<td>0.12±0.08</td>
</tr>
<tr>
<td>$\gamma_{\text{Glycogen}}$</td>
<td>g-COD/g-COD-Substrate</td>
<td>0.09±0.02</td>
<td>0.01±0.00</td>
</tr>
<tr>
<td>$\gamma_{\text{PHB}}$</td>
<td>g-COD/g-COD-Substrate</td>
<td>0.11±0.02</td>
<td>0.18±0.11</td>
</tr>
<tr>
<td>$\gamma_{\text{TAG}}$</td>
<td>g-COD/g-COD-Substrate</td>
<td>0.08±0.04</td>
<td>0.07±0.04</td>
</tr>
<tr>
<td>$\gamma_{\text{Biomass}}$</td>
<td>g-COD/g-COD-Substrate</td>
<td>0.15±0.06</td>
<td>0.28±0.10</td>
</tr>
<tr>
<td><strong>Famine</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{\text{Famine}}$</td>
<td>h$^{-1}$</td>
<td>0.22±0.13</td>
<td>0.22±0.14</td>
</tr>
<tr>
<td>$\gamma_{\text{Biomass}}$</td>
<td>g-COD/g-COD-Substrate</td>
<td>0.06±0.04</td>
<td>0.12±0.08</td>
</tr>
<tr>
<td>Overall $\mu$</td>
<td>h$^{-1}$</td>
<td>0.37±0.24</td>
<td>0.36±0.23</td>
</tr>
<tr>
<td>Overall Biomass yield</td>
<td>g-COD/g-COD-Substrate</td>
<td>0.15±0.02</td>
<td>0.31±0.05</td>
</tr>
</tbody>
</table>

$^a$: poly-β-hydroxybutyrate; $^b$: triacylglycerol. Note that $\mu$ stands for the growth rates, $\gamma$ represents the yields, and $q$ denotes for the specific uptake and production rates.

4.4 Discussion

In this study, we report the storage metabolism by *R. jostii* RHA1 during the feast-famine cycle on glucose and acetate as the sole carbon sources. We observed the production of glycogen, PHB, and TAG by *R. jostii* RHA1 in the feast period. These results agree with the previous observations in growth of *R. jostii* RHA1 in the nutrient limiting conditions (Hernandez et al. 2008).
It seems that *R. jostii* RHA1 can take up glucose and acetate faster in the feast-famine cycle than in the non-limited batch growth condition. The glucose and acetate uptake rates were 3.57±0.07 and 1.82±0.07 g-COD/(g-COD-Xa·h), while these in the non-limited batch growth condition were 0.30±0.23 and 1.14±0.37 g-COD/(g-COD-Xa·h), respectively (Chapter 3, (Tajparast and Frigon 2014)). Therefore, *R. jostii* RHA1 is a fast consumer during feast-famine growth cycles.

The growth rate of *R. jostii* RHA1 in the feast phase on glucose (0.52±0.20 h⁻¹) was twice as large as that in the non-limited growth condition (0.23±0.24 h⁻¹); however, the growth rate of *R. jostii* RHA1 in the feast phase on acetate (0.51±0.18 h⁻¹) was half the amount in the non-limited growth condition (1.03±0.37 h⁻¹). It seems that *R. jostii* RHA1 is a fast growing bacterium in the feast phase on glucose. In contrast, *R. jostii* RHA1 is a slow growing microorganism in the feast phase on acetate. By comparing the overall growth rates of *R. jostii* RHA1 in the feast-famine cycle on glucose (0.37±0.24 h⁻¹) and acetate (0.36±0.23 h⁻¹) with those in the non-limited growth conditions (i.e., 0.23±0.24 and 1.03±0.37 h⁻¹ on glucose and acetate, respectively), this conclusion was further valid (Chapter 3, (Tajparast and Frigon 2014)).

The overall average biomass yields of *R. jostii* RHA1 during the feast-famine cycle on glucose and acetate were 0.15±0.02 and 0.31±0.05 g-COD/g-COD-Substrate, respectively, while these were 0.77±1.39 and 0.90±0.62 g-COD/g-COD-Substrate, respectively, in the non-limited condition (Chapter 3, (Tajparast and Frigon 2014)). That is, the biomass yield in the non-limited condition on glucose was 5 times higher than that in the feast-famine cycle, while the biomass yield in the non-limited condition on acetate was 3 times higher than that in the feast-famine cycle (Chapter 3, (Tajparast and Frigon 2014)). It should be noted that the ± values were standard errors. Although some standard errors were higher than the average values, the model
errors were not significant as it was shown by comparing the $\chi^2$ values with the $h$ values displayed in the footnotes of Tables 4.3 and 4.4. Basically, an error is not significant if $h < \chi^2$.

When glucose was the substrate, the COD accumulation during the feast phase was similar for the three storage compounds: glycogen, PHB, and TAG, as shown by the storage production rates and yields (Table 4.5). However, when acetate was the substrate, essentially no glycogen was accumulated and the COD flux (and yield) towards PHB accumulation was 3 times higher than toward TAG accumulation. However, the relative fluxes of storage compounds during accumulation by *R. jostii* RHA1 in a sudden nitrogen limitation following non-limited growth were PHA > TAG > glycogen (Chapter 3, (Tajparast and Frigon 2014)). This suggests that the metabolic behaviours of the cell should be different in these two different environmental conditions. It seems that this principle promotes early diversion of carbon from the substrate to the storage pools during the feast-famine growth cycle, a principle that does not seem to be the one operating for *R. jostii* RHA1 when a sudden N-limitation occurs. Finally, due to the lipid content of *R. jostii* RHA1 cells, the microorganism tends to accumulate lipid type storage pools especially on fatty acids such as acetic acid.

**Acknowledgments**

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant awarded to DF. MT is grateful to the McGill Engineering Doctoral Award (MEDA) for the financial support of his Ph.D. study.
4.5 References


Tajparast M, Frigon D. 2014. Genome-scale metabolic model of *Rhodococcus jostii* RHA1 (iMT1174) to study the accumulation of storage compounds during nitrogen limited condition. BMC Systems Biology Submitted.


In chapter 4, *R. jostii* RHA1 growth on glucose and acetate during the feast-famine cycles were detailed. When glucose was the substrate, the COD accumulation during the feast phase was similar for the three storage compounds: PHB, TAG, and glycogen. However, when acetate was the substrate, essentially no glycogen was accumulated and the COD flux (and yield) towards PHB accumulation was 3 times higher than toward TAG accumulation. In chapter 5 the genome-scale flux balance analysis was implemented to predict storage metabolism of *R. jostii* RHA1 during the feast-famine cycles on glucose and acetate. In addition, the $^{13}$C-labelling experiments were performed to gain insight into the storage metabolism of *R. jostii* RHA1 during the feast-famine cycles which are detailed in chapter 5. Furthermore, in chapter 5 the fluxes of the central metabolic reactions in the feast and famine phases simulated by the $^{13}$C-metabolic flux analysis were compared.
Chapter 5. Predicting the Accumulation of Storage Compounds by *Rhodococcus jostii* RHA1 in the Feast-famine Growth Cycles Using Genome-scale Flux Balance Analysis

Mohammad Tajparast¹ and Dominic Frigon¹*

¹Microbial Community Engineering Laboratory, Department of Civil Engineering and Applied Mechanics, McGill University, 817 Sherbrooke Street West, Montreal, Quebec, Canada, H3A 2K6

*Corresponding author

Email addresses:

MT: mohammad.tajparast@mail.mcgill.ca

DF: dominic.frigon@mcgill.ca

Short Title: Genome-scale Flux Balance of RHA1 during Feast-famine Cycles

To be submitted to

Biotechnology and Bioengineering
Abstract

Feast-famine cycles in biological systems select for bacterial species that accumulate intracellular storage compounds such as glycogen and poly-β-hydroxybutyrate (PHB). These species survive better the famine phase and resume rapid substrate uptake at the beginning of the feast phase faster than microorganisms unable to accumulate storage. Ecophysiological conditions favouring the selection of species accumulating either storage compounds remain to be clarified. Example questions: Are substrates present in the wastewater the main factor determining this selection? Are there metabolic organizations and strategies deployed by certain species that can explain this selection? These questions were investigated using a genome-scale metabolic modeling approach. In this study, we used the genome-scale metabolic model of a heterotrophic model organism namely *Rhodococcus jostii* RHA1 (iMT1174) and simulated feast/famine growth conditions by means of flux balance analysis (FBA) subject to commonly used metabolic objective functions; here, maximization of the growth rate was set as the main objective function, while minimization of network fluxes, and minimization of metabolic adjustment (*environmental* MoMA) were considered as the sub-objective functions. It turned out that both flux minimization and *environmental* MoMA predicted identical results on glucose, while the former was able to better predict the storage metabolism of *R. jostii* RHA1 on acetate. It showed PHB as the main storage pool on both glucose and acetate that was inline with our experimentally observed storage fluxes. Further, experiments with $^{13}$C-labeled bicarbonate ($\text{HCO}_3^-$) suggested that the fluxes through the central metabolism reactions during the feast phases were similar to the one during the famine phase. That justified the use of the *environmental* MoMA sub-objective function to compare two successive environmental conditions such as the feast and famine phases.
Keywords
Feast-famine cycle; *Rhodococcus jostii* RHA1; flux balance analysis; $^{13}$C-metabolic flux analysis; minimization of metabolic adjustment; minimization of metabolic fluxes
5.1 Introduction

Natural environments like soil and biological wastewater treatment systems are often highly dynamic with regard to availability of external electron donors. Thus, the biomass is submitted to cycles of feast and famine with the feast phase being the period of availability of external electron donors and the famine phase being the starvation period (Van Loosdrecht et al. 1997). The presence of such cycles profoundly affects the microbial ecology and the metabolism of the bacterial species present. One of the main metabolisms found under this cycling is the accumulation (during feast) and consumption (during famine) of storage compounds such as glycogen and poly-β-hydroxybutyrate (Dawes and Senior 1973). Bacterial species accumulating storage compounds during the feast phase outperform the species lacking this ability probably by maintaining the capacity to rapidly respond to a substrate addition such that the metabolic cost of storage in terms of yield is lower than the competition cost to rapidly gain access to the substrate (Frigon et al. 2006). Therefore, understanding and properly modeling the metabolism of the storage compounds is central to predict the population dynamics in activated sludge wastewater treatment systems.

In the current post-genomic era, genomics (sequences of the entire genetic material of a strain) and meta-genomics (genomics of entire communities) will provide useful information to understand the metabolic activity of microorganisms involved in feast-famine cycles. However, the list of genes that an organism possesses (the information contained in the annotated genome sequence of an organism) is insufficient to predict its metabolic behaviour. A genome-scale metabolic model is a mathematical representation of the stoichiometry of the biochemical network. This is one of the tools to make the metabolic predictions. Here, the model organism Rhodococcus jostii strain RHA1 (metabolic model iMT1174, Chapter 3) was used to investigate the ability of flux balance analysis (FBA) to predict storage fluxes in feast-famine cycles. This
bacterial species was selected because of the availability of genomic data (McLeod et al. 2006) and genome-scale metabolic model (Chapter 3), and because it is found in the environment and in activated sludge wastewater treatment systems. Also of special interest for this work, *R. jostii* RHA1 is able to accumulate most of the storage compounds known as: poly-β-hydroxybutyrate (PHB), poly(3-hydroxyvalerate) (PHV), triacylglycerols (TAG), wax esters (WE), α,α-trehalose, glycogen, and polyphosphate (PolyP) (Hernandez et al. 2008).

Solution of the genome-scale metabolic model at steady-state using FBA typically yields a set of underdetermined mass balance equations. The solution space, however, can be investigated by finding solutions optimally satisfying biologically relevant linear objective functions implemented in a linear-programming approach (Kauffman et al. 2003; Reed and Palsson 2003). Defining relevant objective functions for metabolic predictions in the numerous environmental conditions remain at its infancy; and one of the goals of the current work is to evaluate the performance of some of the objective functions available in the literature.

Despite considerable experimental research on activated sludge and pure cultures on the dynamic microbial storage metabolism (Beun et al. 2002; Beun et al. 2000; Carta et al. 2001; Frigon et al. 2006; Goel et al. 1998; Van Aalst-Van Leeuwen et al. 1997), proper metabolic description remain scarce. This impedes the development of proper objective functions. Thus, $^{13}$C-labeling experiments were performed to get insight into the storage metabolism of *R. jostii* RHA1 during the feast-famine cycles. The isotopomer mass distributions of the labelled amino acids were utilized to model the central metabolic pathways of *R. jostii* RHA1 using $^{13}$C-metabolic flux analysis ($^{13}$C-MFA). To provide equivalent sources of $^{13}$C-labelled carbon in the feast and famine, $^{13}$C-bicarbonate was used.
Using the same experimental data, the performance of three objective functions to predict the fluxes of three storage compounds were evaluated. The choice of an objective function is central to the FBA approach. To defined the objective functions used, environmental competition and the niche exclusion principle were first considered (Feist and Palsson 2010), which led to the assumption that the environmental interaction will push the organisms at maximizing its growth rate (which is equivalent to maximizing yield if organic carbon is not secreted). Thus, “the survival of the fittest” infers that the fastest one expressing the optimal responses for adaptation to the environment survives (Ponce De Leon et al. 2008). In light of this principle, the maximal cellular growth rate (or yield) was used as the principal objective function. Sub-objective functions were then coupled with the principal objective functions to represent other metabolic/environmental constraints. It is possible that organisms pursue a global optimal use of their metabolic network by minimizing the total metabolic fluxes (i.e., maximizing cellular efficiency) (Holzhutter 2004). Another possible objective could be to minimize the flux differences between feast and famine growth, which would minimize the metabolic cost of cellular rearrangements between growth phases. This objective function was first developed for the analysis of gene mutations under the name minimization of metabolic adjustment (MoMA) (Segre et al. 2002). Thus, we subbed it *environmental MoMA*.

In this work, we present the first application of genome-scale metabolic modeling in the area of environmental biotechnology in order to better understand the cellular storage accumulation during feast-famine cycles with the emphasis on glycogen, PHB, and TAG. The fluxes of the central metabolic reactions of *R. jostii* RHA1 (the model organism of heterotrophs) simulated with $^{13}$C-MFA in the feast and famine phases on unlabeled glucose and acetate and $^{13}$C-bicarbonate are displayed. The genome-scale metabolic modeling approach and simulation
results for growth of \textit{R. jostii} RHA1 on glucose and acetate as the sole carbon sources are discussed.

\textbf{5.2 Materials and Methods}

\textbf{5.2.1 Metabolic Network Reconstruction and Flux Balance Analysis}

As the first step in reconstructing the cellular metabolic network of \textit{R. jostii} RHA1 (iMT1174), the list of genes in its annotated genome sequence was converted to a list of associated balanced biochemical reactions, which form the basic framework of the model. Subsequently, the model was analyzed in light of literature data to determine if some reactions may have been missing from the list of annotated genes. The reactions are added along with transport reactions to make a physiologically meaningful cellular model. In the case of \textit{R. jostii} RHA1, although annotated, several degradative pathways of various xenobiotics were missing from KEGG (the primary reaction database used (Kanehisa et al. 2004)) and they were added to the model to account for its catabolic ability. Then, the metabolic network was completed by the pathways of seven storage compounds that have also been considered and manually curated in the model: PHB, PHV (Hernandez et al. 2008; Madison and Huisman 1999), TAG, WE (Hernandez et al. 2008; Ishige et al. 2002; Waltermann et al. 2007), \textit{a,a}-trehalose, glycogen, and PolyP. Finally, eight macromolecules were assigned to make up the biomass (e.g., protein, DNA, RNA, phospholipids, small molecules such as electron carriers and coenzymes, peptidoglycan, carbohydrates, and corynomycolic acids). This model is contained in the stoichiometric matrix (S) of the biochemical network with rows and columns corresponding to the metabolites and reactions, respectively.

Briefly, the \textit{in silico} metabolic network of \textit{R. jostii} RHA1 (iMT1174) comprises of 1243 balanced (intracellular) compounds, 1935 unique fluxes, and 1174 open reading frames (ORFs).
The model also includes 330 extracellular compounds that are associated with 518 exchange fluxes. In this model, seven storage compounds were defined independently of the biomass composition as extracellular compounds connected to 11 virtual exchange fluxes to study the variation of storage compound accumulation namely: glycogen, α,α-trehalose, PHB, PHV, TAG, WE, and PolyP.

The gene-protein-reaction (GPR) association, along with the stoichiometric matrix, is a qualitative representation of the cellular network. This model is studied quantitatively by setting up a series of mass balances around each metabolite that can be expressed in matrix notation as:

\[ \frac{dx}{dt} = S \cdot v; \]

where \( x \) is the \((m \times 1)\) vector of the concentrations of the balanced metabolites (i.e., intracellular and macromolecular compounds), \( v \) denotes the \((n \times 1)\) vector of the entire metabolic fluxes, and \( S \) stands for the \(m \times n\) stoichiometric matrix.

The iMT1174 model is then solved by assuming that the cellular network is at steady-state, which simplifies the previous equation to \( S \cdot v = 0 \). This assumption is plausible due to the fast equilibration of intracellular metabolite pools (time-scale of seconds) compared to the time-scale of genetic regulation (minutes) (Segre et al. 2002) and due to high reaction rates of intracellular reactions compared to exchange rates such as substrate uptake, cell growth, and by-product secretion rates (Lee et al. 2006; Roels 1982; Varma and Palsson 1994). Consequently, the output of the model is a distribution of metabolic fluxes through the various chemical reactions (note that the kinetics is not modelled). Since the number of fluxes normally exceeds the number of metabolites \((n > m)\), the problem is said to be underdetermined. In this case, there is a solution space for the fluxes that can be studied by optimizing proper objective functions subject to defined constraints (the mass balance equations as the physicochemical constraints and the inequalities as the enzymatic capacity of the biochemical reactions) (Becker et al. 2007).
One way to simulate metabolic behaviour using the stoichiometric network is to apply objective functions (optimization criteria). For instance, one could experimentally examine the phenotypic responses when activated sludge bacteria are grown in periods of feast and famine conditions. Then, one can find the objective function(s) that properly describe the observed phenotypic responses using the metabolic model of the model organisms under study (here, *R. jostii* RHA1). In this context, the following is the type of question that we aim to answer: is there an objective function that enables us to predict the type of storage compound(s) and their associated metabolic fluxes exhibited by the bacteria?

Although various objective functions exist (Palsson 2006), the maximization of the growth rate was reported to be successful at predicting experimental observations for bacteria (Edwards et al. 2001; Ibarra et al. 2002; Segre et al. 2002). Based on the hypotheses developed earlier, we investigated the value of maximizing the cellular growth rate in order to calculate the unknown metabolic fluxes during the feast and famine cycle. In addition, we developed a nested optimization algorithm to minimize the sum of absolute values (i.e., the Manhattan norm) of all fluxes (Holzhutter 2004) (or minimization of metabolic fluxes). We also adapted a sub-optimization algorithm, so-called minimization of metabolic adjustment (*environmental* MoMA), in order to minimize the metabolic activities between the feast and famine periods by minimizing the Manhattan norm of their flux differences (i.e., \(|v_{\text{Feast}} - v_{\text{Famine}}|\)) (Segre et al. 2002).

Simulating the dynamic metabolic behaviour during the feast and famine cycle, we needed to conceptualize a simplified scenario. We simulated the two phases of growth (i.e., feast and famine phases) independently (except when the *environmental* MoMA sub-objective function was applied). According to our experimental observations, we assumed that the feast phase lasted only 38% and 40% of the cycle time on glucose and acetate, respectively, and was
characterized by a flux distribution (Chapter 4, (Tajparast and Frigon 2014)). Finally, using our assumption that the feast phase lasts 38% and 40% of the cycle time (on glucose and acetate, respectively), the mass balance on the storage compounds was constrained by equating the storage synthesis rate in the feast phase and the storage consumption rate in the famine phase.

Using this scenario, we proceeded to map the flux solutions in the space described by the coordinates of substrate uptake rates and storage metabolic fluxes. To mimic the \( R. jostii \) RHA1 growth condition, we fixed the substrate uptake rates to the experimentally observed values; moreover, we investigated a wide range of the biosynthetic fluxes of the storage compounds in order to assure that we cover all metabolic activities towards storage accumulation. From the simulation results, one can examine the flux distribution of each phase separately or derive an average flux distribution for the entire cycle by computing the time length-weighted average of the two phases.

### 5.2.2 \(^{13}\text{C}\)-Metabolic Flux Analysis

We aim at estimating the reactions involved in central metabolism using \(^{13}\text{C}\)-MFA. The model represents the central metabolism which includes glycolysis and gluconeogenesis, Entner–Doudoroff pathway, tricarboxylic acid (TCA) cycle, pentose phosphate pathway, anaplerotic carboxylation and decarboxylation, storage metabolic reactions, amino acid biosynthetic reactions, and anabolic routes into biomass (Figure A1). We used the openFLUX software application under MATLAB environment (Mathworks Inc, Massachusetts) to solve for the fluxes (Quek et al. 2009). The application is based on the Elementary Metabolite Unit (EMU) framework. Stoichiometric data on growth, substrate uptake rate, storage formation, and on the cellular composition of \( R. jostii \) RHA1 together with mass isotopomer distribution data of the
labeled amino acids (Additional File 1-4) that were produced using the iMS2Flux software (Poskar et al. 2012) were used as model input.

5.2.3 Bacterial Strain and Growth Medium

*Rhodococcus jostii* strain RHA1 was grown on the medium supplemented with a certain amount of labelled carbon source (e.g., $^{13}$C-bicarbonate) as described below (Beun et al. 2000; Vishniac and Santer 1957).

We utilized the medium composed of NH$_4$Cl (6.00 mM), KH$_2$PO$_4$ (1.65 mM), MgSO$_4$ \( \cdot \) 7H$_2$O (1.11 mM), KCl (1.50 mM), and trace element solution (1 mL/L). We also added labelled bicarbonate (100 µL of 2.68 g/L solution) at the beginning of either the feast phase or the famine phase.

The trace element solution contained: Ethylenediamine tetra-acetic acid (EDTA: \( \text{C}_{10}\text{H}_{16}\text{N}_{2}\text{O}_{8} \)) (50.00 g/L), ZnSO$_4$ \( \cdot \) 7H$_2$O (22.00 g/L), CaCl$_2$ (5.54 g/L), MnCl$_2$ \( \cdot \) 4H$_2$O (5.06 g/L), FeSO$_4$ \( \cdot \) 7H$_2$O (4.99 g/L), (NH$_4$)$_6$Mo$_7$O$_{24}$ \( \cdot \) 4H$_2$O (1.10 g/L), CuSO$_4$ \( \cdot \) 5H$_2$O (1.57 g/L), CoCl$_2$ \( \cdot \) 6H$_2$O (1.61 g/L) dissolved in 1000 mL dilution water; adjusted to pH 6.0 with KOH (Vishniac and Santer 1957). The dilution water was supplemented with 55.25 mM MgSO$_4$ \( \cdot \) 7H$_2$O in excess to promote biomass flocculation.

5.2.4 Reactor Condition

The labeling tests were performed using a μ-24 bioreactor (Applikon Biotechnology). It comprises of 24 cells with 5 mL working volume. The dissolved oxygen concentration and temperature of the reactors were controlled, while pH was monitored throughout the test. First, 1.25 mL of the fresh media was added to each cell and 3.75 mL of *R. jostii* RHA1 culture (from the end of the famine phase), which had already been subjected to the steady-state feast-famine
cycle in a 2-L bioreactor prior to the inoculation, was also added to the cell. Several samples at the end of the feast and famine phases were taken by adding 400 µL of concentrated HCl and further centrifuging; solids were washed with deionized water and stored at −80 °C for further analysis of the labelled amino acids using gas chromatography mass spectrometry (GC-MS).

5.2.5 Analysis of Labelled Amino Acids
Proteinogenic amino acids of the biomass were analyzed using GC-MS technique according to Nanchen et al. with some modifications (Nanchen et al. 2007). The biomass samples which had already been stored at −80 °C were thawed and resuspended in 1.5 mL sterile deionized water, vortexed, and homogenized using ultrasonic treatment (60 watts) for 15 min. After resuspension, 100 µL of sample was transferred into a 2 mL microcentrifuge tube, further spun down the cell pellets (at 15800g at room temperature for 15 min). Cell pellets were washed twice by resuspension in 1 mL 0.9% NaCl, and centrifuged at 15800g at room temperature for 15 min. The washed pellets were resuspended in 1 mL of 6 M HCl, and hydrolyzed for 24 h at 110 °C in a well-sealed screw-capped tube to prevent evaporation. The hydrolyzate was dried overnight in a heating block at around 70 °C and under a constant air stream in a fume hood. The dry hydrolyzate was dissolved in 30 µL of a reagent containing 10 mg of Methoxyamine Hydrochloride per 1 mL of Anhydrous Pyridine. 1 µL of the internal standard was added into each sample tube, note that the internal standard was 750 ng/µL of deuterated Myristic acid (so-called D_{27}-myristic acid), the samples were mixed by vortex and sonication for several times, each time taking 10-20 sec. Furthermore, the samples were centrifuged at 15000 rpm for 10 min at room temperature. The samples were transferred into GC-MS vials and cooked at 70 °C for 30 min in a hot block. 70 µL of N-tert-butyldimethylsilyl-N-methytrifluoroacetamide (TBDMS) was added into the vials and cooked at 70 °C for 1 hr in a hot block to derivatize the samples.
GC-MS analyses were performed with an Agilent 5975C mass selective detector coupled to a 7890A gas chromatograph (Agilent Technologies, Santa Clara, CA, USA) fitted with a 7693 autosampler and a DB-5MS+DG capillary column (30 m plus 10 m Duraguard®), diameter 0.25 mm, film thickness 0.25 mm (Agilent J & W, Santa Clara, CA, USA). The GC temperature program started with a 1 min hold at 60 °C followed by a 10 °C/min ramp to 300 °C. Bake-out was at 320 °C for 10 min. The injector and interface to the MS were held at 285 °C. The helium carrier flow rate was held constant at 1.5 mL/min (or a flow rate such that the TBDMS derivative of the D_{27}-myristic acid has a retention time of 18 min). When operated in full scan mode, the scan range was 50 - 700 Da. 1 µL of the sample was injected in splitless mode.

5.2.6 Statistical Analyses

Model II regression was applied to account for the errors in the estimated non-zero fluxes of the central metabolic reactions of *R. jostii* RHA1 in the feast and famine phases, since there were errors associated with the estimated fluxes in both feast and famine phases. Major axis regression was also employed to see the correlation between two sets of the non-zero central metabolic fluxes in the feast and famine phases estimated using ^13^C-MFA (Legendre and Legendre 2012; Sokal and Rohlf 1995). The major axis slope and the coefficient of determination ($R^2$), along with half of the difference between the confidence intervals of slope at 95% confidence interval (i.e., $\alpha = 0.05$) on glucose and acetate were calculated. Here, the confidence interval was calculated with an analytical solution according to (Sokal and Rohlf 1995).

The Hotelling’s multivariate $T^2$ test was also applied to investigate the difference between the observed and predicted storage fluxes in the confidence interval of 95% (i.e., $\alpha = 0.05$) (Brown et al. 2011). The null hypothesis states that there is no difference between experimentally
observed storage fluxes and the predicted ones using the minFluxes and MoMA objective functions. If the \( p \) value is greater than the \( \alpha \) value (i.e., \( p > 0.05 \)), it means that the results are not significant and we fail to reject the null hypothesis. In other words, there is no difference between the measured and predicted storage fluxes.

5.3 Results and Discussion

5.3.1 Central Metabolism Reaction Fluxes during Feast-Famine Cycles

In Chapter 4, the storage metabolic fluxes during the feast-famine cycle were described for \( R. jostii \) RHA1 growing in a sequencing batch reactor on glucose or acetate (Chapter 4, (Tajparast and Frigon 2014)). Here, the fluxes through the central metabolism reactions are further characterized using \(^{13}\)C-MFA with \(^{13}\)C-bicarbonate as the labeled carbon sources. \( R. jostii \) RHA1 was fed with unlabeled glucose or acetate, and \(^{13}\)C-bicarbonate was added to the medium either at the beginning of the feast or the famine phase. The \(^{13}\)C-MFA solution of the models was obtained by adjusting the metabolic fluxes that were experimentally determined in Chapter 4 (Tajparast and Frigon 2014) to their measured values (see Additional Files 1 through 4 for more details). Additionally, the effective \(^{13}\)C-bicarbonate uptake rate was unknown. Thus, \(^{13}\)C-bicarbonate uptake rate was adjusted by minimizing the sum of absolute differences between the calculated and experimentally observed mass isotopomer distribution (i.e., the error) vectors in all cases. The optimal value of \(^{13}\)C-bicarbonate uptake rate was found to be 0.05 mmol/(g-DW·h) by this method. The fluxes determined for simulations with the lowest sum of absolute differences between the calculated and experimentally observed mass isotopomer distribution are presented in Figures 5.1 and 5.2.

According to Figure 5.1, \( R. jostii \) RHA1 exhibited similar metabolic fluxes of the central metabolic reactions on glucose in the feast and famine phases, whereas Figure 5.2 showed that \( R. \)
*jostii* RHA1 exhibited similar metabolic fluxes of the central metabolic reactions on acetate in the feast and famine phases.

Figure 5.1 – Fluxes of the central metabolic reactions of *R. jostii* RHA1 on unlabeled glucose and $^{13}$C-bicarbonate in the feast phase (left panel) and the famine phase (right panel).

Note that unlabeled glucose was fed at the beginning of the feast phase, while $^{13}$C-bicarbonate was fed at the beginning of the feast or famine phase.
Figure 5.2 – Fluxes of the central metabolic reactions of *R. jostii* RHA1 on unlabeled acetate and $^{13}$C-bicarbonate in the feast phase (left panel) and the famine phase (right panel).

Note that unlabeled acetate was fed at the beginning of the feast phase, while $^{13}$C-bicarbonate was fed at the beginning of the feast or famine phase.

Figure 5.3 shows the correlation between the estimated fluxes of the central metabolism reactions of *R. jostii* RHA1 in the feast and famine phases on glucose (Figure 5.3a) and acetate (Figure 5.3b). Accordingly, the fluxes in the feast and famine phases are well correlated showing that the metabolic behaviours of the cell are similar in both the feast and famine phase. This substantiates the use of the *environmental* MoMA sub-objective function in this study.
Figure 5.3 – Correlation plots of the central metabolism reactions of *R. jostii* RHA1 in the feast and famine phases estimated using $^{13}$C-MFA on glucose (a) and acetate (b). Note that 11 reactions were zero in both cases.

As mentioned earlier, because there were errors associated with the estimated fluxes in the feast and famine phases, model II regression was employed to account for errors in both axes in Figure 5.3. Table 5.1 shows the major axis slope and the coefficient of determination ($R^2$), along with half of the difference between the confidence intervals of slope at 95% confidence interval (i.e., $\alpha = 0.05$) on glucose and acetate. In both cases, the slopes were different from zero (i.e., zero is beyond the confidence interval limits) and the $R^2$ values were high showing a good correlation between the central metabolic fluxes in the feast and famine phases.
Table 5.1 – Comparison of the goodness of fit of the non-zero central metabolism reactions of *R. jostii* RHA1 in the famine phase versus the feast phase estimated using $^{13}$C-MFA on glucose and acetate

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Major axis slope&lt;sup&gt;a&lt;/sup&gt;</th>
<th>$R^2$&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>$0.740 \pm 0.140$</td>
<td>0.802</td>
</tr>
<tr>
<td>Acetate</td>
<td>$0.903 \pm 0.054$</td>
<td>0.966</td>
</tr>
</tbody>
</table>

<sup>a</sup> major axis slope ± (difference between confidence intervals of slope)/2 at 95% confidence intervals. <sup>b</sup> Coefficient of determination.

5.3.2. FBA to Predict Storage Compounds Fluxes during Feast-Famine Cycles

The capability of FBA to predict storage compounds’ fluxes during feast-famine cycles was then investigated. Using the experimental data obtained in Chapter 4 for comparison purposes, the fluxes for the synthesis (feast) and consumption (famine) of glycogen, PHB, and TAG, were simulated using the maximization of the growth rate objective function either alone or in conjunction with three sub-objective functions: i) minimizing the Manhattan norm (sum of absolute) of the network fluxes (i.e., $|v|_1$, minimization of fluxes, or minFluxes) (Holzhutter 2004), ii) minimization of metabolic adjustment between the feast and famine periods (*environmental* MoMA) (Segre et al. 2002), iii) minimizing both the Manhattan norm of the flux vectors and their difference over the feast-famine cycle. Although the use of the MoMA objective function is different than the scenario for which it was designed to simulate (i.e., gene mutations), the results of the $^{13}$C-MFA suggested that this could be a proper use of the MoMA principle.

For a given substrate uptake rate (here the one experimentally observed), the storage fluxes are highly sensitive to the growth-associated maintenance energy and to the sub-objective functions used. Figure 5.4 shows the total storage flux as a function of the growth-associated maintenance energy (GAME) on glucose and acetate. The GAMEs were adjusted such that the total storage fluxes corresponded to the experimental ones. For acetate and glucose, respectively,
the GAMEs were: 5.5 and 447.7 mmol-ATP/g-DW for maximization of the growth rate, 5.5 and 451.0 mmol-ATP/g-DW for flux minimization, and 12.0 and 451.0 mmol-ATP/g-DW for environmental MoMA.

Figure 5.4 – Total storage flux as a function of the growth-associated maintenance energy on (a) glucose and (b) acetate. Objective functions are: MoMA, minimization of fluxes, and maximization of the growth rate.

As mentioned earlier, the first objective function investigated to assess the metabolic value of the storage compounds was the maximization of the growth rate. This is equivalent to yield maximization at low growth rates when organic carbon is not secreted. Figure 5.5 shows the contour plots of the total maximum growth rate of *R. jostii* RHA1 as a function of set storage fluxes for three different storage fluxes (glycogen, PHB, and TAG) with glucose and acetate as substrate. The total maximum growth rate decreases with increasing storage fluxes (Figure 5.5), which is due to the metabolic cost of storage transformations. For all other combinations, see Figures A2-A4 in the Appendix.
Figure 5.5 – Contour plots of the weighted average maximum growth rate of *R. jostii* RHA1 vs. a pair of storage fluxes (glycogen vs PHB) at glucose uptake rate of 156.01 Cmmol/(g-DW·h) (a,
b, and c) and acetate uptake rate of 79.51 Cmmol/(g-DW·h) (d, e, and f) at the biosynthetic flux of TAG as 8.81 Cmmol/(g-DW·h) (a, b, and c) and 3.71 Cmmol/(g-DW·h) (d, e, and f).

Note that the substrate uptake rates and the biosynthetic flux of TAG are the experimentally observed fluxes. Here, three objective functions including maximization of growth rate (a, d), minimization of metabolic fluxes (b, e), and MoMA (c, f) were examined. The maximum capacity of all fluxes of the biochemical network is set unbounded. Note that the numbers appear on the isolines show the contour levels corresponding to the weighted average maximum growth rates (d⁻¹) over the feast and famine cycle.

Since the experimentally observed total growth rates were 8.9 d⁻¹ and 8.7 d⁻¹ on glucose and acetate, respectively, we extracted the optimum values of the storage fluxes at these values predicted by the three (sub-) objective functions (e.g., maximization of yield, minimization of fluxes and environmental MoMA), together with the experimentally observed storage accumulation fluxes in the feast phase (Table 5.2). PHB was predicted to be the main storage compound on both glucose and acetate, since the biosynthetic flux of PHB is the highest among the three storage compounds on both substrates (Table 5.2). The calculated storage fluxes for minimization of fluxes and environmental MoMA were identical when glucose was the substrate. These simulations slightly overestimated the glycogen and PHB fluxes compared to the experimental data, while the TAG fluxes on glucose were identical in both experimental and simulated data. On acetate, the minimization of fluxes sub-objective function predicted identical fluxes to the experimental data for PHB and TAG, while this sub-objective function slightly overestimated the glycogen flux. However, the environmental MoMA sub-objective function underestimated glycogen and TAG productions and overestimated PHB production. Finally, the minimization of fluxes sub-objective function in conjunction with the environmental MoMA
sub-objective function showed similar results to those obtained by the *environmental* MoMA sub-objective function alone (data not shown).

**Table 5.2** – Comparison of the experimentally observed and simulated storage fluxes (in Cmmol/(g-DW·h)) at the optimum levels of the three objective functions examined in the feast-famine growth of *R. jostii* RHA1 on glucose and acetate as the sole carbon sources

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Fluxes</th>
<th>Experimental</th>
<th>maxYield</th>
<th>minFluxes</th>
<th>MoMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>Substrate&lt;sup&gt;a&lt;/sup&gt;</td>
<td>156.0±3.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Glycogen</td>
<td>13.9±2.8</td>
<td>7.0-17.4</td>
<td>17.4</td>
<td>17.4</td>
</tr>
<tr>
<td></td>
<td>PHB</td>
<td>15.8±3.1</td>
<td>11.7-19.5</td>
<td>19.5</td>
<td>19.5</td>
</tr>
<tr>
<td></td>
<td>TAG</td>
<td>8.8±4.8</td>
<td>2.2-11.0</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>Acetate</td>
<td>Substrate&lt;sup&gt;a&lt;/sup&gt;</td>
<td>79.5±3.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Glycogen</td>
<td>0.4±0.2</td>
<td>0.0-15.4</td>
<td>0.6</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>PHB</td>
<td>13.0±7.9</td>
<td>0.0-30.2</td>
<td>13.0</td>
<td>16.2</td>
</tr>
<tr>
<td></td>
<td>TAG</td>
<td>3.7±2.3</td>
<td>0.0-4.6</td>
<td>3.7</td>
<td>0.9</td>
</tr>
</tbody>
</table>

<sup>a</sup> Substrate uptake rates are in Cmmol/(g-DW·h). <sup>b</sup> Experimentally observed biosynthetic fluxes of the storage compounds (in Cmmol/(g-DW·h)). <sup>c</sup> Case i: maximizing the growth rate (total biomass yield in Cmol/Cmol). <sup>d</sup> Case ii: minimizing the Manhattan norm of the flux vectors while maximizing the growth rate (in mmol/(g-DW·h)). <sup>e</sup> Case iii: minimizing the Manhattan norm of the difference between the fluxes over the feast-famine cycle while maximizing their growth rates (in mmol/(g-DW·h)). Note that case iv (minimizing both the Manhattan norm of the flux vectors and their difference over the feast-famine cycle while maximizing their growth rates) showed identical results compared to case iii. Note that the values under the maxYield objective function are its optimal range (minimum value - maximum value).

The results showed very similar predictions for the minimization of fluxes and the *environmental* MoMA, with slightly better prediction for the former objective functions on acetate. Furthermore, the use of the two sub-objective functions did not change the predictions. Given that the minimization of fluxes sub-objective function predicted storage metabolism of *R. jostii* RHA1 during the feast-famine cycle on glucose and acetate better, it seems that the organisms optimize the use of its network. However, because the predictions for the two sub-objective functions are very similar for feast-famine cycles, the two sub-objective functions may be somewhat equivalent.
As mentioned earlier, the Hotelling’s multivariate $T^2$ test was used to investigate the relationship between the observed and predicted storage fluxes in the confidence interval of 95% (i.e., $\alpha = 0.05$) (Brown et al. 2011). Table 5.3 shows the results of the Hotelling’s multivariate $T^2$ test. In all cases, the $p$ value was greater than the $\alpha$ value of 0.05 showing that there is no difference between the observed and predicted storage fluxes. Consequently, the current study did not find clear differences in the predictive ability of the minFluxes and MoMA objective functions.

**Table 5.3** – Comparison of the $T^2$ test between the minFluxes and MoMA objective functions for glucose, acetate, and their combination in the feast-famine growth of *R. jostii* RHA1 on glucose and acetate as the sole carbon sources

<table>
<thead>
<tr>
<th>Sample</th>
<th>Objective Function</th>
<th>$T^2$</th>
<th>$F$</th>
<th>df1</th>
<th>df2</th>
<th>$\chi^2$</th>
<th>$D_f$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>minFluxes$^a$</td>
<td>0.0092</td>
<td>0.0092</td>
<td>1</td>
<td>2</td>
<td>NA$^e$</td>
<td>NA$^e$</td>
<td>0.933</td>
</tr>
<tr>
<td></td>
<td>MoMA$^b$</td>
<td>0.0092</td>
<td>0.0092</td>
<td>1</td>
<td>2</td>
<td>NA$^e$</td>
<td>NA$^e$</td>
<td>0.933</td>
</tr>
<tr>
<td>Acetate</td>
<td>minFluxes$^a$</td>
<td>0.0003</td>
<td>0.0003</td>
<td>1</td>
<td>2</td>
<td>NA$^e$</td>
<td>NA$^e$</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>MoMA$^b$</td>
<td>0.000</td>
<td>0.000</td>
<td>1</td>
<td>2</td>
<td>NA$^e$</td>
<td>NA$^e$</td>
<td>1.000</td>
</tr>
<tr>
<td>Glucose + Acetate</td>
<td>minFluxes$^a$</td>
<td>1.972</td>
<td>NA$^e$</td>
<td>NA$^e$</td>
<td>NA$^e$</td>
<td>1.972</td>
<td>1</td>
<td>0.160</td>
</tr>
<tr>
<td></td>
<td>MoMA$^b$</td>
<td>1.982</td>
<td>NA$^e$</td>
<td>NA$^e$</td>
<td>NA$^e$</td>
<td>1.982</td>
<td>1</td>
<td>0.159</td>
</tr>
</tbody>
</table>

$^a$ Case ii: minimizing the Manhattan norm of the flux vectors while maximizing the growth rate.

$^b$ Case iii: minimizing the Manhattan norm of the difference between the fluxes over the feast-famine cycle while maximizing their growth rates.

$^c$ Hotelling’s $T^2$ value.

$^d$ Fisher-distribution value.

$^e$ Not Applicable.

$^f$ Degrees of freedom.

$^g$ $\chi^2$-distribution value.

$^h$ Probability value.

**Acknowledgments**

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant awarded to DF. MT is grateful to the McGill Engineering Doctoral Award (MEDA) for the financial support of his Ph.D. study. Labelled amino acid measurements were performed at the Rosalind and Morris Goodman Cancer Research Centre Metabolomics Core Facility supported by the Canada Foundation of Innovation, The Dr. John R. and Clara M. Fraser Memorial Trust, the Terry Fox Foundation (TFF Oncometabolism Team Grant 116128), and McGill University.
Supporting Information Available

Additional File 1 (Ch5_AF1.xlsx). Reactions involved in the $^{13}$C-MFA model of Rhodococcus jostii RHA1 on glucose and $^{13}$C-bicarbonate during the feast phase, along with its input file of the openFLUX software, and the list of the metabolites.

Additional File 2 (Ch5_AF2.xlsx). Reactions involved in the $^{13}$C-MFA model of Rhodococcus jostii RHA1 on glucose and $^{13}$C-bicarbonate during the famine phase, along with its input file of the openFLUX software, and the list of the metabolites.

Additional File 3 (Ch5_AF3.xlsx). Reactions involved in the $^{13}$C-MFA model of Rhodococcus jostii RHA1 on acetate and $^{13}$C-bicarbonate during the feast phase, along with its input file of the openFLUX software, and the list of the metabolites.

Additional File 4 (Ch5_AF4.xlsx). Reactions involved in the $^{13}$C-MFA model of Rhodococcus jostii RHA1 on acetate and $^{13}$C-bicarbonate during the famine phase, along with its input file of the openFLUX software, and the list of the metabolites.

5.4 References


Appendix

Figure A1 – The central metabolic pathways of *R. jostii* RHA1, along with the amino acid, storage, and biomass biosynthetic reactions.
Figure A2 – Contour plots of the weighted average maximum growth rate of *R. jostii* RHA1 vs. a pair of storage fluxes at glucose uptake rate of 156.01 Cmmol/(g-DW·h) (a, b, and c) and
acetate uptake rate of 79.51 Cmmol/(g-DW·h) (d, e, and f) at the biosynthetic fluxes of the following storage compounds: a: TAG, 8.81 Cmmol/(g-DW·h), b: glycogen, 13.93 Cmmol/(g-DW·h), c: PHB, 15.58 Cmmol/(g-DW·h), d: TAG, 3.71 Cmmol/(g-DW·h), e: glycogen, 0.44 Cmmol/(g-DW·h), f: PHB, 12.98 Cmmol/(g-DW·h).

Note that these fluxes are the experimentally observed storage biosynthetic fluxes. Here, the objective function was maximization of growth rate. The maximum capacity of all fluxes of the biochemical network is set unbounded. Note that the numbers appear on the isolines show the contour levels corresponding to the weighted average maximum growth rates (d⁻¹) over the feast and famine cycle.
**Figure A3** – Contour plots of the weighted average maximum growth rate of *R. jostii* RHA1 vs. a pair of storage fluxes at glucose uptake rate of 156.01 Cmmol/(g-DW·h) (a, b, and c) and
acetate uptake rate of 79.51 Cmmol/(g-DW·h) (d, e, and f) at the biosynthetic fluxes of the following storage compounds: a: TAG, 8.81 Cmmol/(g-DW·h), b: glycogen, 13.93 Cmmol/(g-DW·h), c: PHB, 15.58 Cmmol/(g-DW·h), d: TAG, 3.71 Cmmol/(g-DW·h), e: glycogen, 0.44 Cmmol/(g-DW·h), f: PHB, 12.98 Cmmol/(g-DW·h).

Note that these fluxes are the experimentally observed storage biosynthetic fluxes. Here, maximization of growth rate was the main objective function, and minimization of metabolic fluxes was the sub-objective function. The maximum capacity of all fluxes of the biochemical network is set unbounded. Note that the numbers appear on the isolines show the contour levels corresponding to the weighted average maximum growth rates (d⁻¹) over the feast and famine cycle.
Figure A4 – Contour plots of the weighted average maximum growth rate of *R. jostii* RHA1 vs. a pair of storage fluxes at glucose uptake rate of 156.01 Cmmol/(g-DW·h) (a, b, and c) and
acetate uptake rate of 79.51 Cmmol/(g-DW·h) (d, e, and f) at the biosynthetic fluxes of the following storage compounds: a: TAG, 8.81 Cmmol/(g-DW·h), b: glycogen, 13.93 Cmmol/(g-DW·h), c: PHB, 15.58 Cmmol/(g-DW·h), d: TAG, 3.71 Cmmol/(g-DW·h), e: glycogen, 0.44 Cmmol/(g-DW·h), f: PHB, 12.98 Cmmol/(g-DW·h).

Note that these fluxes are the experimentally observed storage biosynthetic fluxes. Here, maximization of growth rate was the main objective function, and MoMA was the sub-objective function. The maximum capacity of all fluxes of the biochemical network is set unbounded. Note that the numbers appear on the isolines show the contour levels corresponding to the weighted average maximum growth rates (d⁻¹) over the feast and famine cycle.
Connecting Text: Chapter 5 and Chapter 6

In chapter 5, we simulated the *in silico* growth of *R. jostii* RHA1 during the feast-famine cycles on glucose and acetate. Experiments with $^{13}$C-labeled bicarbonate ($\text{HCO}_3^-$) suggested that the fluxes through the central metabolism reactions during the feast phases were similar to the one during the famine phase. Furthermore, analysis of two of the sub-objective functions (i.e., (i) minimization of network fluxes and (ii) *environmental* minimization of metabolic adjustment (MoMA) between feast and famine phases) showed identical results on glucose, while the former was able to better predict the storage metabolism of *R. jostii* RHA1 on acetate. It showed PHB as the main storage pool on both glucose and acetate that was inline with our experimentally observed storage fluxes. So far, we have developed a computational tool to predict storage metabolism during the feast-famine cycle of the pure culture *R. jostii* RHA1; one needs to use this modeling tool and predict the storage metabolism of mixed microbial communities present in activated sludge bacteria. In chapter 6, we predict the storage metabolism of *R. jostii* RHA1 and *E. coli* K12, two model organisms of heterotrophs, during the feast-famine cycle on succinate. We also experimentally observe the dynamics of the storage metabolism of activated sludge bacteria during the feast-famine cycle on succinate which are detailed in chapter 6.
Chapter 6. Genome-Scale Metabolic Modeling to Provide Insight into the Production of Storage Compounds during Feast-Famine Cycles of Activated Sludge

Mohammad Tajparast and Dominic Frigon

Microbial Community Engineering Laboratory, Department of Civil Engineering and Applied Mechanics, McGill University, 817 Sherbrooke Street West, Montreal, Quebec, Canada, H3A 2K6, (Email: mohammad.tajparast@mail.mcgill.ca; dominic.frigon@mcgill.ca).

Published in

Water Science and Technology 2013, 67 (3), pp. 469-476
Abstract

Studying storage metabolism during feast-famine cycles of activated sludge treatment systems provides profound insight in terms of both operational issues (e.g., foaming and bulking) and process optimization for the production of value added by-products (e.g., bioplastics). We examined the storage metabolism (including poly-β-hydroxybutyrate [PHB], glycogen, and triacylglycerols [TAGs]) during feast-famine cycles using two genome-scale metabolic models: *Rhodococcus jostii* RHA1 (iMT1174) and *Escherichia coli* K-12 (iAF1260) for growth on glucose, acetate, and succinate. The goal was to develop the proper objective function (OF) for the prediction of the main storage compound produced in activated sludge for given feast-famine cycle conditions. For the flux balance analysis, combinations of three OF were tested. For all of them, the main OF was to maximize growth rates. Two additional sub-OFS were used: (1) minimization of biochemical fluxes, and (2) minimization of metabolic adjustments (MoMA) between the feast and famine periods. All (sub-)OFS predicted identical substrate-storage associations for the feast-famine growth of the above-mentioned metabolic models on a given substrate when glucose and acetate were set as sole carbon sources (i.e., glucose-glycogen and acetate-PHB), in agreement with experimental observations. However, in the case of succinate as substrate, the predictions depended on the network structure of the metabolic models such that the *E. coli* model predicted glycogen accumulation and the *R. jostii* model predicted PHB accumulation. While the accumulation of both PHB and glycogen was observed experimentally, PHB showed higher dynamics during an activated sludge feast-famine growth cycle with succinate as substrate. These results suggest that new modeling insights between metabolic predictions and population ecology will be necessary to properly predict metabolisms likely to emerge within the niches of activated sludge communities. Nonetheless, we believe that the development of this approach will help guide the optimization of the production of storage
compounds as valuable by-products of wastewater treatment.

**Keywords**
Biomaterials, *Escherichia coli* strain K-12, genome-scale flux balance analysis, *Rhodococcus jostii* RHA1, storage metabolism, wastewater treatment operation
6.1 Introduction

Storage metabolism plays a key role in maintaining stable operation of activated sludge treatment systems as it is partly linked to bulking and foaming problems (Martins et al. 2004). Under the well described feast-famine cycles in these systems, many bacterial species produce storage compounds (including polyhydroxyalkanoates [PHAs], glycogen, and triacylglycerols [TAGs]) that serve as intracellular carbon and energy reserves (van Loosdrecht et al. 1997). Since storage compounds are also biomaterial and biofuel precursors (e.g., bioplastics and biodiesel), wastewater treatment systems also offer the possibility of becoming biomaterial recovery technologies, a focus of active research in the last few years (Kleerebezem and van Loosdrecht 2007). However, predicting and optimizing the production of these compounds remain elusive and a matter of experimental trial and error.

In the last decade, the genome of several bacterial species was published, and for a number of these datasets, the cellular stoichiometry was compiled from the genome annotation and simulated; this process is called genome-scale metabolic modeling and flux balance analysis (FBA). Basically, FBA solves the underdetermined mass balance model for unknown reaction fluxes by optimizing a metabolic objective function subject to constraints at a given steady-state condition. We propose that this is a promising approach to predict and optimize storage compound production in activated sludge. At the center of this approach, it is assumed that the specific substrates and feeding conditions will generate specific environmental niche characterized by the accumulation of a single specific storage compound as the literature suggests for acetate and glucose (Beun et al. 2000; Dircks et al. 2001). Without speculating on the species that will occupy these niches, we suggest that we could compare the results from the genome-scale models of several organisms to obtain valid predictions on the heterotrophic metabolism to emerge in a niche from mixed ordinary heterotrophs. Therefore, FBA could serve
as a potential screening tool to evaluate production of value-added by-products by various wastewaters after a determination of the main compounds contributing to the COD load (e.g., slaughterhouse and/or pulp-and-paper effluent). The development of such systems for carbon recovery from wastewater enhances plant sustainability both from environmental and economic points of view.

As a first attempt to this approach, we examined the predicted storage metabolism during the feast-famine cycle using two genome-scale metabolic models: *Rhodococcus jostii* RHA1 (model iMT1174) (Chapter 3, (Tajparast and Frigon in preparation)) and *Escherichia coli* K-12 (model iAF1260) (Feist et al. 2007). The choice of an objective function is central to FBA, and three combinations were tested. (1) The well-known optimization of the growth rate (Feist and Palsson 2010) was applied alone to both the feast and famine phases simulated independently. (2) In conjunction with growth rate maximization, minimization of the total metabolic fluxes (Holzhutter 2004) in both phases independently (i.e., continuous maximal cellular efficiency) was applied as a sub-objective function. (3) Also in conjunction with growth rate maximization, minimization of the flux differences between feast and famine growth (i.e., minimization of metabolic adjustments, MoMA) (Segre et al. 2002) was used as another sub-objective function. Simulation results were then compared with experimental observations of acetate (Beun et al. 2000), glucose (Dircks et al. 2001), and succinate (this study).

### 6.2 Materials and Methods

#### 6.2.1 Genome-scale Metabolic Models and Flux Balance Analysis

Two genome-scale metabolic models were used: *R. jostii* RHA1 (iMT1174) (Chapter 3, (Tajparast and Frigon in preparation)) and *E. coli* K-12 (iAF1260) (Feist et al. 2007). The *E. coli* K-12 model was modified to include the storage pathways of PHB and TAG. The FBA study
tested maximization of growth rates as the main objective function alone or in conjunction with two sub-objective functions: minimization of metabolic fluxes (Holzhutter 2004) and minimization of metabolic adjustments (MoMA, (Segre et al. 2002)). Basically, the model was studied quantitatively by setting up a series of mass balances around each metabolite that can be expressed in matrix notation as: \( \frac{dx}{dt} = S \cdot V \) where \( x \) is the \((m \times 1)\) vector of the concentrations of the balanced metabolites (i.e., intracellular and macromolecular compounds), \( V \) denotes the \((n \times 1)\) vector of metabolic fluxes, and \( S \) stands for the \(m \times n\) stoichiometric matrix. The model is then solved by assuming that the cellular network is at steady-state, which simplifies the previous equation to \( S \cdot V = 0 \). This assumption is plausible due to the fast equilibration of intracellular metabolite pools (time-scale of seconds) compared to the time-scale of genetic regulation (minutes) (Segre et al. 2002) and due to high reaction rates of intracellular reactions compared to exchange rates such as substrate uptake, cell growth, and by-product secretion rates (Lee et al. 2006; Roels 1982; Varma and Palsson 1994). Consequently, the output of the model is a distribution of metabolic fluxes through the various chemical reactions (note that the kinetics is not modeled). Since the number of fluxes normally exceeds the number of metabolites \((n > m)\), the problem is said to be underdetermined. In this case, there is a solution space for the fluxes that can be studied by optimizing proper objective functions subject to defined constraints (the mass balance equations as the physicochemical constraints and the inequalities as the enzymatic capacity of the biochemical reactions) (Becker et al. 2007). Simulations were conducted using the COBRA toolbox of MATLAB (Becker et al. 2007).

Simulating the continuity between the feast and famine phases of a cycle, we needed to conceptualize a simplified scenario. We simulated the two phases of growth independently (except when the MoMA objective function was applied). Following experimental observations,
we considered that the feast phase lasted only 1.25% for glucose (Dircks et al. 2001), 5.83% for acetate (Beun et al. 2000), and 15% for succinate (experiments conducted in this study) of the cycle time and was characterized by a single flux distribution. Finally, the mass balance on the storage compound simulated was constrained by equating the storage synthesis rate in the feast phase and the storage consumption rate in the famine phase weighted according to their respective proportion of the cycle. The simulated growth rates or specific fluxes are the weighted averages of those of the feast and famine phases. Note that storage metabolism was simulated one at a time; that is, at a given storage condition, synthesis flux of a given storage compound was fixed, while the synthesis fluxes of the other storage pools were set to zero in the feast phase; moreover, storage compounds were excluded from biomass and treated as virtual extracellular compounds.

6.2.2 Lab-Scale Activated Sludge Reactor Fed Succinate

Reactor operation. An Infors-HT model Labfors 2 lab-scale reactor with a 2-L working volume vessel was used. The system was equipped with temperature, dissolved oxygen (DO), pH, and level sensors. Stirring (300 rpm), temperature (set at 20°C), pH (set at 7, controlled by adding 1 M NaOH or 1 M H₂SO₄), aeration airflow (1.5 L/min), feed and dilution water additions, and solids and effluent withdrawals were automatically controlled using the Iris software (v.5; Infors-HT, Switzerland). The reactor was operated in sequencing batch reactor (SBR) mode with a hydraulic residence time (HRT) of 12 h and a solid residence time (SRT) of 1.5 d. The reactor was inoculated with mixed liquor suspended solids (MLSS) from a nearby full-scale conventional activated sludge plant. Continuous operation of the SBR was based on an 8-hour cycle with the following phases: 11 min start period, 14.58 min dilution water addition (1.29 L), 4 min feed addition (36 mL), 404.74 min reaction, 4.85 min excess solids withdrawal to
maintain SRT, 31 min settling, and 9.83 min effluent removal. The steady-state operation of the SBR was achieved when, for at least 5 cycles, the length of the feast period as determined by the DO concentration time profile was invariant.

The medium composition (after combination of feed and dilution water) was similar to previous studies (Beun et al. 2000): succinic acid (3.6 mM, equivalent to 400 mg-COD/L), NH$_4$Cl (2.00 mM), KH$_2$PO$_4$ (1.65 mM), MgSO$_4$.7H$_2$O (0.37 mM), KCl (0.50 mM), and trace element solution (Vishniac and Santer 1957) (1 mL/L). Allylthiourea (10 mg/L) was added to inhibit nitrification.

**Biochemical analysis.** At steady-state, oxygen and CO$_2$ concentrations in the off-gas were monitored using a multi-gas meter (model ATX620, Industrial Scientific Corporation). Several samples were taken during a cycle to determine biochemical constituents according to previously published protocols: total and volatile suspended solids (TSS/VSS) (standard method 2540 D/2540 E; (APHA et al. 1998)), chemical oxygen demand (COD) (standard method 5220 D; (APHA et al. 1998)), NH$_4^+$ concentration by colorimetry ((Rhine et al. 1998); spectrophotometric measurements at 630 nm), glycogen by a hexokinase enzymatic kit and colorimetry (Hexokinase protocol, measurement of NADH concentration at 340 nm, Sigma, St. Louis, MO) (Maurer et al. 1997), and PHB by colorimetry (Law and Slepecky 1961; Paganelli et al. 2011); spectrophotometric measurements at 235 nm). Spectrophotometric measurements of the last three components were performed in microplate using a SpectraMax5 reader (Molecular Devices, LLC, USA).
6.3 Results and Discussion

6.3.1 Genome-Scale Flux Balance Analysis

The development of the modeling approach focused first on determining the procedure to simulate the feast and famine phases of a culture. The approach suggested in previous lumped-reaction metabolic modeling exercises was followed, in which the feast and famine phases were analysed independently (Beun et al. 2000; Dircks et al. 2001). Within this context, the total growth rate (equal to 1/SRT) was the sum of the growth rates calculated in both phases weighted by the proportion of the length of each phase, and a single metabolic condition was assumed to take place in each phase (i.e., a single flux profile occurs in each phase). Because the approach does not consider the dynamic conditions occurring in the reactor at this point, it was necessary to assume a proportion of the cycle for each phase. Depending on the carbon source, different time ratios of the feast phase ($t_{Feast}/t_{Cycle}$) were applied in the simulation studies, such as glucose 1.25%, acetate 5.83%, and succinate 15%. These were similar to the experimentally observed $t_{Feast}/t_{Cycle}$ ratios: acetate 3.33-5.83% (Beun et al. 2000), glucose 0.83-1.25% (Dircks et al. 2001), and succinate 15.10% (below). Note that we presented the simulation results at a constant growth rate depending on the substrate examined according to the experimentally observed overall feast-famine growth rates as follows: glucose 0.2844 d$^{-1}$, acetate 0.2691 d$^{-1}$, and succinate 0.7600 d$^{-1}$ (Figures 1 and 2). With these assumptions, the space of possible storage fluxes was systematically investigated for the two metabolic models and the three substrates by setting substrate uptake rates and storage fluxes (i.e., non-zero for a given storage pool and zero for the two others). For a constant substrate uptake rate, only a subset of storage fluxes were possible for positive growth to occur in both phases (Figures 1 and 2); below the minimum storage flux, positive growth would not occur in the famine phase, and above the maximum flux, positive growth would not occur in the feast phase. According to the literature, positive growth rates
occur in both phases of the feast-famine cycle (Beun et al. 2000; Dircks et al. 2001). Furthermore, since the storage compounds are not part of the biomass, negative growth would represent a net consumption of biomass that would presumably occur only in the absence of storage compounds, an unlikely situation for storage accumulation metabolisms studies here. Finally, note also that the range of possible storage fluxes (x-axis) was different on the various substrate-storage cases, and it was dependent on the substrate, storage, metabolic model, and maintenance energy values (as presented in Figures 1, 2, and 3). For instance, the range for TAG was very small because of the lost of carbon during the feast phase due to the degree of reduction of this storage compound compared to the substrates investigated. Thus, we did not see any TAG accumulation for growth on acetate in the case of R. jostii model and the maintenance energy values presented (Figure 6.1b,d,f).
Figure 6.1 – Comparison of the overall biomass yield of *R. jostii* RHA1 (a and b), the sum of the metabolic fluxes, $|V_{\text{Total}}|$ (c and d), and the total metabolic difference between the feast and famine phases, $|V_{\text{Feast}} - V_{\text{Famine}}|$ (e and f), vs. the specific storage synthesis flux of three storage compounds (simulated one at a time): glycogen, PHB, and TAG at constant growth rate.

The substrates examined are glucose and acetate for which the constant growth rates 0.2844 d$^{-1}$ and 0.2691 d$^{-1}$ are adapted, respectively. Note that the maximum capacity of all fluxes of the
biochemical network is set unbounded; the applied t_{Feast}/t_{Cycle} ratios are: glucose 1.25% and acetate 5.83% which gave rise to different growth rate intervals. Note that the x- and y-axis do not start at zero; however, the y-axis ranges are the same for a given objective function.

Model simulations showed a decrease in the biomass yield with increase in storage flux (Figure 6.1a,b and Figure 6.2a,b, Figure 6.3a,b). This decrease corresponds to the increasing cost for the conversion of substrate to storage before conversion to biomass. Simulation results for glucose and acetate obtained with the *R. jostii* or the *E. coli* models showed similar trends; thus, only the results for the *R. jostii* model are presented (Figure 6.1). Irrespective of the objective function and the models used, the metabolic simulations showed that PHB accumulation optimized the objective functions with acetate as substrate for both modeled ordinary heterotrophs, while accumulation of glycogen optimized them with glucose (Figure 6.1). It is interesting to note that the models predicted well the higher biomass yield on glucose compared to acetate, which is due to the higher energy available for growth on glucose over acetate. Note that the full oxidation of 1 electron equivalent of glucose and acetate releases 41.35 kJ and 27.40 kJ, respectively.

The case is different with succinate as substrate. The predicted associations were the same for all objective functions when the *R. jostii* model was used: PHB accumulation optimize the objective functions over the range of possible storage fluxes (Figure 6.2a,c,e). That is, the yields were highest (Figure 6.2a), and the sum of fluxes (Figure 6.2c) or the feast/famine MoMAs (Figure 6.2e) were lowest with PHB accumulation for *R. jostii* model simulations. However, in the case of the *E. coli* model, glycogen was the storage compound that optimized the objective functions on succinate (Figure 6.2b,d,f) for all objective functions examined. That is, glycogen accumulation tended to maximize the *E. coli* yield on succinate (Figure 6.2b), and
minimize both the total metabolic fluxes (Figure 6.2d) and MoMA (Figure 6.2f). Thus, the two metabolic models had different predictions for succinate as substrate. These results could be due to either (1) differences in biomass compositions between the models (*R. jostii* contains a lot more lipids than *E. coli* due to the presence of mycolic acids) or (2) differences in the metabolic structure (i.e., the reactions present and their connections). To test the first possibility, the biomass compositions of *R. jostii* RHA1 and *E. coli* were switched between models in order to see any effect on storage accumulation on succinate as the sole carbon source. Simulation results with respect to the order of the objective function values remained the same as previously (data not shown, yet similar to Figure 6.2). Therefore, we conclude from these simulations that the optimal storage compound for each of these two organisms is mainly due to the structure of their metabolic network.
Figure 6.2 – Comparison of the overall biomass yield of *R. jostii* RHA1 and *E. coli* (a and b), the sum of the metabolic fluxes, \(|V_{\text{Total}}|_1\) (c and d), and the total metabolic difference between the feast and famine phases, \(|V_{\text{Feast}} - V_{\text{Famine}}|_1\) (e and f), vs. the specific storage synthesis flux of three storage compounds (simulated one at a time): glycogen, PHB, and TAG at constant growth rate of 0.7600 d\(^{-1}\).
The substrate examined is succinate. Note that the maximum capacity of all fluxes of the biochemical network is set unbounded; the applied $t_{\text{Feast}}/t_{\text{Cycle}}$ ratio is 15%. Note that the x- and y-axis do not start at zero; however, the y-axis ranges are the same for a given objective function.

It should be noted that the non-growth associated maintenance energy (NGAME) was set to 0.5 mmol/(g-DW·h) for all simulation results presented (Figure 6.1 and Figure 6.2). A number of simulations at which the NGAME was set to 2 mmol/(g-DW·h) were conducted in order to elucidate its effect on the predictions. The *R. jostii* model showed accumulation of TAG for this elevated maintenance energy (Figure 6.3a) as opposed to those simulated at the NGAME of 0.5 mmol/(g-DW·h) (Figure 6.2a). Similarly for the *E. coli* model, the window of possible glycogen fluxes, that was very small at a low NGAME value (Figure 6.2b), was extended at higher NGAME value (Figure 6.3b). Despite the changes in the window of possible storage compound fluxes, the metabolic values of the storage compounds (i.e., order of objective function values) stayed identical with all applied NGAME values. That is, the *R. jostii* model and the *E. coli* model predicted accumulation of PHB and glycogen on succinate, respectively.

![Graph](image)

**Figure 6.3** – Comparison of the overall biomass yield of *R. jostii* RHA1 and *E. coli* (a and b) vs. the specific storage synthesis flux of three storage compounds (simulated one at a time): glycogen, PHB, and TAG at constant growth rate of 0.7600 d$^{-1}$. 

152
The substrate examined is succinate. Note that the maximum capacity of all fluxes of the biochemical network is set unbounded and the non-growth associated maintenance energy is 2 mmol/(g-DW·h); the applied $t_{\text{Feast}}/t_{\text{Cycle}}$ ratio is 15%. Note that the x- and y-axis do not start at zero.

6.3.2 Lab-Scale Activated Sludge Reactor Fed Succinate

The ambiguity in the predictions on succinate is interesting because it can serve to evaluate the validity of the approach and eventually improve it. We experimentally observed the dynamics of the storage pools and other mixed liquor constituents during the feast-famine growth cycles of activated sludge on succinate as the sole carbon source (Figure 6.4). The soluble COD curve showed that the feast phase lasted about 72.5 min, corresponding to $t_{\text{Feast}}/t_{\text{Cycle}}=15.10\%$. While our analytical techniques detected the presence of both PHB and glycogen in the mixed liquor, the higher dynamics of PHB suggest that it is the main storage compound (Figure 6.4b). It is also possible that biomass components other than glycogen contribute to the glucose level measured in the solids.
Figure 6.4 – Analysis of the dynamics of mixed liquor constituents during a steady-state 8-hour cycle of the activated sludge reactor fed succinate (a). (b) Weight percentage of glycogen and PHB in $X_{OHO}+X_{U,E}$.

Note that amounts of active ordinary heterotrophic biomass+endogenous residue ($X_{OHO}+X_{U,E}$) and soluble COD were divided by 10 and 100, respectively.

6.3.3 Comparison between Flux Balance Analysis Predictions and Experimental Observations

Table 6.1 shows the comparison between predictions by the three objective functions obtained for the two models and the experimental observations for a single storage compound accumulated during the feast-famine cycle growth on a sole substrate. The three substrates simulated in this study were glucose, acetate, and succinate. The composition of the storage pool was experimentally observed for the first two substrates in the previous studies (Beun et al. 2000;
Dircks et al. 2001), and the experiments for succinate were conducted here. All (sub-)OFs predicted identical substrate-storage associations for the feast-famine growth of both models on a given substrate when glucose and acetate were set as sole carbon sources (i.e., glucose-glycogen and acetate-PHB), in agreement with experimental observations. However, in the case of succinate the predictions depended on the network structure of the metabolic models such that the E. coli model predicted glycogen accumulation and the R. jostii model predicted PHB accumulation on succinate. According to our experimental observations, PHB showed higher dynamics as compared to glycogen during an activated sludge feast-famine growth cycle on succinate. This result points to the need to obtain further insights on the links between bacterial physiology and population ecology in order to properly predict the metabolisms likely to emerge in niches of activated sludge microbial communities. The partial success in predicting the storage compounds accumulated on acetate, glucose, and succinate is encouraging. We believe that the development of this approach will help guide the optimization of the production of storage compounds as valuable by-products of wastewater treatment.

Table 6.1 – Comparison between predictions by the three objective functions and experimental observations of single storage compounds accumulated during the feast-famine cycle growth on single substrates

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Max. Growth Rate</th>
<th>Max. Growth Rate + Min. Metab. Fluxes</th>
<th>Max. Growth Rate + MoMA</th>
<th>Experimental Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. jostii</td>
<td>E. coli</td>
<td>R. jostii</td>
<td>E. coli</td>
<td>R. jostii</td>
</tr>
<tr>
<td>Acetate, $t_{\text{Feast}}/t_{\text{Cycle}} = 5.83%$, $\mu_{\text{Total}} = 0.2691 \text{ d}^{-1}$ [1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prediction</td>
<td>PHB</td>
<td>PHB</td>
<td>PHB</td>
<td>PHB</td>
</tr>
<tr>
<td>OF Diff. [%]$^\dagger$</td>
<td>3.10</td>
<td>6.27</td>
<td>28.85</td>
<td>15.11</td>
</tr>
<tr>
<td>Glucose, $t_{\text{Feast}}/t_{\text{Cycle}} = 1.25%$, $\mu_{\text{Total}} = 0.2844 \text{ d}^{-1}$ [2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prediction</td>
<td>Glycogen</td>
<td>Glycogen</td>
<td>Glycogen</td>
<td>Glycogen</td>
</tr>
<tr>
<td>OF Diff. [%]$^\dagger$</td>
<td>0.82</td>
<td>1.15</td>
<td>4.96</td>
<td>5.36</td>
</tr>
<tr>
<td>Succinate, $t_{\text{Feast}}/t_{\text{Cycle}} = 15%$, $\mu_{\text{Total}} = 0.7600 \text{ d}^{-1}$ [3]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OF Diff. [%]$^\dagger$</td>
<td>3.68</td>
<td>10.57</td>
<td>11.08</td>
<td>13.99</td>
</tr>
</tbody>
</table>

$^*$: References: [1] (Beun et al. 2000); [2] (Dircks et al. 2001); [3] (This study).

$^\dagger$: OF Difference was computed at the lowest storage flux between glycogen and PHB [%], this flux was used because this storage flux was possible for all the investigated storage compounds.
6.4 Conclusions

The genome-scale metabolic modeling framework explicitly interprets genomic information in terms of rates and yields. This modeling approach can be potentially applied in various biological wastewater treatment systems such as carbon removal, enhanced biological phosphorus removal (EBPR) and nitrogen removal. As one application, we showed in this paper how one could use the information from several species to infer the metabolism of heterotrophic bacteria in activated sludge microbial communities and predict the accumulation of storage compounds. We propose that this is a promising approach to predict and optimize biological processes even in mixed culture. For example, it could act as a screening tool to evaluate the potential of storage compound production by various wastewaters after a determination of the main compounds contributing to the COD load (e.g., slaughterhouse and/or pulp-and-paper effluent). Once fully developed, this modeling approach may become an additional tool for the developers of new wastewater treatment technologies like the ones focussing on the production of valuable biomaterials.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada Discovery Program and by a McGill Engineering Doctoral Award to MT. The paper was presented at the 3rd IWA-WEF Wastewater Treatment Modeling Seminar (WWTmod 2012) and the fruitful discussions are kindly acknowledged.

6.5 References


Tajparast M, Frigon D. in preparation. Genome-scale metabolic model of Rhodococcus jostii RHA1 (iMT1174) to study the accumulation of storage compounds during nitrogen limited growth. BMC Systems Biology.


Chapter 7 . Summary

7.1 Intellectual Contributions

Studying storage metabolism during nitrogen-limited substrate consumption and feast-famine growth cycles is important since it can provide insight into the storage production for biotechnology applications. It also provides new data on the ecophysiology of heterotrophic bacteria under highly dynamic growth conditions that happen in soil and in activated sludge treatment systems. The contributions of this study to the knowledge are summarized as follows.

1. This study presents the first in silico reconstructed genome-scale metabolic network of the model organism of heterotrophs, R. jostii RHA1.

   The genome-scale metabolic network of R. jostii RHA1 reconstructed in this study includes 1243 balanced (intracellular) metabolites, 1935 unique reactions, and 1174 open reading frames (ORFs). Totally, it also includes 330 extracellular compounds that are associated with 518 exchange reactions. The model also contains all seven storage compounds and the study emphasized on metabolism of glycogen, PHA and TAG. In addition, we made a special effort to include a number of xenobiotics degradation pathways (consisting of 131 transport reactions and 326 conversion reactions) in order to make the model most useful for the research community interested in R. jostii RHA1.

   Different objective functions were applied to solve the underdetermined set of mass balance equations to obtain the unknown fluxes. For the first time the minimization of metabolic adjustment objective function was applied in the environmental context comparing two environmental conditions including non-limited growth followed by a sudden nitrogen-limited substrate consumption.
2. This study provides experimental data on the nitrogen- and non-limited conditions and feast-famine growth cycles of *R. jostii* RHA1 on glucose and acetate.

We have investigated *R. jostii* RHA1 phenotypic behaviours on glucose and acetate in the nitrogen- and non-limited conditions and the feast-famine growth cycles. These data were not available in the literature and it is the first time that they have been produced.

According to our observations, although a mixture of storage compounds (i.e., PHA, TAG, and glycogen) was experimentally observed on both glucose and acetate in the nitrogen-limited condition, PHA turned out to posses the highest storage yield followed by TAG on both substrates.

When glucose was the substrate, the COD accumulation during the feast phase was similar for the three storage compounds: PHB, TAG, and glycogen. However, when acetate was the substrate, essentially no glycogen was accumulated and the COD yield towards PHB accumulation was 3 times higher than toward TAG accumulation.

3. This study presents $^{13}$C-metabolic flux analysis of the central carbon metabolism of *R. jostii* RHA1 in the non-limited and feast-famine growth conditions.

$^{13}$C-labeling experiments have been performed to elucidate the conversion fluxes of the central carbon metabolic pathways present in the *R. jostii* RHA1 network in both non-limited and feast-famine conditions on labeled glucose and acetate in the non-limited growth condition as the sole carbon sources and $^{13}$C-bicarbonate, in association with unlabelled glucose and acetate, in the feast/famine phases. For the first time, a $^{13}$C-metabolic flux analysis approach has been applied in order to investigate the central metabolic pathways of *R. jostii* RHA1. This was used to support the use of the minimization of metabolic adjustment (*environmental* MoMA) sub-
objective function in the environmental context namely comparing the N- and non-limited growth conditions and the feast and famine growth phases. In both environments, $^{13}$C-MFA showed similarities in metabolic behaviours of *R. jostii* RHA1 cells pair wise. Therefore, the *environmental* MoMA sub-objective function can be applied in the environmental context such as feast-famine cycles and nitrogen- and non-limited conditions.

4. This study provides insight into the production of storage compounds during feast-famine growth cycles of activated sludge bacteria.

In this study, we examined the storage metabolism (including glycogen, PHB, and TAG) during the feast-famine cycles using two genome-scale metabolic models: *R. jostii* RHA1 and *Escherichia coli* K-12 for growth on glucose, acetate, and succinate. We examined three objective functions including maximization of the growth rate, minimization of metabolic fluxes, and minimization of metabolic adjustment (*environmental* MoMA) between the feast and famine phase. The last two objective functions were implemented as two sub-objective functions, in conjunction with the maximization of the growth rate as the main objective function.

All (sub-) objective functions predicted identical substrate-storage associations for the feast-famine growth of the above-mentioned metabolic models on a given substrate when glucose and acetate were set as sole carbon sources (i.e., glucose-glycogen and acetate-PHB), in agreement with experimental observations. However, in the case of succinate as substrate, the predictions depended on the network structure of the metabolic models such that the *E. coli* model predicted glycogen accumulation and the *R. jostii* model predicted PHB accumulation. While the accumulation of both PHB and glycogen was observed experimentally, PHB showed higher dynamics during an activated sludge feast-famine growth cycle with succinate as
substrate. Nonetheless, we believe that the development of the genome-scale metabolic modeling approach will help guide the optimization of the production of storage compounds as valuable by-products of wastewater treatment.

7.2 Future Work

In this study, we individually studied the genome-scale metabolic modeling of two heterotrophic microorganisms namely *R. jostii* RHA1 and *E. coli* using a flux balance analysis approach. However, in addition to the bacteria interactions with the environment, they interact with each other in the complex activated sludge system. Therefore, one needs to take into account the meta-genomics data and build a master metabolic network to see the interactions among species present in soil and activated sludge. In this case, one needs to assemble common metabolic pathways of selected microorganisms into a unique master metabolic network, and then add peripheral pathways that are specifically present in individual species as compartments connecting to the master metabolism and/or to each other to account for the microbial interactions.

In this study, the genome-scale metabolic model was applied to study the metabolism of carbon storage compounds during the feast-famine growth cycles and nitrogen- and non-limited conditions. Specifically, we studied the metabolism of glycogen, PHB, PHV, and TAG. One should study the metabolism of the other carbon storage compounds such as α,α-trehalose and wax esters, and non-carbon storage pools such as polyphosphate all together.

In this study, the 13C-metabolic flux analysis was applied to study the central metabolic pathways of *R. jostii* RHA1. Here, we included a few lumped reactions to account for
biosynthetic pathways of amino acids; moreover, the metabolic reactions in the energy metabolism pathways were not considered. In the future, one should develop the network by adding more reactions to account for the detailed amino acid biosynthetic pathways and the energy metabolism to provide more details in the metabolism. Ultimately, one can apply a genome-scale $^{13}$C-metabolic flux analysis of *R. jostii* RHA1 or any other model organism to study cellular metabolism in environmental conditions such as soil and activated sludge treatment systems.

Finally, in this study we *in silico* investigated steady-state conditions. One can apply dynamic flux balance analysis and $^{13}$C-metabolic flux analysis to better understand the dynamics of biological systems. Finally, one can employ dynamic labeling experiments to better understand the bacterial metabolism during the feast-famine cycles.