

Optimization methods and objective functions for analysis of metabolic network models

Présentée le 14 décembre 2020

à la Faculté des sciences de base
Laboratoire de biotechnologie computationnelle des systèmes
Programme doctoral en chimie et génie chimique

pour l'obtention du grade de Docteur ès Sciences

par

Zhaleh HOSSEINI

Acceptée sur proposition du jury

Prof. K. Sivula, président du jury
Prof. V. Hatzimanikatis, directeur de thèse
Prof. R. Senger, rapporteur
Dr L. Küpfer, rapporteur
Prof. M. Dal Peraro, rapporteur

Acknowledgements

First of all, I would like to express my sincerest gratitude to my Ph.D. supervisor, Vassily Hatzimanikatis, for all the opportunities he has given to me, for his support during difficult situations and for all the scientific and life advice. I have learnt a lot from you and I am very grateful of your guidance and scientific enthusiasm during my Ph.D. time.

I would also like to thank the members of my Ph.D. defense committee: Prof. Kevin Sivula, Dr. Lars Küpfer, Prof. Ryan Senger and Prof. Matteo Dal Peraro. Thank you for the insightful review of my thesis, enjoyable discussion during my defense and constructive comments.

I am also very grateful to my collaborators in Environmental Microbiology Laboratory at EPFL, Prof. Rizlan Bernier-Latmani and Dr. Cornelia List. Thank you for providing me with the experimental data and for your ideas and enthusiasm in the joint projects.

I am also thankful to Christine Kupper and Anne-Lene Odegaard for all the help and administrative support.

I would like to thank my wonderful colleagues at LCSB: Many thanks to Misko and Noushin for all the personal and scientific support and advice. I am very grateful to Homa, for being a great friend, for all the coffee and tea time, for endless discussions about everything, and for all your support and kindness during these years. I would also like to thank Jasmin and Anush for wonderful collaborations, for your help and your friendship. Thank you Sofia, Maria, Milenko and Tuure for being great friends and for always being there for small chats and advice. My special thanks go to Meric and Georgios for all the help and advice specially during the first years of my Ph.D. Thank you Daniel for being a great officemate and friend. I would also like to thank Omid for our collaborations in different projects. My big thank goes to the rest of my colleagues and friends: Liliana, Asli, Evangelia, Anastasia, Shakiba, Kiana, Pierre, Robin, Aarti, Yves and Joana for making every day of my Ph.D. life enjoyable.

I am also grateful to my wonderful students, Hugo and Dana. Working with you was a great experience for me.

I am also very thankful to all my teachers and professors, specially my master thesis supervisor Prof. Sayed-Amir Marashi for introducing me to the field of computational biology and for everything he taught me.

My deepest gratitude goes to my beloved friends, Mahdieh and Bahareh. Although you were not always with me but you were always there for me. I am very thankful for all the small and big talks, for the trips we went together, for being able to share every pleasant or unpleasant experiences with you and for all the love and support.

Having a wonderful roommate was one of the most important reasons I could really enjoy living in a foreign country. Thank you Fatemeh for all the great time we spent together and for being such an amazing and caring friend, for all your help since my very beginning days in Lausanne and for so many great memories and fun that I've had with you since then.

I would also like to thank my Iranian friends in Lausanne, whom I cannot name all here. You made living in a foreign country much more enjoyable. My special thanks to Parima, Niloufar, Amirhosein, Ashkan, Mahsa, Aida, Ali, Mohammad, Ahmad, Fatemeh Q., Atena, Bahareh, Farnaz F., Sara, Bahar, Reza and Morteza for all the trips, talks, parties, games and fun we had together.

Last but not least, I would like to thank my family with all my heart. Above all, I thank my parents, for always believing in me and for your unconditional love. I couldn't come this far without your help. I also thank my brothers, Amir and Iraj, and my sister-in-laws, Maryam and Shadi, and my nieces, Zahra and Della, for always being there for me, making me happy and supporting me in every steps of my life.

Finally, I would like to thank Amir, for being so supportive and caring in the final steps of my Ph.D. You kept me going through all the ups and downs and brought energy and love to my life. Thank you very much for making my life full of happiness and wonderful.

Abstract

Computational studies of metabolism aim to systematically analyze the metabolic behaviour of biological systems in different conditions. Reconstruction of genome-scale metabolic network models (GEMs) using genome annotations, literature data, and enzymatic wet-lab experiments is an important step toward the systematic study of metabolism. These models capture the interconnection between different elements of the network by applying stoichiometric balances while taking into account gene-protein-reaction associations. Several methods have already been developed for the analysis of metabolic network models. These methods are mainly used to optimize a single or combination of biologically relevant objective functions subject to stoichiometric, thermodynamic, and other essential constraints. In this thesis, we have summarized the optimization methods and objective functions by classifying them based on biological and mathematical features. Particularly, we suggest reformulations to convert some of the complex optimization classes to simpler ones. One of the interesting reformulations is the conversion of mixed-integer linear fractional programming (MILFP) to mixed-integer linear programming (MILP). We show that this conversion is specifically useful in studying coupling relationships in metabolic network models that are thermodynamically constrained. Coupling is an important concept in metabolic networks that determines how different components of the network such as metabolites or reactions are interrelated. Particularly, flux Coupling Analysis (FCA) is a method, which has been used extensively for evaluating the dependencies between metabolic reactions. FCA has been exploited for several applications such as studying gene essentiality, network evolution, etc. In FCA, two reactions are considered as coupled if the activity of one, constrains the activity of the other. So far, FCA has been used for analyzing metabolic reactions in flux-balanced models. In this work, we developed a new formulation, Thermodynamic Flux Coupling Analysis (TFCA), which calculates flux couplings of metabolic models that are subjected to thermodynamic constraints. With TFCA, we show that adding thermodynamic constraints to a flux-balanced model can significantly change the coupling relationship of reactions of the network. Such thermodynamic constraints can assign directionalities to the reactions and hence reduce the

number of bidirectional reactions (BDR). Decreased BDRs subsequently result in a significant reduction in the solution space. We show that calculating coupling relations helps in reducing the number of combinations of BDRs, which in turn will facilitate the analysis of the metabolic network. In addition to proposing several mathematical reformulations to gain global optimality, we also addressed the issue of finding the proper cellular objective function in different conditions of cellular metabolism. This is not always straight-forward, since the metabolic activities of some organisms are not well-characterized. Metabolism of dormancy phase in some bacteria and parasites is an example of a poorly characterized biological system. In this thesis, we studied the metabolic behaviour of dormant malaria parasite using genome-scale model of *Plasmodium falciparum*. We examined several known and novel objective functions and scored them based on the model's consistency with experimental gene expression data. Our results suggested that minimizing energy dissipation can best describe the metabolic activities of the malaria parasites in the dormancy phase.

In the last chapter of this thesis, we focus on studying another poorly characterized metabolic system that is the process of iron reduction in *Clostridium acetobutylicum*. *C. acetobutylicum* is a well-known model organism for the production of several solvents and organic acids. Research has shown that this organism can reduce Fe(III), but the mechanism behind this reduction is yet to be identified. In this thesis, we analyzed the metabolism of *C. acetobutylicum* using one of its reconstructed genome-scale metabolic network models and experimental transcriptomics data in the presence or absence of Fe(III). By performing several computational studies, we suggested that NAD(P) is involved in the reduction of iron and is the potential physiological electron donor to Fe(III).

Keywords: metabolism, metabolic network models, optimization, global optimality, objective function, flux coupling analysis, *Plasmodium falciparum*, dormancy, *Clostridium acetobutylicum*, Fe (III) reduction.

Résumé

Les études informatiques du métabolisme visent à analyser systématiquement le comportement métabolique des systèmes biologiques dans différentes conditions. La reconstruction de modèles de réseaux métaboliques à l'échelle du génomique (GEM) à l'aide d'annotations des génomes, de données de la littérature et d'expériences enzymatiques en laboratoire humide est une étape importante vers l'étude systématique du métabolisme. Ces modèles détectent les connexions entre différents éléments du réseau en appliquant des bilans stoechiométriques en tenant compte des associations gène-protéine-réaction. Plusieurs méthodes ont déjà été développées pour l'analyse des modèles de réseaux métaboliques. Ces méthodes sont principalement utilisées pour optimiser une seule ou une combinaison de fonctions objectives biologique qui sont soumises à des contraintes stoechiométriques, thermodynamiques et autres. Dans cette thèse, nous avons résumé les méthodes d'optimisation et les fonctions objectives en les classant selon des caractéristiques biologiques et mathématiques. En particulier, nous suggérons des reformulations pour convertir certaines des classes d'optimisation complexes en classes plus simples. L'une des reformulations intéressantes est la conversion de la programmation fractionnaire linéaire en nombres entiers mixtes (MILFP) en programmation linéaire en nombres entiers mixtes (MILP). Nous montrons que la conversion est particulièrement utile pour étudier les relations de couplage dans les modèles de réseaux métaboliques qui sont soumis à des contraintes thermodynamiques. Le couplage est un concept important dans les réseaux métaboliques qui détermine comment les différents composants du réseau tels que les métabolites ou les réactions sont interdépendants. En particulier, l'analyse de couplage de flux (FCA) est une méthode qui a été largement utilisée pour évaluer les dépendances entre les réactions métaboliques. La FCA a été exploitée pour plusieurs applications telles que l'étude de l'essentialité des gènes, l'évolution du réseau, etc. En FCA, deux réactions sont considérées comme couplées si l'activité de l'une contraint l'activité de l'autre. Jusqu'à présent, la FCA a été utilisée pour analyser les réactions métaboliques dans des modèles à flux équilibré. Dans ce travail, nous avons développé une nouvelle formulation, l'analyse de couplage de flux thermodynamique (TFCA), qui calcule les couplages de flux de modèles de métabolisme

soumis à des contraintes thermodynamiques. Avec TFCA, nous montrons que l'ajout de contraintes thermodynamiques à un modèle à flux équilibré peut modifier considérablement la relation de couplage des réactions du réseau. De telles contraintes thermodynamiques peuvent affecter des directionnalités aux réactions et donc réduire le nombre de réactions bidirectionnelles (RBD). Une diminution des RBD entraîne par la suite une réduction significative de l'espace de solution. Nous montrons que le calcul des relations de couplage permet de réduire le nombre de combinaisons de RBD ce qui, à son tour, facilitera l'analyse du réseau métabolique.

En plus de proposer plusieurs reformulations mathématiques pour atteindre l'optimalité globale, nous avons également abordé la question de la recherche de la fonction objectif cellulaire appropriée dans différentes conditions de métabolisme cellulaire. Ce n'est pas toujours simple, car les activités métaboliques de certains organismes ne sont pas bien caractérisées. Le métabolisme de la phase de dormance chez certaines bactéries et parasites est un exemple de système biologique mal caractérisé. Dans cette thèse, nous avons étudié le comportement métabolique du parasite dormant du paludisme en utilisant un modèle à l'échelle du génome de *Plasmodium falciparum*. Nous avons examiné plusieurs fonctions objectives connues et nouvelles et les avons notées en fonction de la cohérence du modèle avec les données d'expression génique expérimentales. Nos résultats suggèrent que minimiser la dissipation d'énergie peut mieux décrire les activités métaboliques des parasites du paludisme en phase de dormance.

Dans le dernier chapitre de cette thèse, nous nous concentrons sur l'étude d'un autre système métabolique mal caractérisé qui est le processus de réduction du fer dans *Clostridium acetobutylicum*. *C. acetobutylicum* est un organisme modèle bien connu pour la production de plusieurs solvants et acides organiques. la littérature montre que cet organisme peut réduire Fe (III), mais le mécanisme derrière cette réduction reste à identifier. Dans cette thèse, nous avons analysé le métabolisme de *C. acetobutylicum* à l'aide d'un de ses modèles de réseaux métaboliques reconstruits à l'échelle du génome et de données transcriptomiques expérimentales en présence ainsi qu'en l'absence de Fe (III). En effectuant plusieurs études informatiques, nous avons suggéré que le NAD(P) est impliqué dans la réduction du fer et est le donneur d'électrons physiologique potentiel à Fe (III).

Mots-clés: métabolisme, modèles de réseaux métaboliques, optimisation, l'optimalité globale, fonction objectif, analyse de couplage de flux, *Plasmodium falciparum*, phase de dormance, *Clostridium acetobutylicum*, réduction du Fe(III).

Table of contents

1	Background	19
2	Classifying optimization methods and objective functions in studying cellular metabolism	25
2.1	Author contribution and publications used in this chapter	25
2.2	Global optimization problem	25
2.3	General overview of metabolic objective functions and optimization methods	26
2.4	Classification of optimization problems	28
2.5	How different optimization classes can be converted to each other	31
2.6	Review Process	33
2.6.1	Classification based on objective function and optimization problem	33
2.6.2	Classification based on study goal	34
2.6.3	Classification based on the need for reformulation	40
2.7	Bi-level optimization	40
2.8	Multi-objective optimization	41
2.9	Effect of adding thermodynamic constraints on problem types	43
2.10	Conclusion	44
3	Flux coupling analysis framework for the study of metabolic network models that include integer variables	46
3.1	Author contribution and publications used in this chapter	46
3.2	Introduction	46
3.3	Methods	48
3.3.1	Problem formulation for FCA	48
3.3.2	Problem formulation for MIFCA	50
3.3.3	Problem formulation for TFCA	52
3.3.4	Metabolic models and experimental data	53
3.3.5	Coupling of subnetworks	54
3.4	Results and Discussion	55
3.4.1	Classification of flux couplings	55
3.4.2	Integration of data in different levels in reduced E. coli model and identification of bidirectional reactions	59
3.4.3	Changes in couplings after constraining the models	60
3.4.4	Coupling and FDPs	63
3.4.5	Coupling of subnetworks	66
3.5	Conclusion	75

4	Cellular objectives of dormancy in malaria parasite	76
4.1	Author contribution and publications used in this chapter	76
4.2	Introduction.....	76
4.3	Materials and methods	77
4.3.1	<i>P. falciparum</i> genome-scale model used for the study	77
4.3.2	Thermodynamics-based flux analysis (TFA).....	77
4.3.3	Flux Variability Analysis (FVA).....	78
4.3.4	Objective functions	79
4.3.5	Integration of transcriptomics data.....	82
4.4	Results and discussion.....	83
4.4.1	Integration of objective functions and hallmarks of dormancy:	83
4.4.2	Computation of alternative flux distributions	84
4.4.3	Constitutive reactions	84
4.4.4	Production of biomass building blocks	85
4.4.5	Scoring combinations of objective functions and conditions.....	86
4.5	Conclusion	89
5	Study of iron reduction in <i>Clostridium acetobutylicum</i>	91
5.1	Author contribution and publications used in this chapter	91
5.2	Introduction.....	91
5.3	Materials and methods	93
5.3.1	Experimental data.....	93
5.3.2	TFA, FVA, sampling and calculation of minimal reaction subnetworks.....	93
5.3.3	Minimal Network Enrichment Analysis	95
5.4	Results and discussion.....	96
5.4.1	Experimental and theoretical growth rates.....	97
5.4.2	Experimental and theoretical by-product productions	97
5.4.3	Changes in FVA and mean flux values with and without iron reduction	98
5.4.4	Minimal reaction subnetwork for fermentation with and without iron reduction	101
5.4.5	Excess ATP production calculation	102
5.4.6	Enrichment analysis with MiNEA	104
5.5	Conclusion	109
6	Conclusions and perspectives	111
7	Appendix	115
8	References	134

List of figures

Figure 2-1. Possible conversion of optimization methods to each other. The blue dashed line is an approximate conversion and the original and the converted problem are not exactly the same.	31
Figure 3-1. Full coupling in metabolic networks. A) Full coupling in bidirectional reactions. In this case, simultaneous existence of positive direction of v_a and negative flux of v_d is not feasible. In addition, simultaneous existence of positive flux of v_d and negative flux of v_a are also not possible (it can be vice versa, i.e., full coupling of positive flux of v_a and negative flux of v_d and full coupling of negative flux of v_a and positive flux of v_d). B) Full coupling in unidirectional reactions.	55
Figure 3-2 Partial coupling in metabolic networks. Partial coupling can only happen for two unidirectional reactions	56
Figure 3-3. Directional coupling in metabolic networks. A) A bidirectional reaction that is directionally coupled to a unidirectional reaction, which means its flux in either direction is coupled to the flux of unidirectional reaction. B) v_a is directionally coupled to reaction v_d . .	57
Figure 3-4. Uncoupled reaction pairs in metabolic networks.	57
Figure 3-5. Conditional (directional) coupling in metabolic networks. A) only positive direction of v_a is coupled to v_d . B) positive direction of v_a is directionally coupled to positive direction of v_d and negative direction of v_d is directionally coupled to negative direction of v_a , negative direction of v_a is uncoupled to positive direction of v_d , and the combination of negative direction of v_d and positive direction of v_a is not feasible.	58
Figure 3-6. Flux directionality profile tree. This is the FDP tree in the model with constraints on metabolite concentrations and flux directionalities. Numbers on the nodes show the number of FDPs after fixing the directionality of some reactions to one specific direction...	66
Figure 3-7. D_1 for all pairs of subnetworks, when no flux is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction.	67
Figure 3-8. D_2 for all pairs of subnetworks, when no flux is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction.	67
Figure 3-9. D_3 for all subnetworks, when no flux is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction.....	68

Figure 3-10. D_1 for all pairs of subnetworks when no flux is maximized and we add the union of reactions of all possible minimal subnetworks	69
Figure 3-11. D_3 for all subnetworks when no flux is maximized and we add the union of reactions of all possible minimal subnetworks.....	70
Figure 3-12. D_1 for all pairs of subnetworks, when NAD production is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction.	73
Figure 3-13. D_1 for all pairs of subnetworks, when acetyl-CoA production is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction.....	73
Figure 3-14. Ratio of number of fully coupled reaction pairs when maximizing for production of each of the building blocks to the number of fully coupled reaction pairs when no flux is maximized	74
Figure 3-15. Ratio of number of directionally coupled reaction pairs when maximizing for production of each of the building blocks to the number of directionally coupled reaction pairs when no flux is maximized	74
Figure 5-1. Experimental and theoretical growth rates, A) with Fe(III) and, B) without Fe(III)	97
Figure 5-2. Flux ranges of experimental by-products production compared with model-based predictions of fermentation, A) with Fe(III) and, B) without Fe(III). The minimal and maximal flux of acetate, butyrate, butanol and H_2 were compared of experimental data (blue) and model predicted data (red) by using NAD(P)H as physiological electron donor.....	98
Figure 5-3. The ratio of mean value of fluxes with and without iron reduction using NAD(P)H. The mean of flux values (in mmol/gDW.h) from model-based samples was used to calculate the ratio of fermentation with iron reduction divided by fermentation without iron reduction for the reactions of glycolysis, fermentation, citric acid cycle and the pentose phosphate pathway. Grey lines show reactions where no flux was present in the model and gaps in the values within the figure legend indicate that these range was not represented. Colours representing values greater than 1 show reactions with higher flux in fermentation with iron reduction and those less than 1 show reactions with higher flux in fermentation without iron reduction.....	100
Figure 5-4. The ratio of mean value of fluxes with and without iron reduction using reduced ferredoxin. The mean of flux values (in mmol/gDW.h) from model-based samples was used to calculate the ratio of fermentation with iron reduction divided by fermentation without	

iron reduction for the reactions of glycolysis, fermentation, citric acid cycle and the pentose phosphate pathway. Grey lines show reactions where no flux was present in the model and gaps in the values within the figure legend indicate that these ranges were not represented. Colours representing values greater than 1 show reactions with higher flux in fermentation with iron reduction and those less than 1 show reactions with higher flux in fermentation without iron reduction.....101

Figure 5-5. Minimal reaction subnetwork for fermentation with and without iron reduction with NAD(P)H as electron donor (only glycolysis and solventogenesis reactions are shown). Black arrows show reactions present in all minimal networks (necessary); green arrows show reactions present in some minimal networks (substitutive) with iron provided and present in all minimal networks of fermentation without iron; and red arrows show reactions present in some minimal networks of fermentation with iron, but not present in any minimal networks of fermentation without iron reduction. Grey arrows are reactions not present in any minimal subnetwork.....103

Figure 5-6. Minimal reaction subnetwork for fermentation with and without iron reduction with ferredoxin as electron donor (only glycolysis and solventogenesis reactions are shown). Black arrows show reactions present in all minimal networks (necessary); green arrows show reactions present in some minimal networks (substitutive) with iron provided and present in all minimal networks of fermentation without iron; and red arrows show reactions present in some minimal networks of fermentation with iron, but not present in any minimal networks of fermentation without iron reduction. Grey arrows are reactions not present in any minimal subnetwork.....104

Figure 5-7. Number of alternative minimal subnetworks in different conditions.107

Figure 5-8. Sizes of alternative minimal subnetworks in different conditions.....107

Figure 5-9. Percentage of common reactions between alternative minimal subnetworks in different conditions.107

Figure 5-10. Percentage of deregulated common reactions in different conditions.....108

List of tables

Table 2-1. classification of objective functions used in the study of metabolic networks.....	36
Table 3-1. Integrating information in different levels to reduced model of E. coli.....	60
Table 3-2. changes in coupling relationships when we add thermodynamic constraints to a non-thermo model.....	61
Table 3-3. changes in coupling relationships when we constrain the metabolite concentrations in the thermodynamically constrained model	62
Table 3-4. changes in coupling relationships when we constrain reaction directionalities in the thermodynamically constrained model with constrained concentrations.....	63
Table 3-5. Impact of having coupled reaction pairs in reducing the number of FDPs in the model	64
Table 3-6. the impact of fixing bidirectional reactions to forward or reverse direction, one by one in the most constrained model.....	65
Table 3-7. Full name of biomass building blocks in E. coli model.....	71
Table 4-1. number of optimal alternative solutions per combinations of condition and objective function	84
Table 4-2. Number of constitutive reactions per combination of condition and objective function	85
Table 4-3. Number of Biomass Building Blocks that can be synthesized in at least on alternative solution per combination of condition and objective function	86
Table 4-4. Consistency score per combination of objective functions and conditions with integrated transcriptomics data	87
Table 4-5. Reactions that cannot be down-regulated in the forced growth condition and Reactions that can only be up-regulated with dissipation minimization as objective function	88
Table 5-1. List of enriched subnetworks by MiNEA. MiNEA was done by maximizing the formation of certain metabolites using three potential electron donors (NADH, NADPH and ferredoxin) for Fe(III) reduction. Underlined are the differences between the potential electron donors within either solid iron reduction or soluble iron reduction. The model was constrained with glucose and Fe(III) fluxes	105

Table 5-2. Up and down regulated genes and corresponding reactions for the biosynthesis of metabolites that are enriched with deregulated genes when NADH is used as electron donor. Essential genes are those genes that are present in all alternative minimal subnetworks and substitutive genes are those genes that are present only in a subset of alternative minimal subnetworks. e and c after the names of some metabolites show that they are extracellular or cytoplasmic, respectively.....108

List of abbreviations

ABE	Acetone-Butanol-Ethanol
BBB	Biomass Building Blocks
BDR	Bidirectional Reaction
CS	Consistency Score
ETC	Electron Transfer Chain
FBA	Flux Balance Analysis
FCA	Flux Coupling Analysis
FCF	Flux Coupling Finder
FDP	Flux Directionality Profiles
FVA	Flux Variability Analysis
GEM	Genome-scale model
GPR	Gene-Protein-Reaction
LFP	Linear Fractional Programming
LP	Linear Programming
MIFCA	Mixed Integer Flux Coupling Analysis
MILFP	Mixed Integer Linear Fractional Programming
MILP	Mixed Integer Linear Programming
MiNEA	Minimal Network Enrichment Analysis
MINLP	Mixed Integer Non Linear Programming
MIQFP	Mixed Integer Quadratic Fractional Programming
MIQP	Mixed Integer Quadratic Programming
MOMA	Minimization of Metabolic Adjustment
QFP	Quadratic Fractional Programming
QP	Quadratic Programming
TFA	Thermodynamics-based flux analysis
TFBA	Thermodynamics-based Flux Balance Analysis
TFCA	Thermodynamics-based Flux Coupling Analysis
TMFA	Thermodynamics-based Metabolic Flux Analysis

1 Background

The focus of biological researches started from study of biological components through genetic, biochemical or molecular biology approaches. These approaches have provided in depth understanding of individual components. However, they do not capture the systematic connections of biological elements and environmental factors that lie behind the ultimate physiology of the cells. Systems biology's aim is to convert detailed information about an organism into a computational platform, such that cellular phenotype can be analyzed and predicted from cellular genotype (1). It can help to realize how all the molecules in a cell can cooperate to create rational physiological functions. The focus of systems biology has shifted to many different areas including stochastic kinetic models and statistical Bayesian networks (2). However, systems-level study of metabolism kept its significance and proved to be a promising and useful research area. This systems-level study requires the detailed information about all the metabolites and reactions occurring in the cells and genes encoding enzymes that catalyze those reactions. Gene to protein to reaction (GPR) associations can connect genome to biochemistry. Technological advances have now facilitated high-throughput characterization of different biological components all at once. One of the major progresses in this area is the development of advanced strategies for genome sequencing of different organisms. This advancement gives a systems-level or genome-scale point of view of metabolism. The combination of full genome sequences with the massive progresses in genetics, molecular biology, and biochemistry has enabled the genome-scale reconstruction of metabolic networks (1,3). The number of reconstructed genome scale metabolic models has increased tremendously in recent years. These metabolic models can be utilized as an illustrative of all bio-chemical reactions that can happen in a cell and contain all the information about metabolic reactions and metabolites that are present in a cell (4). Metabolic networks are very complex due to both large number of metabolic reactions and connections of metabolites that take part in these reactions. A metabolite can be involved in different parts of the metabolism and being used by several enzymes as substrates or produce

as products. Genome-scale metabolic networks have facilitated the study of this complexity. These models can be converted to a mathematical format. More specifically, participation of each metabolite in each reaction can be illustrated in a stoichiometric matrix. The rows and columns of this matrix represent the individual metabolites and reactions respectively. The entries of this matrix are the stoichiometric coefficients of each metabolite in different reactions. Negative and positive coefficients are used for substrates and products respectively. The relation between metabolites and reactions represented by the stoichiometric matrix is a set of linear equations. The important assumption behind the study of metabolic networks is the quasi-steady state assumption which presumes that there is no accumulation or dissipation of metabolites in the cell. These models usually contain a biomass reaction which represents the composition of 1 gDW of cell producing from different building blocks. The metabolic models also include information on possible uptaken and secreted metabolites through boundary reactions. The basic mathematical representation of metabolic networks is the following formulation:

$$\begin{aligned}
 & \min/\max Z = c^T v \\
 & \text{subject to:} \\
 & S \cdot v = 0 \\
 & v_{i,\min} \leq v_i \leq v_{i,\max}
 \end{aligned}$$

Where S is the stoichiometric matrix and v is the vector including the fluxes through all individual reactions in the network. The equality constraint represents the quasi-steady state assumption ensuring there is no net accumulation or consumption of metabolites. v_{\min} and v_{\max} indicates the possible lower and upper bounds of each metabolic reaction. Z show the metabolic objective of the cell which is represented through linear combinations of different fluxes in the network.

Using the above optimization approach, which is called flux balance analysis (FBA) helped researchers to study not only the overall behavior of metabolic networks, but also the contribution of individual metabolic fluxes in this metabolic behavior. In the linear space, the objective function has only one global optimum value. However, since these models are usually underdetermined, the solution to the above optimization problem is not unique. In other words, there are different flux distributions that can satisfy the same global objective

value. One of the reasons that may cause having alternative flux distributions is the bi-directionality of reactions. The experimental or literature data about the reversibility is only available for a subset of enzymes and associated reactions. One of the ways to assign directionalities to the reactions is by applying thermodynamic laws on the reactions of the network. In thermodynamic flux analysis (TFA), a mixed-integer linear programming (MILP) is solved to introduce thermodynamic constraints to the original FBA problem (5). These constraints take into account metabolite concentrations and Gibbs free energy of reactions to determine directionality of all individual reactions of the network. By assigning directionalities to the reactions, the solution space can be reduced significantly, due to the reduction in the number of the bidirectional reactions (BDRs).

Maximizing the flux through the biomass reaction was originally used as the objective function in the above optimization problem. However, it has been shown that living cells have other important goals such as energetic efficiency or optimal resource allocation (6). In addition, genetic and environmental perturbations may force the organisms to use their metabolic capabilities to achieve other biological objectives rather than maximum growth (7). Apart from the study of different objective functions, metabolic network models can be treated as platforms to integrate wide range of experimental data from gene expressions, protein abundances to reaction fluxes and metabolite concentrations. These models have also been extensively used for different applications such as knowledge gap filling of metabolism, drug targeting and overproduction of valuable compounds. Researchers have developed many different optimization methods for the analysis of metabolic networks, where a single or a combination of objective functions are optimized subject to stoichiometric, thermodynamic and other known constraints. There are different classes of optimization problems based on the type of variables, constraints and objective functions. Not all these classes can be solved to global optimality having current solvers. In addition, some of the classes can be converted to other classes using transformation or linearization strategies. In the second chapter of this thesis, we introduce different classes of optimization problems that are used in study of metabolic network models. Then, we review papers about metabolic network model analysis and classify the optimization methods that are used in them based on different properties including problem type, objective function type and the goal of the study. Further, we explain which of these problems can be solved to global optimality and which of them need reformulation. Finally, we discuss how introducing

thermodynamic constraints can change problem type and arise the need for further reformulations. We will show an interesting example of the need for reformulation after the introduction of thermodynamic constraints in chapter three.

Different components of the metabolic network models, such as metabolite concentrations or reaction fluxes are not completely independent of each other and connectivity in the network do has an effect on how these components are coupled. Flux Coupling Analysis (FCA) is a method, which has been used extensively for evaluating the dependencies between metabolic reactions (8). FCA has been exploited for several applications such as studying gene essentiality, network evolution and gap filling. In FCA, two reactions are considered as coupled if the activity of any of them constrains the activity of the other. So far, FCA has been used for analysing metabolic reactions in flux-balanced models. In FCA a fractional programming problem is solved to identify the ratio of any two reactions in the network. having these ratios different types of couplings with different strength are defined between pairs of reactions. The original FCA is a non-linear problem, but can be converted to a linear problem by applying appropriate strategies. FCA is a useful method for researchers that are going to extract biologically meaningful information from metabolic network reconstructions. Investigation of the coupling properties of metabolic networks is important to both understand the organizational principles of metabolic interactions within metabolic networks, and for more efficiently suggesting engineering modifications. Coupling information can be used to suggest ways for inactivating a particular reaction, or making some reactions of interest coupled to each other. In chapter three, we develop a new formulation, Thermodynamic Flux Coupling Analysis (TFCA), which uses mixed integer linear fractional (MILFP) programming to calculate flux couplings for thermodynamically constrained models of metabolism. With TFCA we show that addition of thermodynamic constraints to a flux-balanced model can significantly change the coupling relationship of reactions of the network. We show that calculating coupling relations helps in reducing the number of combinations of BDRs, which in turn will facilitates the analysis of the metabolic network. The MILP formulation of the TFCA allows for analysis of a broader class of problems, with and without thermodynamic constraints that involve integer variables.

In chapter four of the thesis we show another application of our study of objective functions in chapter two. We have studied the cellular objectives of the dormancy phase in malaria parasite. Malaria is an infectious disease caused by *Plasmodium* genus. Some *plasmodium*

species can enter the dormancy phase and leave this phase at a later time in their life cycle. This state is still poorly characterized and the mechanisms of entering or leaving this phase is not known. In this chapter by utilizing an already existed genome-scale metabolic network model of *Plasmodium falciparum* (9), we study different objective functions that can describe the dormancy phase. We then score and rank these objective functions based on the agreement of the model with available experimental gene expression data of the dormant parasite. We show that among all the objective functions that are investigated, minimizing energy dissipation can best describe the physiology of dormant parasite.

In chapter five of this thesis, we are going to analyze the metabolism of *Clostridium acetobutylicum* by utilizing an existing genome-scale model of this organism (10). *C. acetobutylicum* is a gram-positive anaerobic bacterium that is known for its ability to produce different organic acids and solvents. In addition, experimental studies confirm that *C. acetobutylicum* is able to reduce different forms of Fe(III) (solid and soluble) to Fe(II). Iron is an important element for cell survival and is involved in many vital biological processes including energy generation. However, the mechanisms behind and enzymes and cofactors involved in this process are not characterized yet. To investigate the mechanism of iron reduction, we hypothesized three different electron donors for reducing Fe(III) and constrain the metabolic network with experimental metabolite concentration and transcriptomics data. We then analyzed parts of the metabolism that have changes in their activity when Fe(III) is provided for the model. By means of this analysis, we could propose potential mechanism for iron reduction and identify enzymes, reactions and pathways involved in this process.

In this thesis, we targeted investigation of metabolic network models by several approaches. We thoroughly studied the optimization methods and objective functions used in the study of metabolic network models. We analyzed the effects of introducing thermodynamic constraints on optimization methods. We used flux coupling analysis as a tool to study the relation between different reactions of a thermodynamically constrained metabolic network model. To this goal we developed a MILFP problem and converted it to MILP formulation using known strategies. We have also investigated different known objective functions from chapter one and also novel functions, to study the metabolic goal of the latency phase of malaria parasite. At the end, we analyzed the metabolism of *Clostridium acetobutylicum* in

iron reduction condition, using known optimization methods for physiological study and data integration.

Apart from this thesis, I also contributed to the development and refinement of a pipeline that evaluates pathways generated for production of target chemicals. This pipeline integrates the pathways to the genome-scale model of interest and check the thermodynamic feasibility and does the evaluation of yield of embedded pathways. In addition, I contributed in the development of a method for analysis and integration of C^{13} labelling data into metabolic models to study the impact of different physiological states on the distribution of atom labels in the central carbon metabolism. Moreover, I contributed in the development of Optknock framework for thermodynamically constrained models in order to have thermodynamically feasible genetic intervention strategies.

2 Classifying optimization methods and objective functions in studying cellular metabolism

2.1 Author contribution and publications used in this chapter

The work of this chapter is based on the following paper which is under preparation for submission:

Z. Hosseini, MO. Oftadeh, V. Hatzimanikatis (2020) “Global optimization problem to describe different objective functions used in studying cellular metabolism” (in preparation).

This work was done in collaboration with Mohammad Omid Oftadeh, a PhD student in LCSB.

2.2 Global optimization problem

Global optimization is the task of finding the global maximum or minimum of a function or set of functions. The goal of global optimization is to determine not just a local minimum (maximum) but the smallest (largest) local minimum (maximum). In contrast to local optimization in which the fulfillment of the local minimum can be guaranteed (when the gradient is equal to zero), no such general condition exists in global optimization to confirm that the global minimum has been reached (11). In linear programming (LP) and other types of optimization problems, the input variables are subject to some equality or inequality constraints. Linear programming and mixed-integer linear programming (MILP) are problems that can be solved to global optimality using current solvers such as CPLEX. Quadratic programming (QP) and mixed integer quadratic programming (MIQP) are not linear but they can also be solved to global optimality. There are other classes of problems such as linear

fractional programming (LFP) or mixed integer linear fractional programming (MILFP) that are not linear. These classes can be convex or non-convex. However, there are some reformulations that can be applied on these classes that will convert them to LP or MILP problems and therefore their global optimum can be calculated using existing approaches. It should be noted that global optimality is with respect to the objective function. However, the variables are not fixed to a certain value and multiple solutions may exist for a single global objective function.

2.3 General overview of metabolic objective functions and optimization methods

The goal of systems biology is quantitative understanding of functional interactions between the multiple cellular components to eventually predict network, cell and organism behavior. This understanding needs computational models to capture the large numbers of molecular components that can interact within interlinked biochemical networks. Genome-scale metabolic network model reconstruction has been recognized as one of the main modeling approaches for systems-level metabolic studies (12,13). A genome-scale model computationally describes a whole set of stoichiometry-based, mass-balanced metabolic reactions in an organism using gene-protein-reaction (GPR) associations that are formulated on the basis of genome annotation data and experimentally obtained information. Metabolic networks are well established since we know most reactions, their catalyzing enzymes and encoding genes, and how they interact stoichiometrically within a biochemical network. These models allow to predict network capabilities, for example, by predicting metabolic fluxes using flux balance analysis (FBA). To identify optimal solutions in the big solution space, FBA objective functions are defined to solve the system of equations that represent the mass balance, thermodynamic and other constraints (6). While different objectives were proposed for different biological systems, by far the most common assumption is that microbial cells maximize their biomass yield. However, it is shown that maximization of biomass yield might not be necessarily the evolutionary favored way of organisms to live (14). Therefore, alternative objective functions have been proposed for physiological study of metabolic network models such as optimizing enzyme efficiency by minimizing total flux through the

network or optimizing energetic efficiency through maximizing ATP production (6). Combination of different objective functions have been used when a single objective is not enough to describe a metabolic system. Since the objective function of a cell cannot be known a priori, bi-level based optimization methods have been developed to suggest an objective function for a model by minimizing the distance of flux prediction and experimental flux measurements (15,16). In addition, several methods have been developed to integrate and analyze biological data, including different large scale omics data or small scale measurements of metabolite concentrations or fluxes (5,17–19). Furthermore, methods have been proposed to use the metabolic network models for different applications. These applications include drug target identifications in pathogens, genetic modifications for the overproduction of compounds that can already be produced by the organism, strain development for production of new bio-based compounds, and prediction of new metabolic capabilities through gap filling to propose enzymatic functions (20–22).

In the first part of this chapter, we are going to present a classification for optimization methods that have been used in the analysis of metabolic network models. The classifications are based on the linear, quadratic or fractional terms in the equations and also the type of variables (integer or continuous or both) that are used to formulate the problem. We then review papers in metabolic network models analysis and classify them based on the optimization methods and objective functions that they used in their study of metabolism. We show that some of these problems can be convertible to each other. We propose proper reformulation so that we can ensure global optimality for every possible optimization problems and objective functions in these categories. We discuss that by having this global reformulations one can add more complex constraints and variables to an already existed problem and it can still be solved to global optimality. We also propose another classification for existing methods in the analysis of metabolic networks based on their goal of study. One can easily see that many of the methods can be categorized in more than one class. We believe that this classification of papers can serve as a repository of existing methods for metabolic network model study and analysis and also for easily expanding the known approaches to do more complex and thorough analysis of metabolism.

2.4 Classification of optimization problems

We are going to classify the optimization problems in ten groups. A general mathematical formulation for each group is represented:

Linear programming (LP)

$$\begin{aligned} & \min_x c^T x \\ & \text{subject to:} \\ & Ax \leq b \\ & x \geq 0 \end{aligned}$$

LPs are a group of optimization problems with only continuous variables and both objective function and constraints are linear.

Mixed integer linear programming (MILP)

$$\begin{aligned} & \min_{x,y} c_1^T x + c_2^T y \\ & \text{subject to:} \\ & A_1 x + A_2 y \leq b \\ & x \geq 0 \\ & y \in \mathbb{Z}^n \end{aligned}$$

MILPs are optimization problems with both continuous and integer (or in a simpler case binary) variables. The constraints and objective function are still linear combination of variables but can have terms with integer variables.

Linear fractional programming (LFP)

$$\begin{aligned} & \min_x \frac{c^T x + \alpha}{d^T x + \beta} \\ & \text{subject to:} \\ & Ax \leq b \\ & x \geq 0 \end{aligned}$$

LFP is a class of optimization in which the objective function is the division of two linear combinations of variables. The constraints are linear and the variables are continuous.

Mixed integer linear fractional programming (MILFP)

$$\min_{x,y} \frac{c_1^T x + c_2^T y + \alpha}{d_1^T x + d_2^T y + \beta}$$

subject to:

$$Ax + By \leq b$$

$$x \geq 0$$

$$y \in \mathbb{Z}^n$$

MILFPs have both integer and continuous variables. The constraints are linear combinations of both types of variables. However, the objective function is a fractional equation that have both continuous and integer variables.

Quadratic programming (QP)

$$\min_x \frac{1}{2} x^T Q x + c^T x$$

subject to:

$$Ax \leq b$$

$$x \geq 0$$

QPs are like LPs with additional terms in the objective function. These additional terms are the square of continuous variables with some coefficients.

Mixed integer quadratic programming (MIQP)

$$\min_{x,y} \frac{1}{2} x^T Q_1 x + \frac{1}{2} y^T Q_2 y + c_1^T x + c_2^T y$$

subject to:

$$Ax + By \leq b$$

$$x \geq 0$$

$$y \in \mathbb{Z}^n$$

The difference between MIQPs and QPs is that they can have integer variables in the objective function and constraints.

Quadratic fractional programming (QFP)

$$\min_{x,y} \frac{\frac{1}{2} x^T Q_1 x + c_1^T x + \alpha}{\frac{1}{2} x^T P_1 x + d_1^T x + \beta}$$

subject to:

$$Ax + By \leq b$$

$$x \geq 0$$

$$y \in \mathbb{Z}^n$$

QFPs are like LFPs with quadratic terms in addition to linear terms in the objective function.

Mixed integer quadratic fractional programming(MIQFP)

$$\min_{x,y} \frac{\frac{1}{2}x^T Q_1 x + \frac{1}{2}y^T Q_2 y + c_1^T x + c_2^T y + \alpha}{\frac{1}{2}x^T P_1 x + \frac{1}{2}y^T P_2 y + d_1^T x + d_2^T y + \beta}$$

subject to:

$$Ax + By \leq b$$

$$x \geq 0$$

$$y \in \mathbb{Z}^n$$

MIQFPs are a rare problem with fraction of quadratic and linear terms of continuous and integer variables in the objective function. The constraints are linear combinations of both types of variables.

Bi-level optimization

$$\min_x c_1^T x$$

subject to:

$$A_1 x \leq b_1$$

$$\min_x c_2^T x$$

subject to:

$$A_2 x \leq b_2$$

$$x \geq 0$$

Bi-level problems are nested optimization problems in which, in addition of typical constraints of an optimization problem, the outer problem is subject to another optimization problem with its own constraints. One should note that both inner and outer level problems can consist of integer variables or quadratic terms.

Multi-objective optimization

$$\min_{x,y} (f_1(x, y), f_2(x, y), \dots, f_n(x, y))$$

subject to:

$$Ax + By \leq b$$

$$x \geq 0$$

$$y \in \mathbb{Z}^n$$

In multi-objective optimization, a combination of objective functions is optimized subject to linear constraints of continuous or integer variables. Objective functions can be from any of the above classes.

2.5 How different optimization classes can be converted to each other

Not all the classes that are described above can be solved to global optimality having current solvers. we can easily solve LPs, MILPs, QPs and MIQPs. Some of the classes can be converted to each other using proper transformation and introduction of new variables. Figure 2-1 shows the optimization classes introduced in section 2-4 and the ones that can be converted to each other. The blue dashed line is an approximate conversion and the original and the converted problem are not exactly the same.

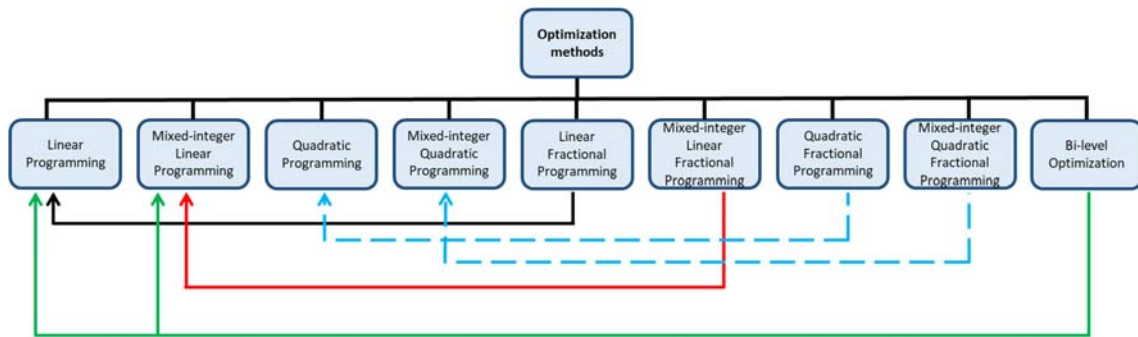


Figure 2-1. Possible conversion of optimization methods to each other. The blue dashed line is an approximate conversion and the original and the converted problem are not exactly the same.

Fractional programming for instance, are nonlinear problems that can be reformulated to linear programming using Charnes-Cooper transformation (23). A linear fractional problem after the transformation will be in the below form.

$$\min_{\hat{x}, t} c^T \hat{x} + \alpha t$$

subject to:

$$A\hat{x} \leq bt$$

$$d^T \hat{x} + \beta t = 1$$

$$\hat{x} \geq 0, t \geq 0$$

This new formulation is the result of multiplication of numerator and denominator of the objective function and both sides of all constraints to a new positive continuous variable, t . Then, xt is replaced by \hat{x} , a constraint is added to force the new denominator to be equal to 1 and the new numerator is minimized. Therefore, in the above reformulation, $x = \frac{\hat{x}}{t}$ and $t = \frac{1}{d^T x + \beta}$. More details about this transformation can be found in literature (8,23).

In addition, MILFP can be converted to MILP using both Charnes-Cooper transformation and Glover's linearization scheme (24). Applying Charnes-Cooper transformation on a MILFP will result in the following problem

$$\min_{\hat{x}, y, t} c_1^T \hat{x} + c_2^T y t + \alpha t$$

subject to:

$$A\hat{x} + B y t \leq b t$$

$$d_1^T \hat{x} + d_2^T y t + \beta t = 1$$

$$\hat{x} \geq 0, t \geq 0$$

$$y \in \mathbb{Z}^n$$

Here, $y t$ is a bilinear term that can be replaced with a continuous variable. It is done by adding a couple of linear constraints based on Glover's linearization strategy. For every bilinear term, three more constraints should be added to the problem:

$$\hat{y} \leq t$$

$$\hat{y} \leq M y$$

$$\hat{y} \geq t - M(1 - y)$$

Here, $\hat{y} = y t$ and the above constraints insure that $\hat{y} = 0$ whenever $y = 0$ and $\hat{y} = t$ whenever $y = 1$. Here, M is a sufficiently large number. After these variable conversions and constraints additions the final problem will look like below:

$$\min_{\hat{x}, t} c_1^T \hat{x} + c_2^T \hat{y} + \alpha t$$

subject to:

$$A\hat{x} + B\hat{y} \leq b t$$

$$d_1^T \hat{x} + d_2^T \hat{y} + \beta t = 1$$

$$\hat{y} \leq t$$

$$\begin{aligned}
\hat{y} &\leq My \\
\hat{y} &\geq t - M(1 - y) \\
\hat{x} &\geq 0, t \geq 0, \hat{y} \geq 0 \\
y &\in \mathbb{Z}^n
\end{aligned}$$

Quadratic fractional programming is a rare class of problems. There is no conversion to global optimality for these kind of problems. Authors in (25,26), for instance, used approximations to convert it to quadratic programming.

Bi-level problems and multi-objective problems have been used a lot in the study of metabolic networks and they also need specific strategies and reformulations for solving. These strategies are discussed in section 2-7 and 2-8.

2.6 Review Process

Here, we review papers that are considering different objective functions for a metabolic network model. We reviewed around 1100 papers that reconstructed, analyzed or integrated data into metabolic network models. At the end we chose the papers that are either published recently (from 2018 to June 2020), or they have more than 100 citations based on Scopus database for more detailed classification and analysis. Among these papers, we extracted 103 unique single or combination of objective functions for analysis of metabolic network models which are summarized in table 7-1 of Appendix. The rest of the papers use already existed methods or objective functions for their studies. We then categorize the papers based on different features.

2.6.1 Classification based on objective function and optimization problem

We classify the optimization types of the methods based on the classes introduced in section 2-4. Specifically, we categorize the papers based on their objective type, i.e., either it consists of linear or quadratic or fractional terms with linear or mixed integer and linear variables, or it has multi-objective function. We also categorized the studies based on the problem types. The problem type can also be a LP, MILP, QP, MIQP, LFP, MILFP or bi-level problem. Table 2-1 shows some example of each category. The complete table can be found in the Appendix.

LPs are used for optimizing biomass (5,27–29), production of a desired metabolite (30,31) and ATP production (32). Having minimum number of changes after a perturbation (33), minimum number of metabolites that are required in a media so that the cell can grow (34), or maximum number of reactions in the network that their activities are consistent with gene expression data (35), have been solved using MILPs.

Examples of methods based on fractional programming in studying metabolism are the study of flux couplings or maximizing ATP production per enzyme reaction (8,36). Flux coupling analysis in thermodynamically constrained metabolic network models is an example of mixed integer linear fractional programming which we will discuss in detail in the next chapter. QPs and MIQPs are used to study minimizing the difference between measured and predicted fluxes (37) or minimizing the sum of squared fluxes in order to have enzyme efficiency (38). People used QFP to optimize maximization of ATP production divided by sum of squared fluxes to optimize for both enzymatic and energetic efficiency (25,26).

2.6.2 Classification based on study goal

We further group the papers based on the purpose of the study. More precisely, we classify the papers in three groups:

2.6.2.1 Papers that are studying physiology of cells using metabolic network modelling

These studies include prediction of maximal biomass production or yield for single species (27–29,39–41) or communities (42–46) for growth efficiency. Maximal ATP production (32,47,48), or maximal of ATP production per enzyme reaction (36) are used as objective function when energetic efficiency of cells are of biological importance. Finding minimal growing reaction set (49,50), minimizing sum (51,52) or squared sum of fluxes (38,46,53) or minimizing number of active reactions (54) are used to optimize for enzymatic efficiency. Finding minimal media (34) or minimal set of exchange metabolites between two organisms so that they can grow simultaneously (55), minimum amount of substrate consumption (45,46,56–59) or minimization of non-essential nutrients uptake rate (60) are examples of objective functions used for optimum resource consumption. Maximizing NAD (NADH, NADP, NADPH) production or consumption from specific pathways or in the specific compartments (61) or the whole model (62) are used for optimizing redox potential of cells.

Minimization of the difference between the flux vectors of wild-type and mutated models is also a very well-known objective function that is used in a method called MOMA (Minimization Of Metabolic Adjustment) for studying the physiology of genetically perturbed cells (7). The distance between acidogenic and solventogenic phases is another target of minimization that have been used by Lee et al (63). Succurro et al., used the distance between consecutive time steps in dynamic flux balance analysis as the objective to be minimized (64). ROOM suggests minimization of number of significant flux changes instead of minimizing the distance between two flux vectors (33). Flux coupling analysis is another physiological study of cells in which authors calculated reaction dependencies in a model in different conditions (8). The objective functions in this paper is to find the minimum and maximum of reaction pairs of study (8). Some studies investigated the maximum production rate of some metabolites in a specific condition or physiology having usually considering a minimum amount for biomass flux (30,31,65,66). Some of these metabolites that their production are mostly targeted for optimization are succinate, lactate, acetate or lipids (67–69). Another attempt in studying physiology is finding essential and synthetic lethal genes without the need for exhaustive search. Authors in (70) used a bi-level programming approach that maximize and minimize biomass in inner and outer levels respectively, subject to desired number of gene deletions to ensure no biomass production after gene deletion. There are also several studies that considered combination of the above-mentioned functions as objective function of the model in a multi-objective type problem (71–74). For instance, combination of biomass maximization and sum of fluxes minimization is a widely used objective function for simultaneously optimize for enzyme efficiency and growth (75,76). We will discuss more about multi-objective problems in section 2-8.

Table 2-1. classification of objective functions used in the study of metabolic networks

	Objective Function	Type of Objective							Type of Problem								Goal of Study			Need for reformulation		References
		LP	MILP	LFP	QP	MIQP	QFP	MO ¹	LP	MILP	LFP	QP	MIQP	QFP	MO	BL ²	Phys ³	App ⁴	DI ⁵	NRN ⁶	RN ⁷	
1	Max. biomass (growth rate)	*							*	*							*			*		(5,39,77)
2	Max. ATP production	*							*								*			*		(32)
3	Min. the squared sum of fluxes				*							*					*			*		(38)
4	Max. ATP per flux unit (sum of fluxes)			*							*						*				*	(36)
5	Max. number of reactions whose activity is consistent with their expression state		*							*									*	*		(35)
6	Min. number of reactions that can produce a		*							*									*	*		(19)

¹ Multi-objective

² Bi-level

³ Physiology

⁴ Application

⁵ Data Integration

⁶ No Reformulation is Needed

⁷ Reformulation is Needed

	specific metabolites																					
7	Min. Growing Reaction Set		*						*							*			*			(50)
8	Outer: Max. bioengineering objective Inner: Max. biomass	*							*						*		*		*			(20)
9	Min. metabolic adjustment	*			*				*		*					*		*	*			(7,78)
10	Min. number of significant flux changes after perturbation		*						*							*		*	*			(33)
11	Min. the difference between fluxes and Min. the total sum of square of fluxes				*						*			*		*		*	*			(79)
12	Outer: Max. bioengineering objective Inner: weighted Max. of biomass and Min. of sum of fluxes	*								*					*	*		*		*		(80)

2.6.2.2 Papers that are integrating experimental data such as transcriptomics, metabolomics or fluxomics data, into metabolic network models

The general idea of most of the papers in this category is using LP or MILP formulations to maximize the consistency (or minimize the inconsistency) between reaction activities and their corresponding expression states (17,35,81–83). There are also studies that have just used the transcriptomics data to put more constraints on lower and upper bounds of fluxes and then used the constrained model for further analysis (84). MiNEA is a method that uses gene expression changes between two conditions to find deregulated metabolic tasks (19). In a recent study Sarkar et al., used single polymorphism data to minimize the deviation from a reference flux state plus the deviation from mass action kinetics (85). Minimizing the difference between flux predicted from metabolic modelling and flux calculated from experimental proteomic data is the objective of a study by Yizhak et al (37). Minimizing the error from experimental measurement of proteomics or transcriptomics data is considered as part of the objective function in a study by Tian et al (86). Minimizing the error between experimentally constrained and unconstrained model is another way of using experimental gene expression data (87). MOMA and ROOM can also be considered in this category if the reference flux that is used for distance minimization is the result of a fluxomics study (7,33). Metabolite concentrations is another source of data that have been used in studies of genome-scale models. In Thermodynamics-based flux analysis (TFA), metabolite concentrations have been used for constraining genome-scale models using thermodynamic constraints (5). It has also been used for finding the consistency between model metabolite production/consumptions with experimental measurements (88). Mo et al., used external metabolomics measurement to connect them to internal fluxes (89). Experimental gene knock-out analysis have been used to minimize the number of false positives and false negatives between model predictions and experimental measurements (90). GrowMatch for instance, uses a bi-level algorithm with maximum biomass and minimum biomass as the objective functions of outer and inner problems, respectively. It Then resolves a false positive case by finding the minimum number of deletions that are needed so that the model do not produce biomass (21). For false negative cases the problem is resolved by adding minimum number of reactions to the model so that the model produces biomass (21).

2.6.2.3 Papers that are using the computational studies of metabolism for different kind of applications

The mostly used application is optimizing the production of a desired metabolite and it is usually done via a bi-level optimization problem. Optknock uses a bi-level problem with maximizing production of the target metabolite in the outer problem and maximization of biomass production in the inner problem (20,91). There are also multi-objective studies with simultaneously optimizing for production of metabolites and another biologically meaningful objective function (92). Choi et al., maximized biomass flux while forcing the production of target flux and found the increased fluxes as targets for amplification (93). OptForce identifies minimum number of flux knock-outs, upregulations or downregulations in order to ensure a minimum target overproduction through a bi-level problem (94). Moreover, Pharkya et al., used a bi-level problem to identify reaction inhibition, activation or eliminations for maximizing the flux of biochemical target having combination of maximum biomass and minimum sum of fluxes as the objective function of inner problem (80). Maximizing the minimum production of target metabolite while having biomass is another way of formulating the bi-level problem for this purpose (95). By having maximum target production in the outer problem by a constraint on number of knock outs, or having minimum number of knockouts with a constraint on minimum target production one can formulate another bi-level formulation to study target overproduction. The objective of inner problem can be minimizing the production of target metabolite (96). In some papers, authors add minimum number of non-native reactions to the model so that it can produce some percentage or the maximum theoretical yield that is calculated in a previous step (97,98). OptOrf maximizes the desired production and minimizes the number of gene deletions or over productions with a penalty in the outer problem. In the inner problem, it maximizes biomass production (99). Many problems simply did single gene knock outs to find the effect of gene knock outs on production of desired metabolites (100–103). The correlation between reactions of the model and the target production reaction when the target reaction is maximized have also been studied to find highly correlated pairs (104). In addition, authors have maximized the availability of NADPH for natural product biosynthesis (105).

Studying diseases and drug target identification is another field of application of metabolic network models (22). A bi-level optimization framework has been used to infer oncoenzymes in hepatocytes. The outer objective of this problem is finding maximum similarity of model

fluxes and experimental measurements of Warburg effect and the inner objective is quadratic summation of fluxes (106). People also studied the effect of drugs on biomass production by limiting the flux through reactions that are associated to the drug (107).

Optimization methods have also been extensively used in improving the accuracy of models through gap filling (21). Finding the minimum number of reactions that are needed to be added to the model to have less blocked reactions and dead-end metabolites in the model is an example (108).

2.6.3 Classification based on the need for reformulation

Moreover, we analyzed if and how some of the computational problems in these papers needed reformulation based on the information in figure 2-1. About 60% of the methods that are part of our review need some sort of reformulation in order to have global optimality solving with current solvers. In other word, all the papers that are using FLP, MILFP, multi-objective and bi-level programming need reformulations. Details about these reformulations can be found in section 2-5, 2-7 and 2-8.

2.7 Bi-level optimization

Bi-level optimization-based procedures have been utilized in many strain design, gap filling and community modelling problems. The key concept in bi-level problems, is to use an inner and an outer level problem. The inner problem assigns fluxes in the network to optimize for the objective of interest of the inner problem in response to changes imposed by the outer level constraints. The outer problem optimizes another objective of interest by identifying appropriate changes to the network. Bi-level problems are mainly solved by converting them into a single-level LP or MILP (depending on the original problem type) by adding the dual constraints of the inner problem to the constraints of the inner and outer problems and imposing the strong duality condition. Strong duality is a condition, in which the value of the primal optimal objective function and the value of the dual optimal objective function are equal. The following problem is the single-level conversion of a bi-level optimization problem with only linear variable and constraints (as described in section 2-4 of this chapter):

$$\min_x c_1^T x$$

subject to:

$$\begin{aligned}
A_1 x &\leq b_1 \\
A_2 x &\leq b_2 \\
c_2^T x &= b_2^T y \\
A_2^T y &\geq c_2 \\
x, y &\geq 0
\end{aligned}$$

The first and second constraints are the constraints of outer and inner problems respectively. The third constraint applies the strong duality condition. The fourth constraint is the dual constraint associated to primal variables of the inner problem and the new y variable is the dual variable associated to the primal constraints.

Having binary variables in the original problem may make the conversion of the problem to a single-level problem more challenging. For instance, Bilinear terms may appear in the objective function of the dual problem. These terms can be linearized using the available standard techniques that were discussed previously in this chapter. In most of the bi-level methods, binary variables are the variables of the outer problem and are considered as parameters in the inner problem. However, there are some formulations in which the decision about the values of some of the binary variables are made in the inner problem. In these cases, the single-level conversions are more complex and need the use of complementary slackness condition. One can find more details and examples of the conversion of bi-level problems with binary variables in the following papers (80,109,110).

2.8 Multi-objective optimization

Since there is no single objective function that can comprehensively describe cellular behavior in a specific condition, several studies have considered multi-objective optimization for this purpose. In multi-objective optimization, one solution might be optimal for one objective, but suboptimal for the other. There are several methods to solve multi-objective optimizations. These methods include weighted sum method, ε -constraint method and the goal programming method. In the weighted sum method, one can simply assume weights for different objectives and then optimize for the sum of these weighted objectives (111).

$$\min_{x,y} (w_1 f_1(x, y) + w_2 f_2(x, y) + \dots + w_n f_n(x, y))$$

subject to:

$$Ax + By \leq b$$

$$x \geq 0$$

$$y \in \mathbb{Z}^n$$

In ε -constraint method, one objective function is chosen to be optimized and an upper or lower bound (whether the objective function is to be minimized or maximized) is considered for all other objective functions. These bounds are imposed by adding new constraints to the original problem (112).

$$\min_{x,y} (f_1(x,y))$$

subject to:

$$Ax + By \leq b$$

$$f_i(x,y) \leq \varepsilon_i \quad i = 2, 3, \dots, n$$

$$x \geq 0$$

$$y \in \mathbb{Z}^n$$

In goal programming approach, one can minimize the deviation of each objective function from a pre-defined goal. This goal can be the optimum value for that single function, even if having this optimum value is not feasible in the multi-objective problem (113).

$$\min \sum_i \delta_i^+ + \delta_i^-$$

subject to:

$$Ax + By \leq b$$

$$f_i(x,y) + \delta_i^+ - \delta_i^- = g_i$$

$$x \geq 0$$

$$\delta_i^+, \delta_i^- \geq 0$$

$$y \in \mathbb{Z}^n$$

Here, δ_i^+ and δ_i^- are the negative and positive violations from g_i , the optimum value (or goal) of i th objective function.

There are many other methods proposed for solving multi-objective problems and they are discussed in more details elsewhere (114,115). In our search, we found several combinations

of objective functions that have been used in the study of metabolic network models. For instance, Schuetz et al., used 55 single objective functions including maximum ATP, biomass, acetate and carbon dioxide yields, and maximum sum of absolute fluxes to study microbial metabolism (71). They also studied all pairs and triplets of these single objectives to further investigate what combination of objectives can better explain the physiology of the studied cells (71). Multi-objective optimization has also applications in strain design for simultaneous maximization of biomass and desired product (92,116) and minimization of number of knockouts (116).

2.9 Effect of adding thermodynamic constraints on problem types

Thermodynamics-based flux analysis (TFA) is an MILP formulation which confines the directionality of reactions by introducing thermodynamic constraints. This is done by taking into account Gibbs free energy of reactions and metabolite concentrations (5). Therefore, a TFA problem is more constrained compared to a FBA problem and the results are more precise due to refined reaction directionalities. Binary variables are introduced in TFA to link Gibbs free energy of reactions to flux of reactions in the network. Therefore, there is a change from LP to MILP when we convert a FBA problem to a TFA problem. Introducing thermodynamic constraints can be done in all of the methods that we have discussed in previous sections. In other words, all the problems can have a thermodynamically curated version where directionality of reactions are constrained based on TFA formulation. If the original problem is already having integer variables, i.e., MILP or MIQP, then introducing thermodynamic constraints have no effect on the problem type. However, if the problem is an LP or QP, it will be converted to MILP or MIQP after introducing thermodynamic constraints.

An example of this change in problem type can be seen when one wants to compute flux couplings in a thermodynamically constrained metabolic network model. Original flux coupling analysis is a fractional programming problem, which can be converted to linear programming using Charnes-Cooper transformation as explained in section 2-5 (8). Thermodynamic flux coupling analysis (TFCA) however, is a MILFP which can be converted to a MILP problem using Glover's linearization scheme. TFCA is discussed in more detail in the next chapter. Optknock is another example of problems that require reformulation after

introducing thermodynamic constraints. Original Optknock formulation is a bi-level problem (20). Optknock includes integer variables that are defined in the outer problem and are used as parameters in the inner problem. In other words, the inner problem in Optknock does not impose any constraints on the integer variables. However, if the inner problem is a TFA formulation, then one needs to use complementary slackness condition to convert the bi-level problem to a single level problem with duality theorem. FOCAL and OptReg are the examples of methods that use the same approach in order to solve their bi-level formulations (80,109). As a conclusion, in order to globally optimize a thermodynamically constrained problem further reformulation might be needed due to introduction of integer variables.

2.10 Conclusion

In this chapter, we propose an inventory of objective functions that can be investigated using the general classes of optimization formulations. We provide classifications for the optimization problems that are generally used in the analysis of metabolic network models. We discuss about the possible conversion of classes to each other so that at the end every optimization method can be formulated as an LP, MILP, QP or MIQP. One of the most interesting classes of the problems are FLPs and MIFLPs. They can have many interesting applications in metabolic engineering and in the study of branching points in metabolism and calculation of optimal flux ratios between branches. Introducing binary variables in order to enable flux deletions will turn such problems to MILFPs. However, FLPs and MIFLPs are not used as widely as other classes, since the problems are not solvable using well-known solvers such as CPLEX or Gurobi. With the reformulations that we provide in this chapter, these classes can be easily converted to LPs and MILPs. Enabling gene or flux knock-outs are not the only usage of Integer variables in the study of metabolic networks. They can also link the thermodynamic properties of reactions and compounds such as Gibbs free energy of reactions and compound formations to the metabolic fluxes as it is done in TFA formulation (5). Introducing new thermodynamic constraints using integer variables Henry et al., could produce flux distributions that do not contain any thermodynamically infeasible reactions (5). We should emphasize that, with the proper reformulations as suggested in this chapter, one can easily expand the use of thermodynamic constraints in other optimization problems used in systems biology studies. We realized from our literature search that there is more tendency

toward using multi-objective and bi-level optimizations, recently. Although solving these problems are more challenging, it shows that we are heading toward the right direction, knowing the fact that not a single objective function can describe the metabolic goals of living cells. Further attempts should be made in order to solve these problems as efficiently and accurately as possible.

3 Flux coupling analysis framework for the study of metabolic network models that include integer variables

3.1 Author contribution and publications used in this chapter

The work of this chapter is based on the following paper which is under preparation for submission:

Z. Hosseini, V. Hatzimanikatis (2020) “A Flux coupling analysis framework for the study of metabolic network models that include integer variables”.

3.2 Introduction

The number of reconstructed genome scale metabolic models has increased tremendously in recent years. These metabolic models include all known and hypothetical biochemical reactions that can operate in a cell and contain all the information about metabolic reactions and metabolites that are present in a cell (4). Different properties of these models, such as metabolite concentrations or reaction fluxes are not completely independent of each other and the structure of the network determines how these components are coupled (8,117). For example, due to stoichiometric coupling, certain fluxes can always correlate or anti-correlate, while others can correlate or anti-correlate depending on the conditions and on the kinetic properties. Knowledge of such coupling relations is important for interpreting physiology and metabolic engineering, and computational approaches are required to identify and characterize these couplings.

Flux coupling analysis (FCA) is one of the computational approaches that is developed to identify and characterize the coupling between any pair of metabolic reactions, and set of coupled reactions, in metabolic network models (8). A couple of methodologies have been

developed so far for computing flux coupling relations in metabolic network models. The original flux coupling finder (FCF) algorithm uses a fractional programming to calculate the ratio between two fluxes. If this ratio for two reactions is always fixed, then these reactions are fully coupled. If it is not fixed, then there is a range of classification for the type of coupling. The FCF belongs to the class of linear programming problems known as linear fractional programming (LFP) problems. Some methods have been developed to improve the performance and speed of the FCF formulation (118,119).

FCA has been used in a large number of applications so far: studying gene evolution (120), identifying genetic interactions (121) and gene/reaction essentiality (122), gap filling of metabolic networks (123) etc. Investigation of the coupling properties of metabolic networks is important to both understand the organizational principles of metabolic interactions within metabolic networks, and also for more efficiently suggesting engineering modifications (8,109). Therefore, FCA is a useful method for both computational biologists and experimentalists.

Many useful algorithms for metabolic network analysis and optimization are based on mixed-integer LP (MILP) formulations (5,33,35,49,97,109). For example, ROOM is a MILP algorithm that identifies the number of flux changes compared to the wild-type after a gene knock-out (33). Another MILP formulation is introduced for identifying the minimal set of metabolic reactions that can support growth on different substrates (49). FOCAL is another MILP-based method that identifies conditions such that a reaction become essential for a specific phenotype (109). However, current versions of FCA do not account for integer variables. Here, we introduce MIFCA (mixed integer flux coupling analysis), a new formulation for computing flux couplings in a problem with integer variables. MIFCA is based on a mixed integer linear fractional programming (MILFP), which is converted to an MILP, based on some proposed problem reformulations (11).

One of the important classes of MILP formulations is the thermodynamics-based flux analysis (TFA) (4) which is also referred to as thermodynamics-based metabolic flux analysis (TMFA) (5) and thermodynamics-based flux balance analysis (TFBA) (124,125). In TFA, the directionality of metabolic reactions is constrained based on their Gibbs free energy which also accounts for metabolite concentrations (5). Thus, by applying thermodynamic constraints one can define reaction directionalities more precisely and classify the reactions as bidirectional and unidirectional, depending on the thermodynamically feasible range of

their directionality (4,124,126). The application of thermodynamic constraints within TFA decreases the possible flux space (4,124), allows us to study the energetic feasibility of native and heterologous pathways (4,124), and therefore it could result in more truthful predictions (9,127).

We also introduce here TFCA (thermodynamics flux coupling analysis) as a new class of problems which is based on the MIFCA formulation and it includes thermodynamic constraints. We show that when thermodynamic constraints render some reactions unidirectional, the coupling classification of the reactions is also changing. These effects and the properties and performance of the TFCA formulation are illustrated in a metabolic network model of *E. coli*.

The thermodynamic constraints in metabolic network models constrain many fluxes to a unique direction and thus they reduce significantly the solution space (4,124,126). However, we still have some bidirectional reactions (BDR), which in turn result in alternative flux directionality profiles (FDPs). FDPs are important when we want to study a specific physiology or in study of kinetic models, when one must reduce the complexity of the system as much as possible. Theoretically, with n BDRs we can have 2^n FDPs, which makes it difficult to analyze the network. However, the degrees of freedom of the network are less, since the activities of some reactions are coupled to the activities of others. We show how coupling relationships of the reactions can capture this reduction in the solution space.

3.3 Methods

3.3.1 Problem formulation for FCA

In FCA algorithm the minimum and maximum ratio of every pair of reactions are calculated. Because the problem is a fractional program, a transformation in variables is needed to convert it to a LP problem. The details of this transformation is discussed in (8) and the formulation of FCA is shown below:

For any two fluxes, v_1 and v_2 , the maximization or minimization of their respective ratios in a metabolic model is described mathematically as (8):

$$\text{Maximize (minimize)} \ v_1/v_2$$

$$\text{Subject to } \sum_{j=1}^M S_{i,j} v_j = 0 \quad \forall i \in N \quad (1)$$

$$v_j^{uptake} \leq v_{j,max}^{uptake} \quad \forall j \in M_{transport} \quad (2)$$

$$v_j \geq 0 \quad \forall j \in M \quad (3)$$

To convert the problem to a linear problem a variable t is multiplied to both denominator and numerator of the objective function and to both sides of each constraint. Then the following problem is obtained:

$$\text{Maximize (minimize) } v_1 \cdot t / v_2 \cdot t$$

$$\text{Subject to } \sum_{j=1}^M S_{i,j} v_j \cdot t = 0 \quad \forall i \in N \quad (4)$$

$$v_j^{uptake} \cdot t \leq v_{j,max}^{uptake} \cdot t \quad \forall j \in M_{transport} \quad (5)$$

$$v_j \cdot t \geq 0 \quad \forall j \in M \quad (6)$$

$$t \geq 0 \quad (7)$$

The following problem is equivalent to the above problem with proper variable change and addition of a new constraint:

$$\text{Maximize (minimize) } \tilde{v}_1$$

$$\text{Subject to } \tilde{v}_2 = 1 \quad (8)$$

$$\sum_{j=1}^M S_{i,j} \tilde{v}_j = 0 \quad \forall i \in N \quad (9)$$

$$\tilde{v}_j^{uptake} \leq \tilde{v}_{j,max}^{uptake} \cdot t \quad \forall j \in M_{transport} \quad (10)$$

$$\tilde{v}_j \geq 0 \quad \forall j \in M \quad (11)$$

$$t \geq 0 \quad (12)$$

where S is the stoichiometric matrix in which the rows represent metabolites and the columns represent reactions, v . Each reaction through an enzyme has been split in two terms representing the forward and the backward reaction. This separation is performed so that the flux through each reaction can be constrained to be greater than or equal to zero. The $S_{i,j}$ element in S is the stoichiometric coefficient of each metabolite i in each reaction j . The vector \tilde{v} is the fluxes of reactions multiplied by t and t is the auxiliary variable that is used for transformation of the LFP into LP.

3.3.2 Problem formulation for MIFCA

In models with integer variables, the original FCF variable transformation results in bilinear terms in the formulation, which leads to a mixed integer linear fractional programming (MILFP) problem. MILFPs are a class of optimization problems with a fractional objective function and integer variables. MILFPs can be converted to MILPs (128) in the formulation described below.

In the most general case, we consider the reactions to be catalytically reversible. This implies that the enzymes are able to catalyse the reactions in both directions: the *forward* and *backward* reactions. The directionality of the reaction depends then on the displacement of the reaction from thermodynamic equilibrium. Therefore, in the problem formulation, for each reaction i we define the net flux V_i as the difference between two fluxes, V_i^f and V_i^b , the net-forward and the net-backward flux, respectively, with the reference direction chosen arbitrarily. In addition, for each reaction i we introduce two binary variables, z_i^f and z_i^b for the net fluxes V_i^f and V_i^b , respectively. We then have the following constraints for the variables that correspond to each reaction i :

$$V_i = V_i^f - V_i^b \quad (13)$$

$$-V_i^{b_max} \leq V_i \leq V_i^{f_max} \quad (14)$$

$$0 \leq V_i^f \leq z_i^f \cdot V_i^{f_max} \quad (15)$$

$$0 \leq V_i^b \leq z_i^b \cdot V_i^{b_max} \quad (16)$$

$$z_i^f + z_i^b \leq 1 \quad (17)$$

Equations (13) to (16) are the bounds for the net fluxes and equation (17) guarantees the solution admits only one feasible directionality. If a reaction is known, or hypothesized, to be catalytically unidirectional the corresponding z variables are set accordingly. For example, if reaction V_j is catalytically feasible in the backward direction, then we set $z_j^b = 1$ and $z_j^f = 0$. For any two fluxes, v_1 and v_2 , the maximization or minimization of their respective ratios in a general metabolic model with integer variables is described mathematically as:

$$\begin{aligned} & \text{Minimize (maximize) } v_1/v_2 \\ & \text{Subject to } S \cdot v = 0 \end{aligned} \quad (18)$$

$$0 \leq v_i \leq z_i v_{max} \quad \{i = 1, \dots, r\} \quad (19)$$

$$z \in \{0,1\} \quad (20)$$

Here S is the stoichiometric matrix, v is the flux vector, and r is the total number of reactions. Here, the reactions are split into forward and backward directions and therefore the variables show net forward and net backward fluxes. z_i is a binary variable associated to each reaction flux. As described before, there is a z variable for each backward and forward reaction, v . For each reaction if z_i is equal to zero, then v_i is zero.

By multiplying the numerator and denominator of the objective function as well as all constraints by a positive variable t , an equivalent problem is obtained, which is an MILFP problem.

$$\text{Minimize (maximize)} \quad v_1 \cdot t / v_2 \cdot t$$

$$\text{Subject to } S \cdot v \cdot t = 0 \quad (21)$$

$$0 \leq v_i \cdot t \leq z_i \cdot t v_{max} \quad \{i = 1, \dots, r\} \quad (22)$$

$$z \in \{0,1\} \quad (23)$$

Then, the MILFP problem can be transformed into the following equivalent MINLP problem:

$$\text{Minimize (maximize)} \quad \tilde{v}_1$$

$$\text{Subject to } \tilde{v}_2 = 1 \quad (24)$$

$$S \cdot \tilde{v} = 0 \quad (25)$$

$$0 \leq \tilde{v}_i \leq z_i \cdot t v_{max} \quad \{i = 1, \dots, r\} \quad (26)$$

$$z \in \{0,1\} \quad t \geq 0 \quad (27)$$

The only nonlinear terms in the above MINLP problem are the bilinear terms $z_i \cdot t$, which are the products of a binary variable z_i and a continuous variable t . This type of bilinear terms can be linearized by introducing a number of auxiliary variables and constraints, following the Glover's linearization scheme (24,128). We introduce a set of auxiliary variables w_i , for each bilinear term such that $w_i = z_i \cdot t$. Thus, the MINLP problem can be further linearized into an equivalent MILP problem, which is given below:

$$\text{Minimize (maximize)} \quad \tilde{v}_1$$

$$\text{Subject to } \tilde{v}_2 = 1 \quad (28)$$

$$S \cdot \tilde{v} = 0 \quad (29)$$

$$0 \leq \tilde{v}_i \leq w_i v_{max} \quad \{i = 1, \dots, r\} \quad (30)$$

$$w_i \leq t \quad \{i = 1, \dots, r\} \quad (31)$$

$$w_i \leq M \cdot z_i \quad \{i = 1, \dots, r\} \quad (32)$$

$$w_i \geq t - M \cdot (1 - z_i) \quad \{i = 1, \dots, r\} \quad (33)$$

$$z \in \{0,1\} \quad t \geq 0 \quad w \geq 0 \quad (34)$$

where M is a sufficiently large number. Constraint (32) ensures that if z_i is zero, then w_i should be zero; constraints (31) and (33) imply that if z_i is one, then w_i should be equal to t .

3.3.3 Problem formulation for TFCA

Since in thermodynamically constrained models there are integer variables, FCA in these models can be considered as a MIFCA problem. The final MILP formulation for TFCA after variable transformation and linearization is as follows:

Minimize (maximize) \tilde{v}_1

$$\text{Subject to} \quad \tilde{v}_2 = 1 \quad (35)$$

$$S \cdot \tilde{v} = 0 \quad (36)$$

$$0 \leq \tilde{v}_i \leq w_i v_{max} \quad \{i = 1, \dots, r\} \quad (37)$$

$$\widetilde{\Delta_r G'_i} - K \cdot t + K w_i < 0 \quad \{i = 1, \dots, r \mid \Delta_r G_i^{\circ} \text{ is known}\} \quad (38)$$

$$\Delta_r G_i^{\circ} \cdot t + RT \sum_{j=1}^m n_{i,j} \ln(\widetilde{x_j}) = \widetilde{\Delta_r G'_i} \quad \{i = 1, \dots, r + L \mid \Delta_r G_i^{\circ} \text{ is known}\} \quad (39)$$

$$\widetilde{\Delta_r G'_i} - K w_i < 0 \quad \{i = r + 1, \dots, r + L\} \quad (40)$$

$$w_i + \sum_{j=1}^r \alpha_{i,j} w_j \leq \sum_{j=1}^r \alpha_{i,j} \cdot t \quad \{i = r + 1, \dots, r + L\} \quad (41)$$

$$w_i \leq t \quad \{i = 1, \dots, r + L\} \quad (42)$$

$$w_i \leq M \cdot z_i \quad \{i = 1, \dots, r\} \quad (43)$$

$$w_i \leq M \cdot y_i \quad \{i = r + 1, \dots, r + L\} \quad (44)$$

$$w_i \geq t - M \cdot (1 - z_i) \quad \{i = 1, \dots, r\} \quad (45)$$

$$w_i \geq t - M \cdot (1 - y_i) \quad \{i = r + 1, \dots, r + L\} \quad (46)$$

$$w_i^f + w_i^b \leq t \quad \{i = 1, \dots, R\} \quad (47)$$

$$z, y \in \{0,1\} \quad (48)$$

(Equations (35), (36), (37), (42), (43) and (44), and (45) and (46) in TFCA are the same as (28), (29), (30), (31), (32) and (33) in MIFCA, respectively). Constraint (47) is equal to constraint (17) and it shows that for each original reaction in R (before splitting) only one of the forward or backward reactions can carry flux.

Here S is the $m \times r$ stoichiometric matrix, v is the flux vector, m is the total number of metabolites and r is the total number of split reactions. z_i is a binary variable associated to each reaction flux. v_i can carry flux if z_i is equal to one. If z_i is equal to zero, then the flux through reaction i should be zero. Equation (38) ensures that a reaction flux cannot be positive unless $\Delta_r G'_i$ is negative. Equation (39) associates Gibbs free energy of reactions to metabolite activities. Equation (40) is the thermodynamic feasibility constraint for the reactions with unknown $\Delta_r G'^\circ$ and equation (41) excludes flux distributions that involve flux through the set of reactions that comprise an infeasible lumped reaction. y_i is the binary variable which is set to zero if the lumped reaction is thermodynamically feasible and one otherwise. In these equations, L is the number of lumped reactions, and $\alpha_{i,j}$ is a coefficient equaling one if reaction j is one of the original reactions with unknown $\Delta_r G'_i$ that makes up the lumped reaction i . The detailed explanation about these constraints is available in (5). Overall, the TFCA formulation adds only a number of linear variables and constraints into the MIFCA formulation.

3.3.4 Metabolic models and experimental data

We apply the proposed methods on metabolic network of *Escherichia coli*, for which we have standard datasets of metabolite concentrations (129). We applied TFCA for a reduced model of the *E. coli* genome-scale model *iJO1366* (130). The reduced model was produced by Dr. Georgios Fengos using the reduction methods redGEM and lumpGEM (131,132). Briefly, having central metabolism and subsystems of interest, redGEM reduces genome-scale models into core models in a consistent manner. It uses graph-based approaches and optimization methods to minimize the loss of information (131). lumpGEM, on the other hand, generates subnetworks for the production of any target metabolite of interest in the model, from pre-defined core precursors. Every metabolite and cofactor that is part of the generated subnetwork is stoichiometrically balanced. LumpGEM can also merge all the reactions of the subnetwork into elementally balanced lumped reactions (132). Together,

lumpGEM and redGEM can be used for generation of balanced reduced core metabolic network models.

3.3.5 Coupling of subnetworks

In reduced metabolic network model of *E. coli* that we are using in this study, there are lumped reactions that link the core reactions of the model to biomass reaction. The core reactions are reactions that belong to one of the following subsystems: (i) Glycolysis/Gluconeogenesis, (ii) Pentose Phosphate Pathway, (iii) Pyruvate Metabolism, (iv) Citric Acid Cycle, (v) Glyoxylate Metabolism, and (vi) Electron Transport Chain. For production of each biomass building block several reactions are needed which are merged together to form lumped reactions. This group of reactions are called subnetworks and after the addition of lumped reactions to the model they are removed from the reduced model. In our case, the subnetworks are with minimum size and are not unique, i.e., there may be more than one minimal subnetwork for production of a biomass building block. For analyzing coupling between subnetworks, we removed the biomass reaction and added boundary reactions for biomass building blocks. Lumped reactions are also removed and the corresponding subnetwork is added back to the model. We then computed the coupling between every pair of reactions that are parts of the subnetworks. We define two types of degrees of coupling, $D_1(a, b)$ and $D_2(a, b)$, between any pairs of subnetworks a and b , as follows:

$$D_1(a, b) = \frac{\text{number of coupled reaction pairs that only one of them is in subnetwork } a \text{ and the other is in subnetwork } b}{\text{total number of not common pairs between } a \text{ and } b}$$

$$D_2(a, b) = \frac{\text{number of all coupled reaction pairs that one of them is in subnetwork } a \text{ and one in subnetwork } b}{\text{total number of reaction pairs between } a \text{ and } b}$$

It should be noted that in $D_2(a, b)$ all the coupled reaction pairs between subnetworks are taken into account even if they are common between the two subnetworks. However in $D_1(a, b)$, only the couplings between uncommon reaction pairs are counted. To analyze the coupling within each subnetwork we also define $D_3(a)$ for any subnetwork a as follows:

$$D_3(a) = \frac{\text{number of all coupled reaction pairs that both of them are in } a}{\text{total number of reaction pairs in } a}$$

In order to analyze how optimization of production of each biomass building block affects the

degrees of coupling between subnetworks, we maximized the production of each biomass building block, fix the lower and upper bounds of the corresponding reaction to the maximum value and then computed all reaction couplings again.

3.4 Results and Discussion

3.4.1 Classification of flux couplings

The different types of couplings that can arise when we compute the coupling relationships for each pair of reactions in the network have been classified as follows (8):

- (i) Directional coupling ($v_a \rightarrow v_d$), if a non-zero flux for v_a implies a non-zero flux for v_d but not the reverse.
- (ii) Partial coupling ($v_a \leftrightarrow v_d$), if a non-zero flux for v_a implies a non-zero, though variable, flux for v_d and vice versa.
- (iii) Full coupling ($v_a \Leftrightarrow v_d$), if a non-zero flux for v_a implies not only a non-zero but also a fixed flux for v_d and vice versa.

Fully coupled reactions are then two reactions that are always active simultaneously and the ratio of their fluxes is always a fixed value (Figure 3-1). However, when we consider thermodynamics, it is possible that only one direction for one of the fluxes is allowed, and this can reduce the ranges of the fluxes.

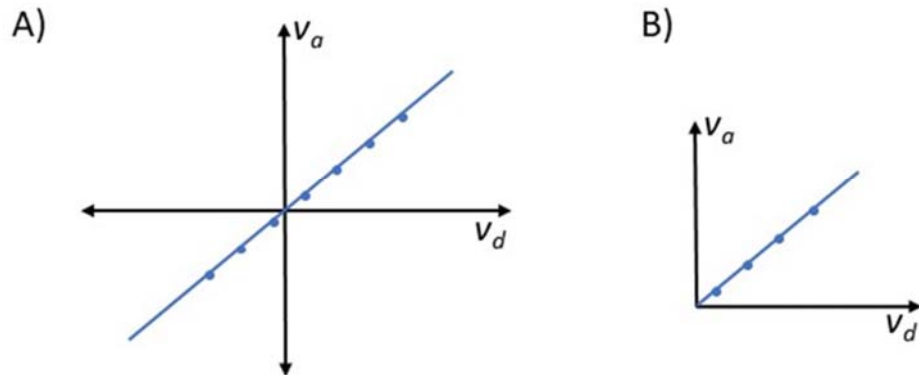


Figure 3-1. Full coupling in metabolic networks. A) Full coupling in bidirectional reactions. In this case, simultaneous existence of positive direction of v_a and negative flux of v_d is not feasible. In addition, simultaneous existence of positive flux of v_d and negative flux of v_a are also not possible (it can be vice versa, i.e., full coupling of positive flux of v_a and negative flux of v_d and full coupling of negative flux of v_a and positive flux of v_d). B) Full coupling in unidirectional reactions.

Partially coupled reactions are two reactions that are always active *simultaneously* but the flux ratio of the two reactions can change between two values (Figure 3-2). This further implies, that the feasible ranges of one of the fluxes, v_a (or v_d), depends on the value of the coupled flux, v_d (or v_a). In other words, for a fixed value of one of the fluxes, v_a (or v_d), the other flux, v_d (or v_a) is constrained with a range that depends on the value of the fixed flux. For the case of partial coupling, the solution space is always constrained in a single quadrant.

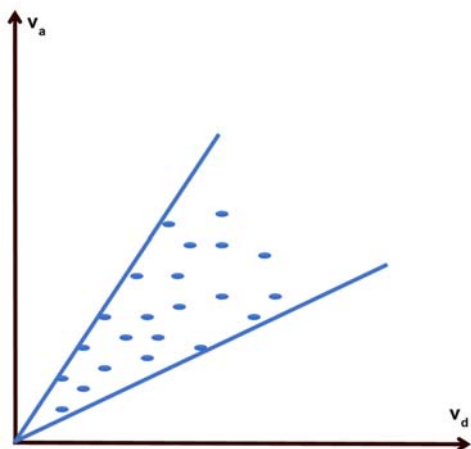


Figure 3-2 Partial coupling in metabolic networks. Partial coupling can only happen for two unidirectional reactions

In a pair of *directionally coupled* reactions the activity of only one of the reactions is dependent on the activity of the other (Figure 3-3). If one of the reactions is bidirectional we will have panel A. In this case, positive and negative fluxes of the bidirectional reaction (v_a) are dependent to the flux of unidirectional reaction (v_d). If, we have two unidirectional reactions either due to stoichiometry alone or due to thermodynamic constraints, and v_a is directionally coupled to v_d then we will have panel B.

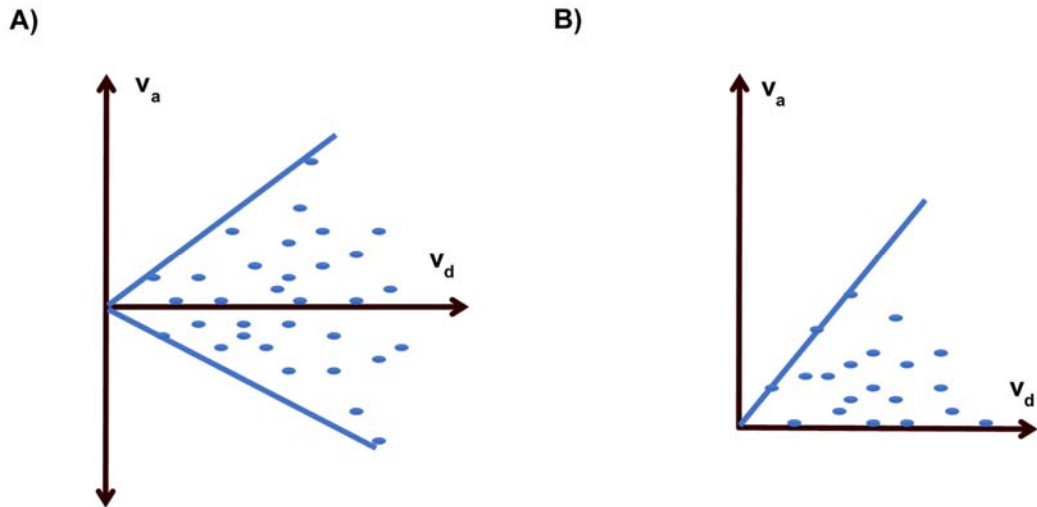


Figure 3-3. Directional coupling in metabolic networks. A) A bidirectional reaction that is directionally coupled to a unidirectional reaction, which means its flux in either direction is coupled to the flux of unidirectional reaction. B) v_a is directionally coupled to reaction v_d .

Uncoupled Reactions are reaction pairs not falling into one of three categories (i)-(iii). Uncoupled reactions can happen for all combinations of unidirectional or bidirectional reactions and the activities of none of the reactions in the pair is coupled to the activity of other (Figure 3-4).

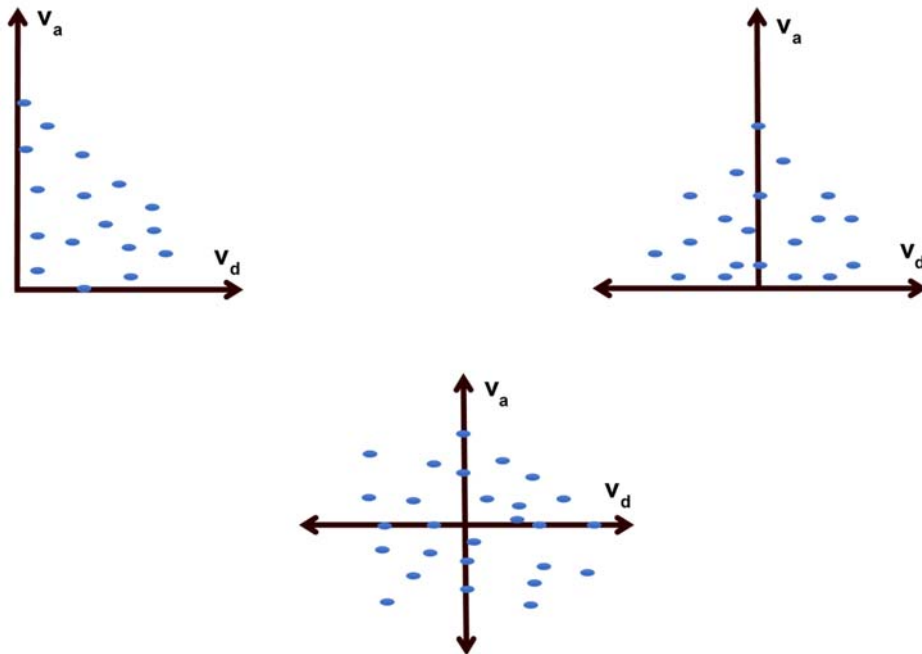


Figure 3-4. Uncoupled reaction pairs in metabolic networks.

Conditional (directional) coupling is the new case we introduce here and corresponds to the case when the type of coupling of the reaction is conditional to the directionality of the fluxes. In other words, the type of coupling may change when at least one of the fluxes is switching direction. This additional classification affects only the uncoupled reactions pairs, i.e., these pairs were classified as uncoupled in the original flux coupling formulation. This case is illustrated in Figure 3-5 with two examples that involve bidirectional fluxes. As it is shown in the panel (B), we will have 4 cases; (i) The positive direction of v_a is directionally coupled to positive direction of v_d ; (ii) Negative direction of v_d is directionally coupled to negative direction of v_a ; (iii) Positive direction of v_d and negative direction of v_a are not coupled to each other; (iv) and finally the combination of negative direction of v_d and positive direction of v_a is not possible. If only one of the reactions is bidirectional (panel A), positive direction of v_a is directionally coupled to v_d , but negative direction of v_a is uncoupled to v_d . The right panel is particularly important, because it can tell us that the number of combinations of bidirectional reactions in this case is 3 instead of 4.

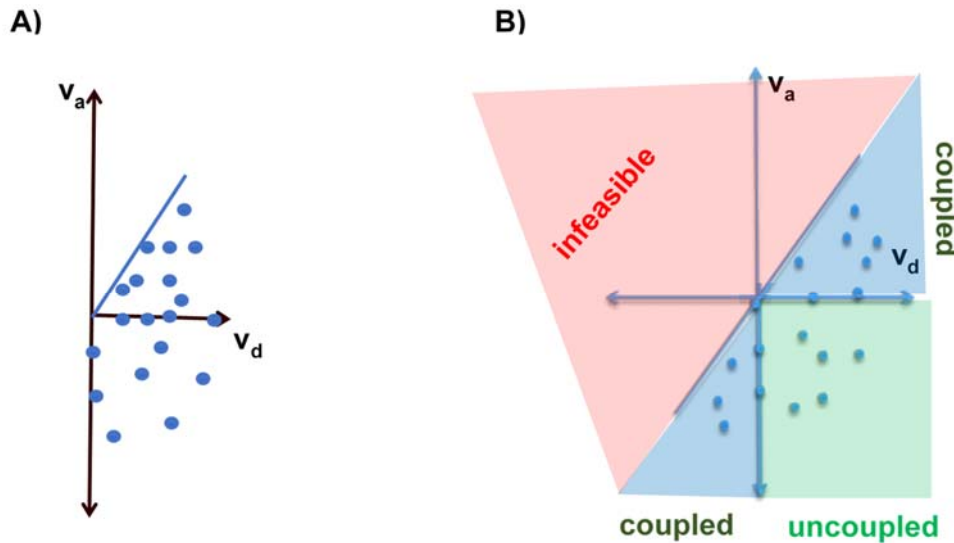


Figure 3-5. Conditional (directional) coupling in metabolic networks. A) only positive direction of v_a is coupled to v_d . B) positive direction of v_a is directionally coupled to positive direction of v_d and negative direction of v_d is directionally coupled to negative direction of v_a , negative direction of v_a is uncoupled to positive direction of v_d and the combination of negative direction of v_d and positive direction of v_a is not feasible.

In general, it is possible to have a general classification of coupling relationships and one that considers the directionality of the reactions. Classification based on directionalities is important since in the cells the fluxes operate under a specific flux profile with specific flux

directionalities. Therefore, to understand cellular physiology it is important to draw conclusions that are based on the flux directionalities.

3.4.2 Integration of data in different levels in reduced *E. coli* model and identification of bidirectional reactions

We studied a reduced *E. coli* model derived from the genome-scale model of Orth *et al.* (130), using the model reduction methods redGEM and lumpGEM (131,132). The reduced model has 322 intracellular metabolites and 201 intracellular reactions. The model described explicitly 6 subsystems, as originally defined in the genome-scale model: (i) Glycolysis/Gluconeogenesis, (ii) Pentose Phosphate Pathway, (iii) Pyruvate Metabolism, (iv) Citric Acid Cycle, (v) Glyoxylate Metabolism, and (vi) Electron Transport Chain.

In our analysis here, we considered the physiology studies of *E. coli* growth in a batch reactor under aerobic condition reported by McCloskey *et al.* (129), and we integrated the experimental data for specific uptake fluxes, specific production fluxes, and specific growth rates. We included data about the standard Gibbs free energy of reaction, $\Delta G'^{\circ}$, which allowed to consider the effects of thermodynamic feasibility on the allowable directionality of the reaction fluxes, and to include the information for the metabolite concentration.

When considering thermodynamic feasibility, we can assume a general range for the concentrations based on prior knowledge and tighter ranges for the measured metabolites in the studied physiology. We have to stress that the metabolites measurements are partial, i.e., measurements data are not available for all the metabolites of the model.

We used the reduced model, since this model is less complex and at the same time is consistent with the genome-scale model of *E. coli*. Then, we integrated experimental data in different levels to this model to see the impacts of different constraints on the coupling of reaction pairs in the model.

We considered four different cases that depend on the type of metabolite data used and the assumption about reaction directionality and the model constraints: (i) Model *NT*: all reactions are reversible and there is no thermodynamics-based constraint; (ii) Model *ToC*: all reactions are reversible with thermodynamics-based constraints; (iii) Model *TcC*: all reactions are reversible with thermodynamics-based constraints and ranges for the measured metabolites; and (iv) Model *TcCR*: pre-assigned directionalities for the irreversible reactions

according to the reference genome-scale model. The rest of the reactions considered reversible with thermodynamics-based constraints and ranges for the measured metabolites. For all models we imposed the constraints on the specific uptake fluxes, specific production fluxes, and specific growth rates. Model NT is a regular FBA model, and models ToC, TcC, and TcCR are TFA models with different constraints on thermodynamics and intracellular fluxes. Table 3-1 summarizes the information integrated in the models in each step.

In metabolic models the reactions can be classified as catalytically irreversible and catalytically reversible (4). Catalytically irreversible reactions have pre-assigned directionality and they can operate only if they are thermodynamically feasible in the catalytically feasible direction. Catalytically reversible reactions can operate in both directions depending on the constraints imposed on the network. For a specific physiology and set of constraints, the reversible reactions can be classified as bidirectional if, subject to the constraints, they can operate in both directions based on a bidirectionality analysis. In this analysis, for every reaction we perform a flux variability analysis and if a reaction can operate in both directions is classified as bidirectional reaction (BDR). For the four models introduced here, we also performed a bidirectionality analysis based on the Gibbs free energy of reactions.

last column of the table 3-1 shows the number of bidirectional reactions. The third columns in the middle show the kind of data that are integrated on the model in each step.

Table 3-1. Integrating information in different levels to reduced model of *E. coli*

Model	Thermodynamic constraints	Experimental concentrations	Pre-assigned directionalities	# of BDRs
NT	-	-	-	19
ToC	+	-	-	15
TcC	+	+	-	15
TcCR	+	+	+	7

3.4.3 Changes in couplings after constraining the models

We investigated the impact of addition of constraints to the models in each step. The following three tables show the changes in flux couplings when we move from one state of the model to the other. Table 3-2 shows the transition from non-thermo model to the

thermodynamically constrained model with open concentrations. As it can be seen, there are a large number of changes in flux coupling relations. For instance, there are 200 reaction pairs that their coupling changed from uncoupled to directionally coupled. This shows that the effect of thermodynamic constraints on other parts of the network make the change on the coupling type of these pairs, since the directionality of the reaction pair itself is not changed. We can also see that there are 15 reaction pairs that their coupling relationships change from uncoupled to conditionally directionally coupled. In these cases, changes in the directionality of reaction pair itself has an effect in the change in coupling relationship. In total, this table shows that thermodynamic constraints have a very important effect on the coupling relationships in the model.

Table 3-2. changes in coupling relationships when we add thermodynamic constraints to a non-thermo model

		ToC									
NT											
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	2	25	-	-	-	54	-	-
		-	-	-	1	-	12	-	2	-	-
		-	-	-	200	-	-	-	-	-	-
		-	-	-	-	-	45	-	-	-	-
		-	-	-	1	-	15	-	-	-	-

Table 3-3 shows the changes in couplings when we introduce the concentration data to the thermodynamically constrained model. Even in this table we see some changes in the couplings, but the number of changes are not as large as the previous table. That is because of the fact that thermodynamic constraints have already their effect on the network.

Table 3-3. changes in coupling relationships when we constrain the metabolite concentrations in the thermodynamically constrained model






















		TcC									
ToC											
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	3	131	-	-	-	-	-	-
		-	-	-	-	-	2	-	-	-	-
		-	-	-	-	-	-	-	-	-	-

Table 3-4, shows the transition from the model with constrained concentrations to the model with having additionally the data of flux directionalities, which is gathered from different papers. As it is shown, we have again some changes from uncoupled reaction pairs to directionally coupled reaction pairs. The interesting change that is seen in this table is the change from bidirectional uncoupled pairs to bidirectional pairs that are conditionally directionally coupled. Since the directionality of reactions themselves didn't change, again it shows that changes in other, maybe remote parts of the network, makes these changes in coupling. This is particularly interesting for analyzing FDPs, since it will help in reducing the number of FDPs in the network, which is discussed in the next section.

Table 3-4. changes in coupling relationships when we constrain reaction directionalities in the thermodynamically constrained model with constrained concentrations

		TcCR									
TcC											
		-	-	-	-	-	-	-	-	-	-
		1	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	70	-	-	-	59	-	-
		-	-	-	-	-	-	-	-	-	-
		-	-	-	266	-	-	-	-	-	-
		-	-	-	3	-	8	-	-	-	-
		-	-	-	-	-	-	2	-	-	-

3.4.4 Coupling and FDPs

In this section, we will discuss how finding the coupling relations can help in reducing the number of FDPs in thermodynamically constrained metabolic network models. Table 3-5 shows the summary of the results.

Second column of the table shows the number of bidirectional reactions in the models. The third column shows the theoretical number of flux directionality profiles that can exist in the models. As we expect, some of the bidirectional reactions are coupled to each other, and therefore not all of the theoretical FDPs are possible. The number of coupled bidirectional reaction pairs of each model and the number of reduced FDPs are presented in the third and fourth columns, respectively. For instance, in the last and most constrained model, the number of FDPs is changed from 128 (2^7) to 30 ($2^{4.9}$). This reduction in the number of FDPs is helpful in reducing the complexity of the model and making the study of the physiology of the cells much easier.

Table 3-5. Impact of having coupled reaction pairs in reducing the number of FDPs in the model

Model	BDRs	Theoretical FDPs	Coupled pairs* (fully coupled and conditionally directionally coupled pairs)	Reduced FDPs
NT Non-thermo	19	2^{19}	22	$2^{14.23}$
ToC Thermo-open concentrations	15	2^{15}	7	$2^{12.16}$
TcC Thermo-constrained concentrations	15	2^{15}	7	$2^{12.16}$
TcCR Thermo-constrained concentrations and fluxes	7	2^7	9	$2^{4.9}$

In the most constrained model, we started to fix bidirectional reactions in forward and reverse directions one by one, and then compute the coupling relations again. The results are shown in Table 3-6. As we can see fixation of some of the reactions such as Transketolase1, Ribulose 5-phosphate 3-epimerase and Fumarase in reverse direction, will result in a great reduction of FDPs. While, for reactions such as Transketolase2, Transaldolase and Triose phosphate isomerase, fixation of reactions in any direction would have almost the same effect on the number of FDPs. These kinds of analysis can help in making decisions on how to rank FDPs for different studies. Particularly, it can help when we are interested in reducing the number of FDPs to a very small number such as in analyzing kinetic models of metabolism, where the number of unknown parameters is already very high.

Table 3-6. the impact of fixing bidirectional reactions to forward or reverse direction, one by one in the most constrained model

Fixed direction	# of BDRs	# of coupled pairs	# of FDPs
Nothing Fixed	7	9	30
R_ Transketolase1 (R_TKT1)	1	0	2
F_ Transketolase1 (F_TKT1)	6	4	28
R_ Ribulose 5-phosphate 3-epimerase (R_RPE)	3	1	6
F_ Ribulose 5-phosphate 3-epimerase (F_RPE)	5	1	24
R_ Fumarase (R_FUM)	3	0	8
F_ Fumarase (F_FUM)	6	6	22
R_ Transaldolase (R_TALA)	6	8	16
F_ Transaldolase (F_TALA)	5	4	14
R_ Transketolase2 (R_TKT2)	5	4	14
F_ Transketolase2 (R_TKT2)	4	0	16
R_ Triose phosphate isomerase (R_TPI)	6	9	15
F_ Triose phosphate isomerase (F_TPI)	6	9	15

To further classify the FDPs, we constructed a FDP tree that shows how we can reduce the number of FDPs step by step by fixing reactions to one direction (Figure 3-6). In this tree, the number in the root node shows the number of FDPs in the model. By going through each edge we fix directionality of one of the bidirectional reactions to a specific direction. The output node of each edge shows the reduced number of FDPs after the directionality fixation. In the leaves of the tree we cannot reduce the number of FDPs anymore, since the remaining bidirectional reactions are not coupled to each other anymore. In other words, they can be active in either direction independent of each other.

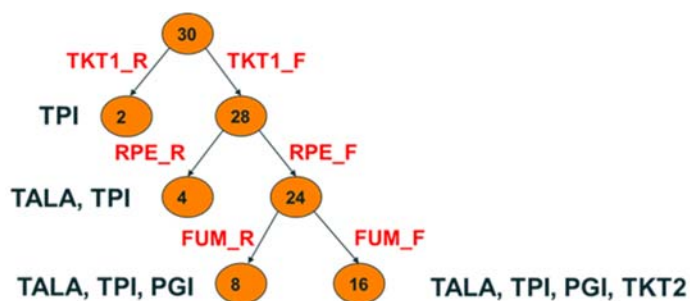


Figure 3-6. Flux directionality profile tree. This is the FDP tree in the model with constraints on metabolite concentrations and flux directionalities. Numbers on the nodes show the number of FDPs after fixing the directionality of some reactions to one specific direction

3.4.5 Coupling of subnetworks

In this section, we study how different parts of the metabolism are coupled to each other. For this purpose, we studied the coupling between different pairs of subnetworks that are corresponding to production of different biomass building blocks in the model. One should note that there may be several combination of lumped reactions that satisfy minimum growth requirement and at the end only one set of these lumped reactions are used for model reduction. To study subnetwork coupling we need to add back subnetworks corresponding to each lumped reaction to the model and remove the lumped reaction. However, since we can have different lumped reactions for each building block which are corresponding to different (or similar) subnetworks, we also analyzed the subnetwork coupling when the union of all the reactions in all the possible subnetworks are added back to the model after the removal of the lumped reactions.

Figure 3-7 and 3-8 show D_1 and D_2 for all pairs of subnetworks, when no flux is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction, respectively. Figure 3-9 shows D_3 for all subnetworks, when no flux is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction. Table 3-7 shows full name of metabolites used in these figures.

In 36% of the subnetwork pairs $D_2(a, b) > D_1(a, b)$ and in 10% of the subnetwork pairs $D_1(a, b) > D_2(a, b)$. Although one may expect that D_2 should be bigger than D_1 since we are considering more couplings in D_2 , it is not always the case. It is mainly happening in the subnetwork pairs that the reactions that are common between them are less coupled to each other which means in general they have a lower degree of intra-coupling or D_3 . We believe that D_1 is a better metric for analyzing degree of coupling between two different subnetworks

since D_2 is also taking into account coupling between reaction pairs that both of them are

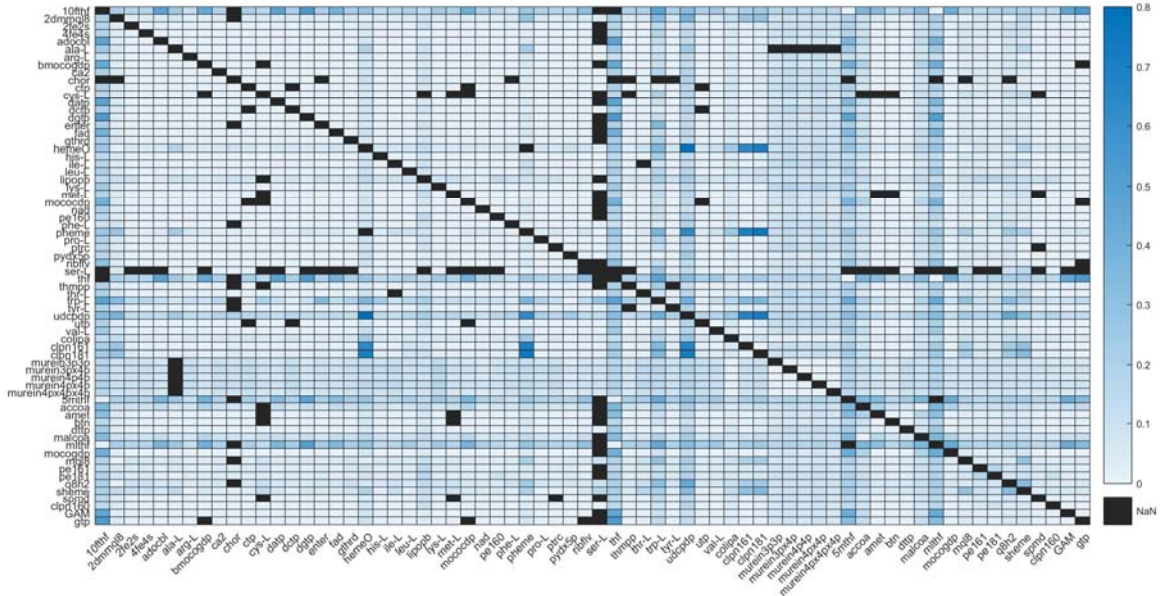


Figure 3-7. D_1 for all pairs of subnetworks, when no flux is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction.

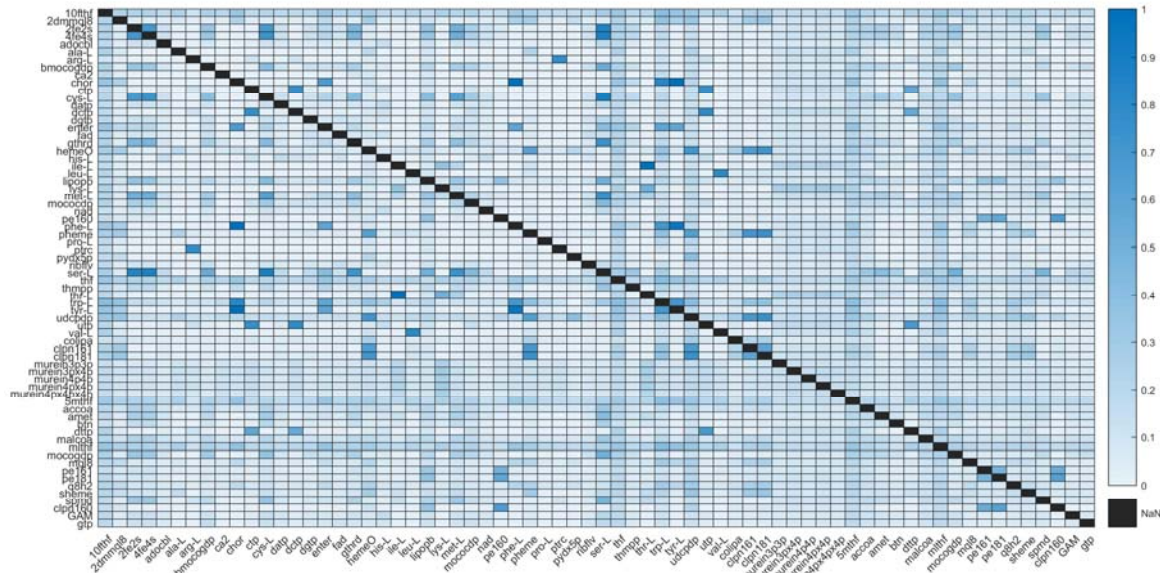


Figure 3-8. D_2 for all pairs of subnetworks, when no flux is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction.

inside one subnetwork and therefore it cannot truly show inter-coupling between subnetworks. Therefore, for the rest of our analysis we are considering D_1 and D_3 as measures of inter- and intra- subnetwork couplings respectively. As it can be seen in figure 3-7 we can have degrees of coupling from 0 to 0.8 between different pairs of subnetworks. One can see that seemingly unrelated parts of the metabolism such as subnetworks for biosynthesis of

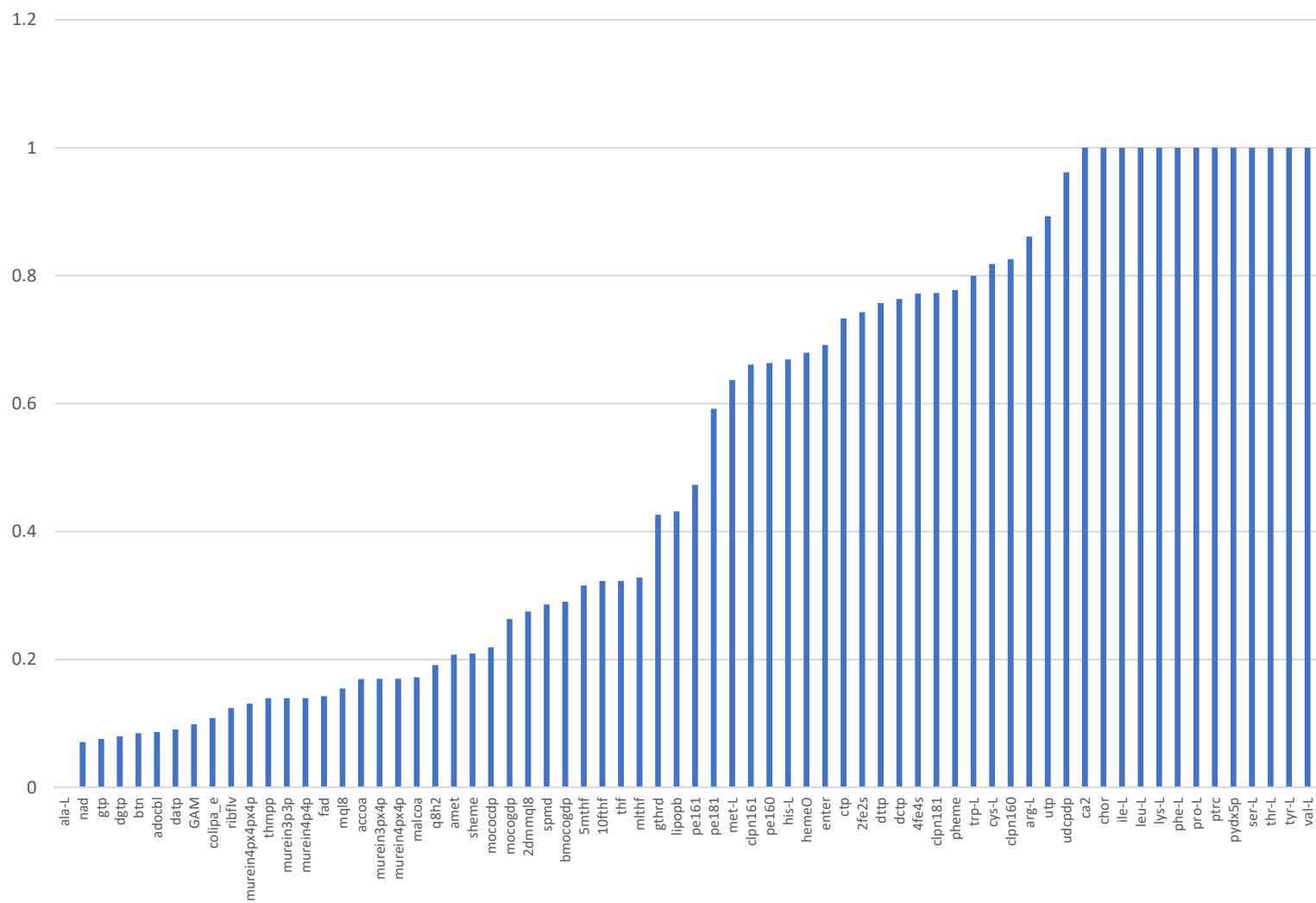


Figure 3-9. D_3 for all subnetworks, when no flux is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction

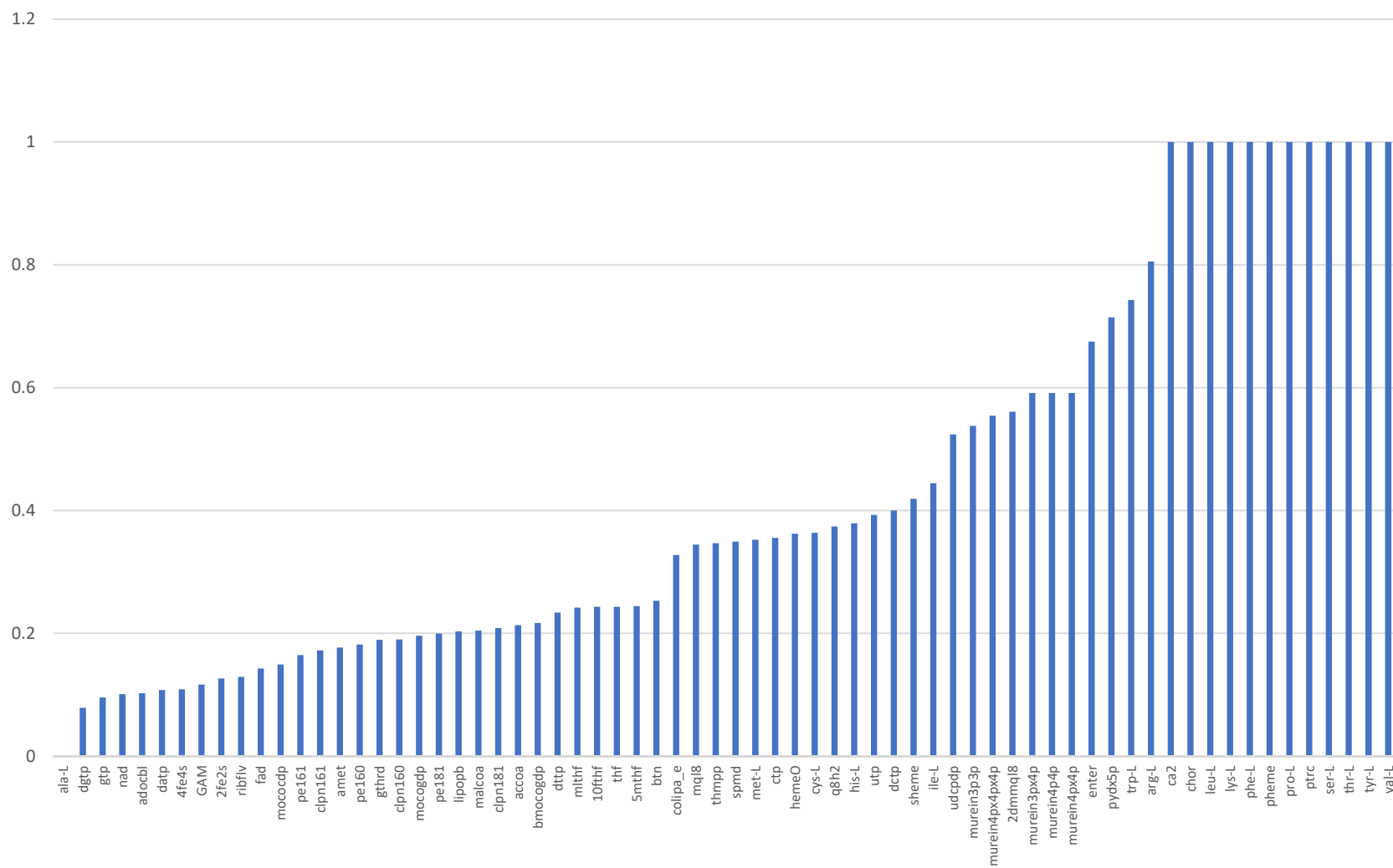


Figure 3-11. D3 for all subnetworks when no flux is maximized and we add the union of reactions of all possible minimal subnetworks

If we compare figures 3-10 and 3-11 to figures 3-7 and 3-9 respectively, we can see that when the network is more flexible, i.e., we add more reactions to the network, in general the degrees of coupling is less. For example, we can see that the maximum number for D_1 is decreased from 0.8 to 0.25 as it can be seen in figures 3-7 and 3-10, respectively. In addition,

Table 3-7. Full name of biomass building blocks in E. coli model

Metabolite abbreviated name	Metabolite full name	Metabolite abbreviated name	Metabolite full name
10fthf	10-Formyltetrahydrofolate	ser-L	L-Serine
2dmmql8	2-Demethylmenaquinol 8	thf	5,6,7,8-Tetrahydrofolate
2fe2s	[2Fe-2S] iron-sulfur cluster	thmpp	Thiamine diphosphate
4fe4s	[4Fe-4S] iron-sulfur cluster	thr-L	L-Threonine
adocbl	Adenosylcobalamin	trp-L	L-Tryptophan
ala-L	L-Alanine	tyr-L	L-Tyrosine
arg-L	L-Arginine	udcpdp	Undecaprenyl diphosphate
bmocogdp	bis-molybdopterin guanine dinucleotide	utp	Uridine-5'-triphosphate
Ca ²⁺	Calcium	val-L	L-Valine
chor	chorismate	colipa	core oligosaccharide lipid A
ctp	Cytidine triphosphate	clpn161	cardiolipin (tetrahexadec-9-enoyl, n-C16:1)
cys-L	L-Cysteine	clpn181	cardiolipin (tetraoctadec-11-enoyl, n-C18:1)
datp	Deoxyadenosine triphosphate	murein3p3p	two linked disaccharide tripeptide murein units (uncrosslinked, middle of chain)
dctp	Deoxycytidine triphosphate	murein3px4p	two disaccharide linked murein units, tripeptide crosslinked tetrapeptide (A2pm->D-ala) (middle of chain)
dgtp	Deoxyguanosine triphosphate	murein4p4p	two linked disaccharide tetrapeptide murein units (uncrosslinked, middle of chain)
enter	Enterochelin	murein4px4p	two disaccharide linked murein units, tetrapeptide corsslinked tetrapeptide (A2pm->D-ala) (middle of chain)
fad	Flavin adenine dinucleotide oxidized	murein4px4px4p	three disaccharide linked murein units (tetrapeptide crosslinked tetrapeptide (A2pm->D-ala) & tetrapeptide corsslinked

			tetrapeptide (A2pm->D-ala)) (middle of chain)
gthrd	Reduced glutathione	5mthf	5-Methyltetrahydrofolate
hemeO	Heme O	accoa	Acetyl-CoA
his-L	L-Histidine	amet	S-Adenosyl-L-methionine
ile-L	L-Isoleucine	btn	Biotin
leu-L	L-Leucine	dttp	Deoxythymidine triphosphate
lipopb	lipoate (protein bound)	malcoa	Malonyl-CoA
lys-L	L-Lysine	mlthf	5,10-Methylenetetrahydrofolate
met-L	L-Methionine	mocogdp	molybdopterin guanine dinucleotide
mococdp	molybdopterin cytosine dinucleotide	mql8	Menaquinol 8
nad	Nicotinamide adenine dinucleotide	pe161	phosphatidylethanolamine (dihexadec-9enoyl, n-C16:1)
pe160	phosphatidylethanolamine (dihexadecanoyl, n-C16:0)	pe181	phosphatidylethanolamine (dioctadec-11-enoyl, n-C18:1)
phe-L	L-Phenylalanine	q8h2	Ubiquinol-8
pheme	Protoheme	sheme	Siroheme
pro-L	L-Proline	spmd	Spermidine
ptrc	Putrescine	clpn160	cardiolipin (tetrahexadecanoyl, n- C16:0)
pydx5p	Pyridoxal 5-phosphate	GAM	D-Glucosamine
ribflv	Riboflavin	gtp	Guanosine triphosphate

the coupled or uncoupled subnetworks are different comparing figures 3-7 and 3-10. For example, lysine (subnetwork 24) and disaccharide linked peptide (subnetwork 47) are the most coupled subnetworks in figure 3-10. This shows that selection of subnetworks and lumped reactions have an impact on coupling relationship for the production of different biomass building blocks.

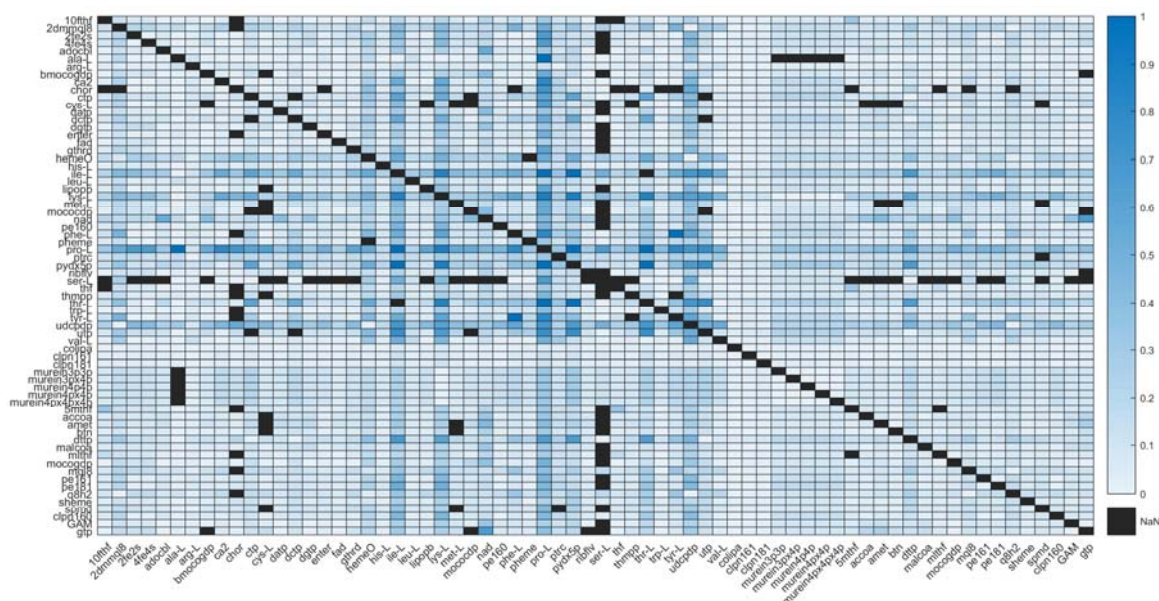


Figure 3-12. D_1 for all pairs of subnetworks, when NAD production is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction.

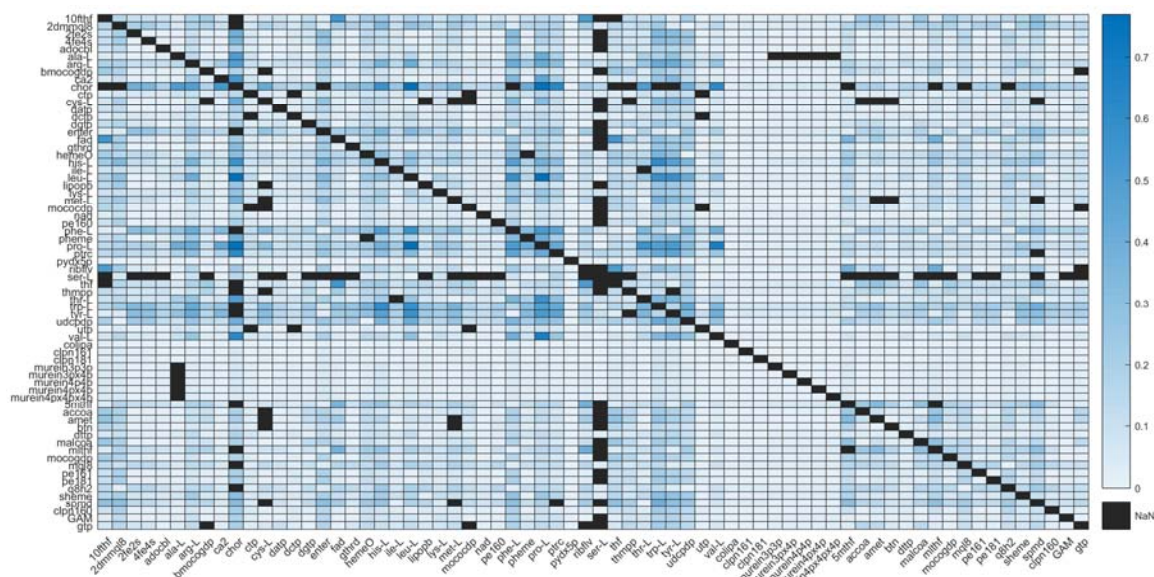


Figure 3-13. D_1 for all pairs of subnetworks, when acetyl-CoA production is maximized and we only add the reactions of one subnetwork corresponding to each lumped reaction.

When we maximize the production of biomass building blocks the network is more constrained and therefore the degrees of couplings are higher compared to when we do not maximize for any flux. This is shown for NAD and acetyl-CoA in figures 3-12 and 3-13.

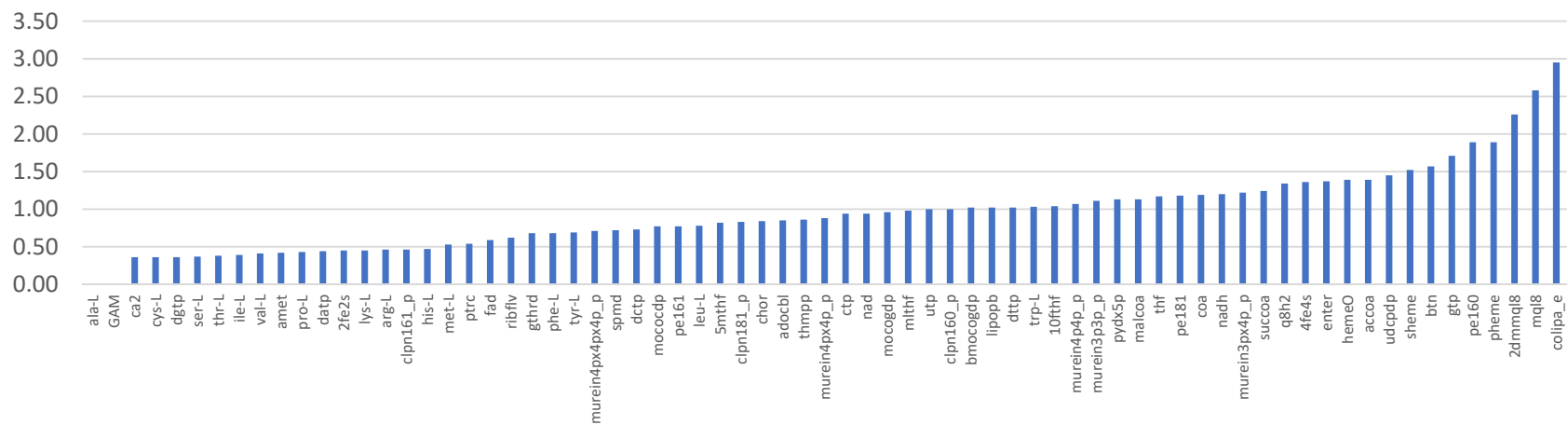


Figure 3-14. Ratio of number of fully coupled reaction pairs when maximizing for production of each of the building blocks to the number of fully coupled reaction pairs when no flux is maximized

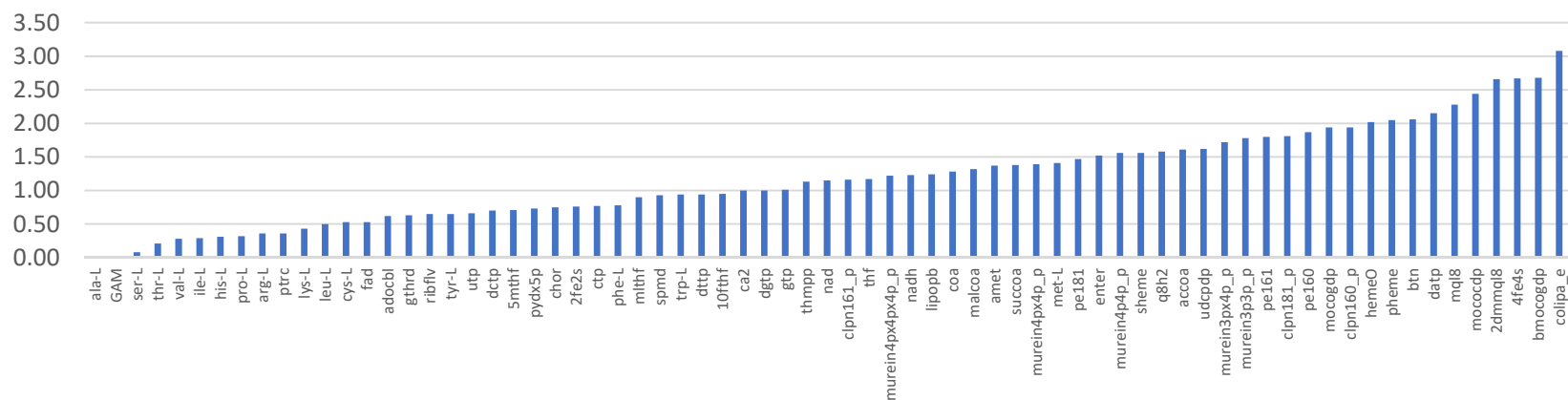


Figure 3-15. Ratio of number of directionally coupled reaction pairs when maximizing for production of each of the building blocks to the number of directionally coupled reaction pairs when no flux is maximized

Figure 3-14 and 3-15 show the ratio of number of fully and directionally coupled reaction pairs when maximizing for production of each of the building blocks to the number of fully and directionally coupled reaction pairs when no flux is maximized, respectively. As it can be seen, for both fully and directionally coupled reaction pairs in most of cases the ratio is more than one which shows the dependency between the fluxes are higher when we constrain the network. In case of fully coupled reaction pairs, for the ratios that are less than one at least one of the reaction pairs have become blocked therefore the coupling does not exist anymore. In case of directionally coupled reaction pairs, when the ratio is less than one, the reactions are either blocked or reaction pairs became fully coupled after the maximization of the building block.

3.5 Conclusion

In this chapter, we contribute to the analysis of metabolic networks by developing an optimization method for the analysis of coupling and dependencies in fluxes.

Analysis of flux couplings in metabolic network models is important to understand how different components of the network are interacting with each other and controlling the activity and availability of each other. In particular, in thermodynamically constrained metabolic network models, we can reduce the number of possible flux directionality profiles. Subnetworks coupling is a good systemic measure to see if the activity of different parts of the metabolism that may be seemingly unrelated are actually coupled to each other. As an ultimate goal, coupling analysis can be used for design of new phenotypes and engineering metabolism of organisms. This can be done by detecting parts of the network that removal of them can make a certain flux to be coupled with another one.

Furthermore, several known (novel) pathways may exist that can produce a metabolite of interest in a metabolic network. However not all of them have the same priority for being used in metabolic engineering purposes. Pathways are ranked based on different criteria such as pathway length, yield, thermodynamic feasibility and etc. Coupling analysis can introduce new ways of ranking pathways. For instance, those pathways that the production of the desired metabolites are coupled with biomass reaction may be of more interest for being used in experimental evaluations.

4 Cellular objectives of dormancy in malaria parasite

4.1 Author contribution and publications used in this chapter

The work of this chapter is based on the following paper which is under preparation for submission:

A. Chiappino-pepe, **Z. Hosseini**, E. Vayena, H. Frammery, V. Hatzimanikatis (2020) “Objective function, essentiality, and metabolic hallmarks of dormant malaria parasites” (in preparation).

The parts that are presented in this thesis was done in collaboration with Dr. Anush Chiappino-pepe and Hugo Frammery.

4.2 Introduction

Malaria is one of the most fatal infectious diseases which is caused by *Plasmodium* genus parasites. *P. Falciparum* and *P. vivax* are two species that play significant roles in malaria disease in humans. Parasite life cycle in human body consists of entering liver and infecting hepatocytes, replicating, rupturing hepatocytes, entering blood stream, infecting red blood cells, rupturing these cells and again entering blood stream. It should be noted that the parasite is not pathogenic outside red blood cells (133). Some species of malaria parasite such as *P. vivax* can form hypnozoites when they enter liver cells. Hypnozoites reduce their metabolic activity and energy consumption and remain dormant in hepatocytes. They can restart their activities after a period of time which can vary from weeks to months. After reactivation, hypnozoites can go through the same life cycle as the parasites that did not enter the dormancy phase (134). Although there are some hypotheses, it is still not known why parasites enter or leave the dormancy phase or what is the metabolic activity of parasite in dormancy (134–136).

Most malaria drugs are not effective on the dormant parasite. Hence, with the currently available drugs the dormant parasites should first be activated to leave the dormancy phase which is not very desirable. The alternative option is to have drugs that can target essential biological tasks in the dormant phase itself. Therefore, in depth understanding of metabolic activities of hypnozoites is essential to find new drug targets. Metabolic network analysis of malaria parasite gives a valuable opportunity to study the metabolism of this dangerous pathogen systematically. Thermodynamically curated genome-scale model of *Plasmodium falciparum* has been reconstructed in 2017 (9). For fast growing parasites maximizing biomass is a reasonable objective function that can be optimized in mathematical study of the model as it is done in (9). However, this is not the case for the dormant parasite since it is non-growing. Therefore, it is important to determine what are the potential cellular objectives of malaria parasite in the dormant phase in order to study its metabolism. In this chapter, we aim to analyse the dormant malaria parasite using an updated version of the genome-scale model of *P. falciparum*. Although this species of malaria parasite does not enter the dormant phase, it is the most comprehensive GEM that is available and there is significant similarity between this species and other hypnozoite forming species. Our results suggest that minimizing energy dissipation is the most consistent objective function with the experimental gene expression data of *P. vivax*.

4.3 Materials and methods

4.3.1 *P. falciparum* genome-scale model used for the study

iPfa2 consists of 429 genes and 1317 reactions. It includes 261 extracellular and 1170 intracellular metabolites. The model has five compartments: the cytosol, the mitochondrion, the apicoplast, the endoplasmic reticulum and the nucleus. The iPfa2 is also thermodynamically curated for 69% of the reactions.

4.3.2 Thermodynamics-based flux analysis (TFA)

TFA is a MILP formulation that integrates metabolite concentrations into the flux-balanced model in order to impose second law of thermodynamics on model reactions. In addition to

the mass balance constraint of FBA problem, the following constraints describe TFA formulation:

$$0 \leq v_i \leq z_i v_{max} \quad \{i = 1, \dots, r\}$$

Where, v is the vector representing the flux through reactions of the model and z is the vector of binary variables associated to each reaction. If z_i is equal to 1 then reaction i can carry flux, otherwise the flux through i th reaction should be equal zero. Every reaction in TFA formulation is split to forward and backward reactions with positive fluxes. r here, is the total number of forward and backward reactions.

$$z_{i,forward} + z_{i,backward} \leq 1$$

The above constraint that applies to all pairs of forward and backward reactions ensure that only one of the associated binary variables can be equal to 1. In other words, in each flux distribution, one reaction can only have flux in one direction.

$$\Delta_r G'_i - K + K z_i < 0$$

For all the reactions that have known $\Delta_r G'$, the above constrained insures that z_i can only be 1 if corresponding $\Delta_r G'_i$ is negative. K here, is a sufficiently large number.

$$\Delta_r G_i'^{\circ} + RT \sum_{j=1}^m s_{i,j} \ln(x_j) = \Delta_r G'_i$$

Finally, the above constraint applies the second law of thermodynamics on reactions of the model. $\Delta_r G_i'^{\circ}$ is the standard Gibbs free energy change of reaction i in the model. R and T are ideal gas constant and temperature in Kelvin, respectively. $s_{i,j}$ is the stoichiometric constraint of metabolite j in reaction i and $\ln(x_j)$ is the natural logarithm of the activity of metabolite j .

4.3.3 Flux Variability Analysis (FVA)

FVA is a technique used to determine the range of possible values, each reaction can take under the constraints imposed on the model. First, the objective function is optimized. Then, in a new constraint, a certain percentage of the optimal value of the objective is set as the lower or upper bound of the objective function. Finally, the objective function is changed to be the minimization or maximization of each flux. As a result, the allowable range for each reaction is computed.

4.3.4 Objective functions

A set of six objective functions were chosen for this study. Their formulation is described below. These objective functions were analyzed with the thermodynamics constraints included in the model.

4.3.4.1 Minimization of the sum of all fluxes

The objective is to minimize the sum of the absolute value of all fluxes

$$\text{Minimize } \sum_{i \in Rxn} |v_i|$$

where Rxn is the set of all reactions in the model, and v_i is the net flux through reaction i . Since in TFA formulation, fluxes are split between forward and backward directions and all flux values are positive, the above objective function can be simplified as follows:

$$\text{Minimize } \sum_{i \in Rxn} v_i$$

4.3.4.2 Minimization of the number of active reactions

The objective is to find the minimal network of reactions that makes the model feasible.

$$\text{Minimize } \sum_{i \in Rxn} z_i$$

where Rxn is the set of all reactions in the model, and z_i is a binary variable equal to 0 if there is no flux through reaction i . This is ensured by the following constraint from TFA formulation:

$$0 \leq v_i \leq z_i v_{max}$$

4.3.4.3 Minimization of the sum of all fluxes through the uptake reactions

The objective is to minimize the sum of the absolute value of the fluxes through all the uptake reactions.

$$\text{Minimize } \sum_{i \in Upt} |v_i|$$

where Upt is the set of all uptake reactions in the model, and v_i is the net flux through reaction i .

This is implemented in a similar way to the problem of minimizing the sum of all fluxes. Here, only the backward direction of the exchange fluxes is taken into account so that we only minimize the flux through uptake reactions.

4.3.4.4 Minimization of the number of active uptakes

In this problem, the objective is to find the minimal set of uptakes that have to be active to render the model feasible.

$$\text{Minimize } \sum_{i \in Upt} z_i$$

where Upt is the set of all uptake reactions and z_i is the binary variables associated to backward direction of each uptake flux in the model.

4.3.4.5 Minimization of energy dissipation

The objective is to minimize the energy dissipated by the system, that is the sum over the product of the Gibbs free energy change of all the internal reactions in the model and the flux through the reactions.

$$\text{Minimize } \sum_{i \in IntRxn} v_i \Delta_r G'_i$$

where $IntRxn$ is the set of all internal reactions in the model, excluding exchange reactions and transports between the external compartment and the cytosol. v_i is the net flux through reaction i , and $\Delta_r G'_i$ is the Gibbs free energy change of reaction i .

Since both v_i and $\Delta_r G'_i$ are variables, this is a Mixed Integer Non Linear Programming (MINLP) problem. However, it can be shown that for all reactions that are not boundary reactions, $\Delta_r G'_i$ can be approximated by $\Delta_r G_i^{\circ}$, the standard Gibbs free energy change of reaction, which is a constant. The usage of this approximation allows to transform the non-linear problem into a linear problem that can be solved with the regular MILP solver used for the other objective functions.

4.3.4.6 Maximization of ATP yield per flux unit

The objective is to find maximum ATP that can be produced per some of all fluxes:

$$\text{Maximize } \frac{v_{ATP}}{\sum_{i \in Rxn} v_i}$$

v_{ATP} is the flux of unused ATP produced, Rxn is the set of all reactions in the model, and v_i is the flux through forward or backward direction of reaction i .

To implement this problem, an exchange reaction is added to the model to be able to simulate an ATP secretion. Maximizing the production of unused ATP can consequently be done by maximizing the flux through this new exchange reaction.

This problem is inherently an MILFP but we will apply a linearization process to transform it into an equivalent MILP (128). This transformation relies on the fact that an equivalent optimization problem is:

$$\text{Maximize } \frac{v_{ATP} \cdot t}{\sum_{i \in Rxn} v_i \cdot t}$$

where t is a continuous positive variable such that

$$\sum_{i \in Rxn} v_i \cdot t = 1$$

If the above equation holds, then the denominator of objective function can be simplified and the objective function of the problem becomes

$$\text{Maximize } v_{ATP} \cdot t$$

This can be transformed into a linear problem by performing the variable substitution

$$\hat{v}_i = v_i \cdot t$$

which results in the following general MILP problem:

$$\text{Maximize } \hat{v}_{ATP}$$

Subject to

$$\sum_{i \in Rxn} \hat{v}_i = 1$$

$$S \cdot \hat{v} = 0$$

$$LB \cdot t \leq \hat{v} \leq UB \cdot t$$

$$t \geq 0$$

where S is the stoichiometric matrix, \hat{v} is the vector containing all the substituting flux variables, and LB and UB are the vectors holding respectively the lower and upper bounds of the original flux variables.

Additionally, the vector y holding all the binary variables contained in the model also needs to be converted into a new vector w by introducing a set of additional constraints as follows (128):

$$w = y \cdot t$$

$$w \leq t$$

$$\begin{aligned}
w &\leq M \cdot y \\
w &\geq t - M(1 - y) \\
w &\geq 0
\end{aligned}$$

where M is a sufficiently large number representing the upper bound of t . The details about these additional constraints are explained in chapter two and three.

For each objective function, all the alternative solutions that can achieve the optimal value of the objective function are identified. This is done by first performing an optimization of the objective, identifying the optimal value and setting it as a constraint. Then the set of active reactions of the solution achieving this optimum are identified and an integer cut constraint is added to the model to prevent the solver from finding the same exact set of active and inactive reactions again. This procedure is repeated till all the alternative optimal solutions are identified.

4.3.5 Integration of transcriptomics data.

Recent experimental analyses on dormant parasites have been performed to determine gene expression levels for dormant *P. vivax* (137). These data were used for our consistency assessment which is explained later in this chapter. Since the experimental measurements were performed on *P. vivax*, and the iPfa2 GEM is for *P. falciparum*, the equivalence between the genes of the two species were identified using a blasting technique, available in the RAVEN toolbox (138).

The genes that have equivalence in the iPfa2 model are sorted in ascending order of expression level. Then, by manually setting two thresholds, the genes are classified in three categories: lowly expressed genes, normally expressed genes and highly expressed genes. The Gene-Protein-Reaction (GPR) rules are used to determine what are the corresponding upregulated and down-regulated reactions. A gene highly expressed is likely to produce more of the enzyme it is coding for which results in a larger flux in the corresponding reaction.

In the next step, constraints are added to the model that can be switched on or off to force the reaction fluxes to be in their high flux regime, for upregulated reactions, or in their low flux regime, for downregulated reactions.

4.4 Results and discussion

4.4.1 Integration of objective functions and hallmarks of dormancy:

The aim of this study is to understand the metabolic activities of hypnozoites through analyzing different objective functions for dormancy phase. The list of objective functions used in this study are as follows and discussed in details in the materials and methods section:

- i) Minimization of the sum of all fluxes;
- ii) Minimization of the number of active reactions;
- iii) Minimization of the sum of all fluxes through the uptake reactions;
- iv) Minimization of the number of active uptakes;
- v) Minimization of energy dissipation;
- vi) Maximization of ATP yield per flux unit.

Although metabolic behavior of hypnozoites are poorly characterized, there are some postulations for this phase known as hallmarks of dormancy (139,140). We used some of these hallmarks as constraints imposed on the model. In particular, we considered seven conditions for the model by applying appropriate constraints:

- i) Forced glucose uptake; by applying a small flux through the reaction uptaking glucose from environment.
- ii) Forced lactate production; by applying a small flux through the reaction secreting lactate to the environment.
- iii) Forced ATP synthesis; by applying a small flux through the ATP synthesis reaction in electron transfer chain (ETC).
- iv) Forced ATP maintenance; by applying a small flux through the ATP phosphohydrolysis reaction in cytoplasm.
- v) Forced biomass production (this is not a hallmark of dormancy); by applying a small flux through biomass reaction.
- vi) Combination of five conditions above;
- vii) Combination of five conditions above except biomass production;

Since in dormancy no growth is observed, in all of the above conditions except conditions v and vi, a small positive upper bound is considered for the biomass condition to simulate for the absence of growth.

4.4.2 Computation of alternative flux distributions

All the conditions presented above are combined separately with all the objective functions described before. For each combination, all the alternative solutions were computed, which are different sets of active and inactive reactions that can achieve the same optimal value for the objective function. Table 4-1 shows the number of alternative optimal solutions identified for each combination of condition and objective function.

Table 4-1. number of optimal alternative solutions per combinations of condition and objective function

	Glucose uptake	Lactate secretion	ATP synthesis	ATP maintenance	growth	All conditions	All but growth
Min. sum of all fluxes	296	20	274	429	853	827	209
Min. number of active fluxes	1	2	2	28	5	1	1
Min. sum of all fluxes through uptakes	1	15	5	4	227	512	11
Min. number of uptakes	1	9	7	18	7	12	6
Min. dissipation	8	7	1	7	37	3	13
Max. ATP yield per flux unit	253	151	161	201	339	318	125

4.4.3 Constitutive reactions

Constitutive reactions are the reactions that are active in all alternatives. Table 4-2 shows the number of constitutive reactions for each combination of objective function and condition. As it can be seen, maximizing growth has the highest number of constitutive reactions. It is due to the fact that maximizing growth needs the production of a certain amount of all biomass building blocks which makes the system less flexible.

Table 4-2. Number of constitutive reactions per combination of condition and objective function

	Glucose uptake	Lactate secretion	ATP synthesis	ATP maintenance	growth	All conditions	All but growth
Min. sum of all fluxes	8	3	13	16	506	504	69
Min. number of active fluxes	12	8	24	1	304	579	149
Min. sum of all fluxes through uptakes	15	10	73	23	402	398	42
Min. number of uptakes	3	4	39	7	477	578	51
Min. dissipation	288	336	357	284	604	705	282
Max. ATP yield per flux unit	31	26	31	19	392	396	60

4.4.4 Production of biomass building blocks

In this section, we identified biomass building blocks (BBBs) that can be produced in each combination of objective function and condition. Table 4-3 shows the summary of this result. It should be noted that the numbers show how many BBBs can be produced in at least one alternative optimal solution. Therefore, these number of BBBs are not necessarily produced at the same time, in one solution. In addition to conditions where forced biomass production is imposed and hence all BBBs are produced, we have three more combinations in which all 86 BBBs are produced: maximization of ATP yield per flux unit with forced glucose uptake or lactate secretion and minimization of dissipation having all hallmarks of dormancy. We further analyzed these three conditions and realized that only for the latter case all BBBs can be produced at the same time in one alternative solutions and therefore, biomass can also be produced.

Table 4-3. Number of Biomass Building Blocks that can be synthesized in at least on alternative solution per combination of condition and objective function

	Glucose uptake	Lactate secretion	ATP synthesis	ATP maintenance	growth	All conditions	All but growth
Min. sum of all fluxes	47	27	48	52	86	86	50
Min. number of active fluxes	3	16	42	85	86	86	20
Min. sum of all fluxes through uptakes	6	46	45	34	86	86	51
Min. number of uptakes	21	42	45	46	86	86	47
Min. dissipation	85	85	37	85	86	86	86
Max. ATP yield per flux unit	86	86	21	31	86	86	58

4.4.5 Scoring combinations of objective functions and conditions

As explained in materials and methods section, we added constraints for up-regulated and down-regulated reaction. After adding these constraints, the system is forced to work at its optimal value with respect to the objective function, and the constraint carried by the condition studied is also enforced. Then, the number of the gene expression-based constraints that are fulfilled is maximized as the new objective of the problem. This number of fulfilled constraints represents the consistency between the experimental data and the combination applied on the model. It is called the Consistency Score (CS) and it serves as a metric to score and rank the objective functions and conditions. Table 4-4 summarizes CS for all combinations of condition and objective function. It should be noted that maximum theoretical CS that can be achieved is 89.

Table 4-4. Consistency score per combination of objective functions and conditions with integrated transcriptomics data

	Glucose uptake	Lactate secretion	ATP synthesis	ATP maintenance	growth	All conditions	All but growth
Min. sum of all fluxes	69	69	69	69	55	55	69
Min. number of active fluxes	11	18	14	17	1	1	12
Min. sum of all fluxes through uptakes	75	75	75	75	61	61	75
Min. number of uptakes	83	83	76	82	69	69	79
Min. dissipation	83	83	83	83	69	69	82
Max. ATP yield per flux unit	69	69	69	69	55	55	69

As it can be seen in the table, for all of the objective functions that we investigate, as soon as having growth is forced to the model, the CS is dropped. It can be explained by the fact that growth is not a hallmark of dormancy. Therefore, forcing the model to activate flux for all the reactions needed for growth is not consistent with experimental gene expression data of dormancy. We further examined what are the inconsistencies that are only observed when biomass production is forced to the model. We found eight reactions that their corresponding genes are down-regulated based on the gene expression data. However, enforcement of biomass production makes these reactions to have higher fluxes and therefore the corresponding down-regulating constraint is not fulfilled. They are indicated in table 4-5. Interestingly, all these reactions are related to protein biosynthesis. This finding shows that dormant parasite is reducing the activity of protein synthesis, and it is not something that is in agreement with having growth. Some of the current drugs for non-dormant malaria parasite are targeting these protein biosynthesis genes (141,142). However, in the dormant phase, parasite itself aims to downregulate these genes and therefore such drugs are not effective for killing parasite in the dormancy.

Table 4-5. Reactions that cannot be down-regulated in the forced growth condition and Reactions that can only be up-regulated with dissipation minimization as objective function

Reactions that cannot be down-regulated in the forced growth condition	
Reaction	Subsystem
$\text{H}_2\text{O} + \text{dATP} + \text{thioredoxin disulfide} \rightarrow \text{ADP} + \text{thioredoxin}$	Purine metabolism
$\text{ATP} + \text{L-tryptophan} + \text{H}^+ + \text{tRNA(Trp)} \rightarrow \text{diphosphate} + \text{AMP} + \text{L-tryptophanyl-tRNA(Trp)}$	Aminoacyl-tRNA biosynthesis
$\text{ADP} + \text{orthophosphate} + \text{L-glutamate} + \text{glutaminyI-tRNA} \rightarrow \text{H}_2\text{O} + \text{ATP} + \text{L-glutamine} + \text{L-glutamyl-tRNA(Gln)}$	Aminoacyl-tRNA biosynthesis
$\text{ADP} + \text{orthophosphate} + \text{L-glutamate} + \text{L-asparaginyI-tRNA(Asn)} \rightarrow \text{H}_2\text{O} + \text{ATP} + \text{L-glutamine} + \text{L-aspartyl-tRNA(Asn)}$	Aminoacyl-tRNA biosynthesis
$\text{ATP} + \text{H}^+ + \text{L-asparagine} + \text{tRNA(Asn)} \rightarrow \text{diphosphate} + \text{AMP} + \text{L-AsparaginyI-tRNA(Asn)}$	Aminoacyl-tRNA biosynthesis
$\text{ATP} + \text{L-glutamate} + \text{tRNA(Gln)} \rightarrow \text{diphosphate} + \text{AMP} + \text{L-glutamyl-tRNA(Gln)}$	Aminoacyl-tRNA biosynthesis
$\text{ATP} + \text{L-tryptophan} + \text{H}^+ + \text{tRNA(Trp)} \rightarrow \text{diphosphate} + \text{AMP} + \text{L-tryptophanyl-tRNA(Trp)}$	Aminoacyl-tRNA biosynthesis
$\text{ATP} + \text{L-glutamate} + \text{H}^+ + \text{tRNA(Glu)} \rightarrow \text{diphosphate} + \text{AMP} + \text{L-glutamyl-tRNA(Glu)}$	Aminoacyl-tRNA biosynthesis
Reactions that can only be up-regulated with energy dissipation minimization as objective function	
Reaction	Subsystem
$\text{H}_2\text{O} + \text{ATP} + \text{L-selenomethionine} \rightarrow \text{orthophosphate} + \text{diphosphate} + \text{Se-adenosylselenomethionine}$	Purine metabolism
$\text{H}_2\text{O} + \text{Se-adenosyl-L-selenohomocysteine} \rightarrow \text{adenosine} + \text{selenohomocysteine}$	Purine metabolism
$\text{NAD}^+ + (\text{S})\text{-lactate} \rightarrow \text{NADH} + \text{pyruvate} + \text{H}^+$	Pyruvate metabolism
$\text{NAD}^+ + 3\text{-mercaptolactate} \rightarrow \text{NADH} + \text{H}^+ + \text{mercaptopyruvate}$	Cysteine and methionine metabolism
$3\text{-hydroxyoctadecanoyl-CoA} \rightarrow \text{H}_2\text{O} + (2\text{E})\text{-octadecanoyl-CoA}$	Fatty acid metabolism
$\text{NAD}^+ + \text{orthophosphate} + \text{D-glyceraldehyde 3-phosphate} \rightarrow \text{NADH} + 3\text{-phospho-D-glycerol phosphate}$	Glycolysis
$\text{phenylpyruvate} \rightarrow 2\text{-hydroxy-3-phenylpropenoate}$	Phenylalanine metabolism
$3\text{-(4-hydroxyphenyl)pyruvate} \rightarrow 2\text{-hydroxy-3-(4-hydroxyphenyl)propenoate}$	Tyrosine metabolism

Another important observation from table 4-4 is that in all conditions, the maximum CS is achieved with minimizing dissipation as the objective function. This suggests that among the objective functions that we investigated, minimizing dissipation is the most consistent cellular objective with experimental gene expression data. We further investigated what are the constraints that can only be fulfilled having this objective function. Eight reactions that should be upregulated based on the data, can only be up-regulated with minimizing dissipation as the objective function. This is the case for all seven different conditions. These reactions are also available in table 4-5. This suggests that not all biological pathways have equally reduced their metabolic activities upon dormancy. For the dormant parasite to stay viable, some parts of its metabolism need to be upregulated. Our results suggest that this increase in metabolic activity can only be captured if the cellular objective is minimizing energy dissipation.

4.5 Conclusion

The aim of this chapter was to characterize metabolic behavior of dormant malaria parasite in the liver stage using metabolic network modeling. For this purpose, we used iPfa2, genome-scale metabolic network model of *P. falciparum* and gene expression data of *P. vivax* in the dormant phase. In order to simulate dormancy, we considered seven different metabolic conditions and six different objective functions. We did all of our analysis for the forty-two combinations of objective functions and conditions. Our computational platform for the study of dormant parasite helped us understand the metabolic activity of hypnozoites. Specifically, we investigated the flexibility of the networks, constitutive reactions, and ability of model to produce biomass building blocks in each combination. With integrating gene expression data of dormant parasite into the model we could calculate consistency score for each combination of dormancy conditions and objective functions. Our results suggest that minimizing energy dissipation throughout the whole metabolic network can best describe the metabolic behavior of dormant parasites. It should be noted that minimizing dissipation is introduced in this work for the first time to be the objective function of a metabolic modeling study. It emphasizes again the importance of taking into account thermodynamic properties of reactions and metabolites in analysis of metabolic network models. Minimization of energy dissipation is the best objective function in terms of capturing the upregulation of reactions in different pathways. It suggests that although this complex objective function is rewiring

the fluxes so that the energy dissipation is minimized, but it can realize that some parts of the metabolism still need to be upregulated. In addition, some of the reactions in protein synthesis pathways are down-regulated in hypnozoites. This can be of interest in guiding drug targeting for this disease, since some of the currently used drugs are knocking down the enzymes associated to these reactions (141). The fact that in the dormant phase these reactions are down-regulated, shows that these drugs may trigger the parasite to go through the dormant phase. Therefore, an effective treatment can be a drug that target both dormant and non-dormant phases of malaria parasite at the same time.

5 Study of iron reduction in *Clostridium acetobutylicum*

5.1 Author contribution and publications used in this chapter

The project presented in this chapter is the result of our collaboration with Prof. Rizlan Bernier-Latmani and Dr. Cornelia List from Environmental Microbiology Laboratory at EPFL. The experimental measurements were done by our collaborators and the computational studies were done by Zhaleh Hosseini. Some of the results are published in the following paper:

C. List, Z. Hosseini, K. Lederballe Meibom, V. Hatzimanikatis, R. Bernier-Latmani (2019) “Impact of iron reduction on the metabolism of *Clostridium acetobutylicum*”, Environmental Microbiology, 21(10):3548–63.

5.2 Introduction

Clostridium acetobutylicum is an anaerobic gram-positive bacterium that can convert various sugars to acetate, butyrate and solvents such as acetone, butanol, and ethanol (143–145). *C. acetobutylicum* has been studied as the model organism for the production of solvents on an industrial scale through the acetone-butanol-ethanol (ABE) fermentation (63,146,147).

The life cycle of *C. acetobutylicum* consists of an initial growth phase in which the organism produces butyrate and acetate (the acidogenic phase). Production of these organic acids cause the pH to decrease. Then, the metabolism shifts to the production of acetone, ethanol and butanol (the solventogenic phase). Additionally, the bacterium produces H₂ and CO₂ (145,148). Studies showed that providing additional electron acceptors such as Fe(III) for fermentative bacteria, helps in generating energy and pH stabilization. In addition, the impact of iron reduction on growth yield, glucose consumption and fermentation has been discussed

(149,150). Most microorganisms need iron as an essential element for their growth and survival. Iron serves as a main cofactor for the activity of some enzymes and can be used directly as an electron donor or acceptor. Iron is cycled between its ferric (Fe(III)) and ferrous (Fe(II)) states and bacteria take up these forms differently. The reduced form of iron is more soluble compared to ferric iron. Thus, it is easier to be acquired from the environment. However, some bacteria can use the oxidized form of iron as an electron acceptor, producing Fe(II). Iron reduction can be used not only for direct energy gain from electron transport to Fe(III), but also for enhanced fermentation (148,151–153). *C. acetobutylicum* can be used as a model to study the process of iron reduction and extracellular electron transfer in a gram-positive organism and to uncover the mechanism of extracellular electron transfer and influence of iron reduction on its metabolism. Our collaborators in the laboratory of environmental microbiology at EPFL showed that *C. acetobutylicum* can reduce soluble and solid Fe(III). Iron reduction affect pH balancing and carbon and electron flow. They showed that the amount of hydrogen, butanol, organic acids and ATP productions are changed if iron is provided (148).

C. acetobutylicum ATCC 824 is the first sequenced *Clostridium* and it can serve as a model organism for clostridial metabolism in general (135,154). There are a couple of reconstructed genome scale models of *C. acetobutylicum* that will help in genetic, biotechnological and physiological researches of different clostridia with annotated genomes (10,63,143,154,155). In this chapter, we study the metabolism of *Clostridium acetobutylicum* by using the genome-scale model for this organism which is reconstructed in 2014 by Dash et al (10). We are using the data provided by our collaborators in Environmental Microbiology Laboratory at EPFL to study mechanisms and reactions involved in iron transfer and reduction process. For this purpose, we performed the thermodynamic curation of the model to allow the integration of metabolomics data to the genome scale model of this organism (5). Thermodynamics-based metabolic flux analysis (5), flux variability analysis (FVA) (156), and flux sampling (157) was used to investigate *C. acetobutylicum* metabolism. Metabolic modelling showed an increase in some of the fluxes of central pathways if iron is provided. In addition, computational analysis suggests that more alternative reactions are possible with iron reduction compared to the condition with no iron. In addition, transcriptomics data was integrated to the model to analyse the enriched metabolic pathways with iron reduction using Minimal Network Enrichment Analysis (MiNEA) (19). MiNEA finds the minimal reaction sets need to fulfil

different metabolic tasks. Then, having transcriptomics data MiNEA shows what minimal subnetworks are enriched with over or under expressed genes. We showed that Minimal subnetworks for some of the amino acids, RNA and cross-linked peptidoglycans biosynthesis are enriched after iron is provided for reduction by the cell. Our analysis suggests that NADH and NADPH are more probable to have role in iron reduction in *Clostridium acetobutylicum*. In addition, ATP and FAD minimal subnetworks are significantly deregulated which is in agreement with experimental data and shows that iron reduction affect energy generation in this model organism.

5.3 Materials and methods

5.3.1 Experimental data

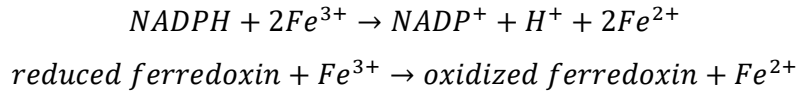
Time-point measurements of concentrations of some metabolites used in constraining metabolic network model. Experiments were done providing 4% glucose and 40 mM soluble Fe(III) for the cells. Glucose and Fe(III) consumptions and acetate, lactate, butyrate, ethanol, acetone, butanol, Fe(II), pyruvate, H₂ and CO₂ productions were measured at 0, 2, 4, 6, 8, 12, 18 and 24 hours after cultivation of the microbes. These concentrations were used to fit polynomial equations to calculate specific uptake and secretion rates in exponential growth phase. OD600 data was used to calculate specific growth rate for different time points.

For RNA sequencing analysis, the initial concentrations were 0.5% glucose and 2.5 mM solid Fe(III) and 4% glucose and 40 mM soluble Fe(III). Time series concentration of glucose consumption, Fe(II) production and OD600 data were available were used for polynomial fitting. Analysis of transcriptomics data provided us with upregulated and down regulated genes when the iron is available (in solid or soluble form) compared to conditions where no iron is provided for the cells.

5.3.2 TFA, FVA, sampling and calculation of minimal reaction subnetworks

iCac802 was used as the genome-scale model in this study (10). The original model contains 1457 reactions and 1247 metabolites. To be able to model iron reduction, the three following reactions were added to the model one at a time:



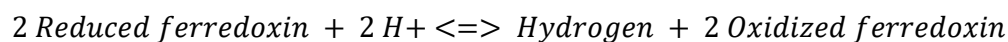


The uptake and secretion flux rates calculated in previous section were used to constrain the corresponding uptake and secretion rates in the model and to compare the results of model simulation with experimental measurements. Metabolic network models are usually studied using flux balance analysis (FBA) (158). In FBA, an objective function in the model is optimized subject to stoichiometric and capacity constraints. Here, thermodynamic-based flux balance analysis (TFA) was used for studying iron reduction in iCac802. In TFA, metabolite concentrations and Gibbs free energy of reactions are also taken into account (5). To check the consistency between model predictions and experimental data, all above-mentioned constraints for metabolite secretions and uptakes, (except the growth rates) were put as lower and upper bounds of the corresponding fluxes in the metabolic network. The experimental standard deviations were used to calculate the lower and upper bounds put into the model. Then, the model was optimized by maximizing the biomass production and it was compared with the experimental data of growth rates. After maximizing biomass production, the lower bound for biomass reaction was set to 90% of its maximum and then Flux variability analysis was performed to identify reactions that shifted fluxes when the appropriate iron reduction reaction is added to the model. Flux variability analysis calculates maximum and minimum possible flux for all reactions of the network by having suboptimal biomass and other constraints of interest. Sampling of the solution space was done to see the change in the mean value of the reactions involved in glycolysis, citric acid cycle, fermentation and pentose phosphate pathway. For doing so, the minimum amount of acetate needed to be produced from central fluxes to fulfil experimental constraints was calculated and was set as a new constrain to the model. In addition, the reactions from acetoacetyl-CoA to acetoacetate and butyryl-CoA to butyrate were forced to only carry fluxes in the forward direction. After sampling 5000 points, the mean value of the reactions in the iron-reducing condition was divided by the mean value of the reactions with no iron condition. Moreover, to compare the experimental and theoretical flux ranges for acetate, butyrate, butanol and H₂, the bounds on these fluxes were relaxed and the maximum and minimum possible fluxes were calculated for each. For this calculation, all other experimental constraints were imposed and lower bound for biomass reaction was set to 90 percent of its maximum. To calculate ATP production, the predetermined bounds on ATP maintenance reaction were

relaxed and maximum flux for this reaction was calculated for each of the iron-reducing reactions, while maintaining the experimental constraints. The results were compared with the no Fe(III) case. Finally, all alternative minimal reaction subnetworks were constructed by setting all experimental data as constraints. Metabolic graphs were drawn using yEd graph editor. It should be noted that in all the above computations the experimental bounds on pyruvate secretion makes the model infeasible having ferredoxin as electron donor for iron reduction. Therefore, the bound on this flux is relaxed only with ferredoxin as cofactor.

5.3.3 Minimal Network Enrichment Analysis

iCac802 had originally 803 genes. We added CA_C0028, gene corresponding to hydrogenase enzyme, to the model and linked it to the following reaction:



Minimal Network Enrichment Analysis (MiNEA) was done to understand what parts of metabolism are involved in iron reduction and if they are enriched in deregulated genes (19). We performed MiNEA for biosynthesis of all biomass building blocks and for NADH, reduced ferredoxin, H₂, Lactate, butanol and Fe(II). We studied the synthesis networks of these metabolites using MiNEA, where we called the synthesis of a metabolite as a metabolic task. The inputs of MiNEA are metabolic network model of the organism, desired metabolic tasks and experimental gene expression data. MiNEA enumerates all the thermodynamically feasible minimal-size networks that are active for the fulfillment of the desired metabolic task.

A mixed-integer linear programming (MILP) problem is solved in MiNEA to optimize the objective of maximization of the number of reactions that cannot carry flux, while maintaining a certain threshold for the production of the desired task. We used 95% of the maximum possible production of the metabolic task as the threshold. All the alternative minimum-size networks for the synthesis of metabolic tasks were calculated. In the next step of MiNEA, the deregulated genes that are different between the two conditions (here solid and soluble Fe(III) vs. no Fe(III)) are determined. A reaction can have three states based on its associated genes: up-regulated, down-regulated or un-regulated. A reaction is considered as up (down) regulated if its associated genes are only up (down) regulated. If a reaction is associated to

both up and down regulated genes, it is considered as un-regulated. In the last step, the significantly deregulated minimal networks are identified using a hypergeometric test on deregulated reactions. For instance, to see if a minimal network is significantly upregulated, the probability of having a minimal network with the same size, with equal or more upregulated reactions and with as few as possible down-regulated reactions is calculated. This probability shows the p-value for the significance of upregulation of the corresponding minimal network. More details about the MILP formulation and the way p-value is calculated for the significance of deregulation of each subnetwork can be found in (19). For this analysis, lower bounds and upper bounds of glucose uptake and Fe(II) secretion are constrained based on experimental concentration data. In addition, the upper bounds of all the by-product productions are also constrained based on available experimental data. The model is not constrained with the specific growth rate data, since we wanted the minimal subnetworks to be specific to the desired tasks and not to contain all the pathways required for growth.

5.4 Results and discussion

The metabolic network model allows analysis of several aspects of the metabolism: i) comparison of the model predicted flux ranges for each by-product and growth rate to their experimentally calculated flux ranges; ii) computing flux ranges of all individual reactions within the metabolic network, allowing the comparison of specific reaction contributions into the whole metabolic system between different conditions (i.e., with and without iron); iii) the minimum number of reactions that would support the transformation of substrate to products as observed in the experimental data; iv) analysis of enriched minimal subnetworks for different metabolic tasks to better understand what parts of the networks are significantly affected based on the gene expression data.

The metabolic model of *C. acetobutylicum* did not include any reaction for iron reduction. Iron reducing reactions either with NAD(P)H or with reduced ferredoxin as cofactor were added to the model to analyse the mechanism of iron reduction in this organism.

5.4.1 Experimental and theoretical growth rates

We compared maximum biomass production from model with the calculated fluxes for growth from time dependent OD600 data. The model is constrained with experimental glucose consumption and by-product secretion data. The physiological electron donor for iron reduction for this analysis is NAD(P)H. The error bars are associated to the standard deviation from different replications of the experiments. As shown in Figure 5-1, in time points of 4-hour and 6-hour model predictions are in experimental range or less. This suggests that model may lack some reactions or pathways to be able to achieve more growth and proper gap filling of the network may result in better predictions. In 8-hour, cells slowly start to enter lag phase, however this is not recognizable by a stoichiometric model and therefore the predicted growth is slightly higher compared to experimental value. It should be noted that if Ferredoxin is used as the electron donor, all the experimental constraints cannot be set for the model. In other words, if all by-product productions are constrained based on the experimental data, the model is infeasible. If we relax the constraints on pyruvate production the model becomes feasible. However, the growth predictions are not in agreement with the experimental growth rates. This is one of the first evidences that suggests NAD(P)H is more probable to be the physiological electron donor for iron reduction in *C. acetobutylicum*.

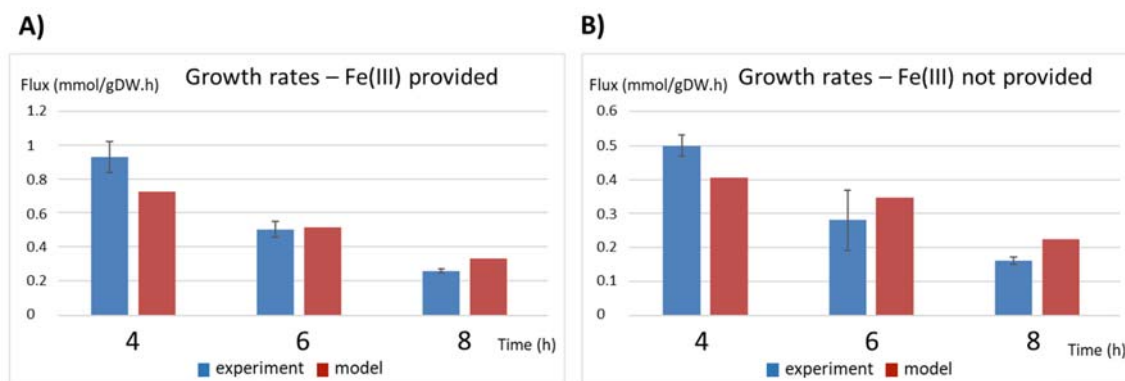


Figure 5-1. Experimental and theoretical growth rates, A) with Fe(III) and, B) without Fe(III)

5.4.2 Experimental and theoretical by-product productions

The flux ranges (minimal and maximal fluxes in mmol/gDW.h) of the experimental by-products (acetate, butyrate, butanol and H₂) produced at three time points (4, 6 and 8 h) were compared with model-predicted fluxes. For this analysis lower bound of the biomass flux is set to 90 percent of the values obtained in the last part and the experimental constraint on

the flux of these four by-products were relaxed one by one. It can be seen that the model largely agrees with the experimental data (Figure 5-2). More specifically, the experimental data mainly fell within the predicted flux ranges except for the 8-hour time point. At that time point, the model predicted a higher minimal and maximal flux for butyrate and acetate compared to what was obtained experimentally, when Fe(III) was not provided. If Fe(III) is provided, the model predicted minimal and maximal flux ranges that were higher for acetate, but lower for H₂ and butyrate compared to experimental data. Since the 6-hour time point shows more agreement with experimental data, both in terms of growth rate and by-product productions, for the rest of our analysis constraints imposed on the model are based on the experimental data of this time point.

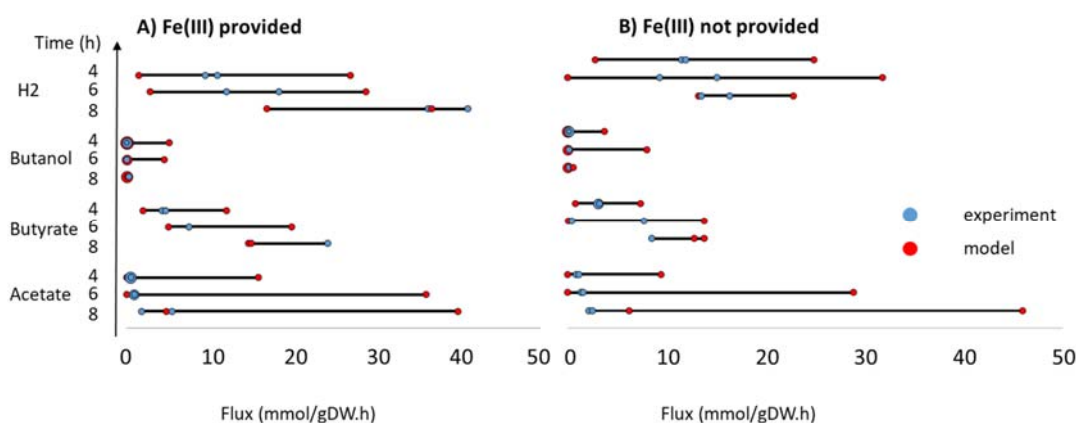


Figure 5-2. Flux ranges of experimental by-products production compared with model-based predictions of fermentation, A) with Fe(III) and, B) without Fe(III). The minimal and maximal flux of acetate, butyrate, butanol and H₂ were compared of experimental data (blue) and model predicted data (red) by using NAD(P)H as physiological electron donor.

5.4.3 Changes in FVA and mean flux values with and without iron reduction

Flux variability analysis (FVA) was performed to determine whether changes are observed in the metabolic flux ranges with and without iron reduction for central metabolic pathways at time point 6 h. The results obtained with 90 percent of maximum biomass as the lower bound of biomass reaction and all other experimental constraints imposed on the model. If Fe(III) was provided, there was a distinct shift in the flux ranges of some reactions compared to the condition with no iron. The most interesting changes in flux ranges was observed for reactions of glycolysis (glucose 6-phosphate to fructose-6-phosphate and phosphoenolpyruvate to pyruvate) and the TCA cycle (oxaloacetate to α -ketoglutarate). The minimum possible flux for these reactions when iron is provided is higher compared to the maximum possible flux of

them when no iron is available. Furthermore, in the riboflavin metabolism, the conversion of riboflavin to Flavin mononucleotide and FAD was distinctly shifted towards higher values with iron reduction which is in agreement with experimental results (148). Finally, a large number of anabolic reactions exhibited a higher flux range in the iron case relative to the no iron case. To visualize the shift in fluxes, 5000 feasible flux distributions were generated by flux sampling to get the mean values for each flux. Then, the ratio (reaction flux with iron reduction/reaction flux without iron reduction) of the mean values are calculated for each reaction. Figure 5-3 and 5-4 show reactions of glycolysis, TCA cycle, pentose phosphate pathway and solventogenesis with NAD(P)H and reduced ferredoxin as the physiological electron donor for iron reduction respectively. Colors show different values for the ratio of fluxes. In most reactions of the central metabolism, the mean of the flux values is higher with iron reduction (>1), with some exceptions. Using NAD(P)H as the physiological electron donor for iron reduction, a number of reactions exhibited greater flux during fermentation without iron reduction: most significantly, a reaction producing ATP (acetyl phosphate + ADP \rightarrow acetate + ATP), one reaction consuming ATP (ribose 5-phosphate + ATP \rightarrow 5-phosphoribosyl 1-pyrophosphate + AMP), the reactions for butanol production, the reaction of d -xylulose 5-phosphate to ribose 5-phosphate and the reaction of 2-deoxy-d -ribose 1-phosphate to glyceraldehyde 3-phosphate (Figure 5-3). Using reduced ferredoxin, the result is similar, but higher flux ranges are obtained with fermentation without iron reduction for another reaction in the citric acid cycle: succinate to succinyl-CoA consuming ATP and the reduction of NADP with reduced ferredoxin to produce oxidized ferredoxin and NADPH (Figure 5-4).

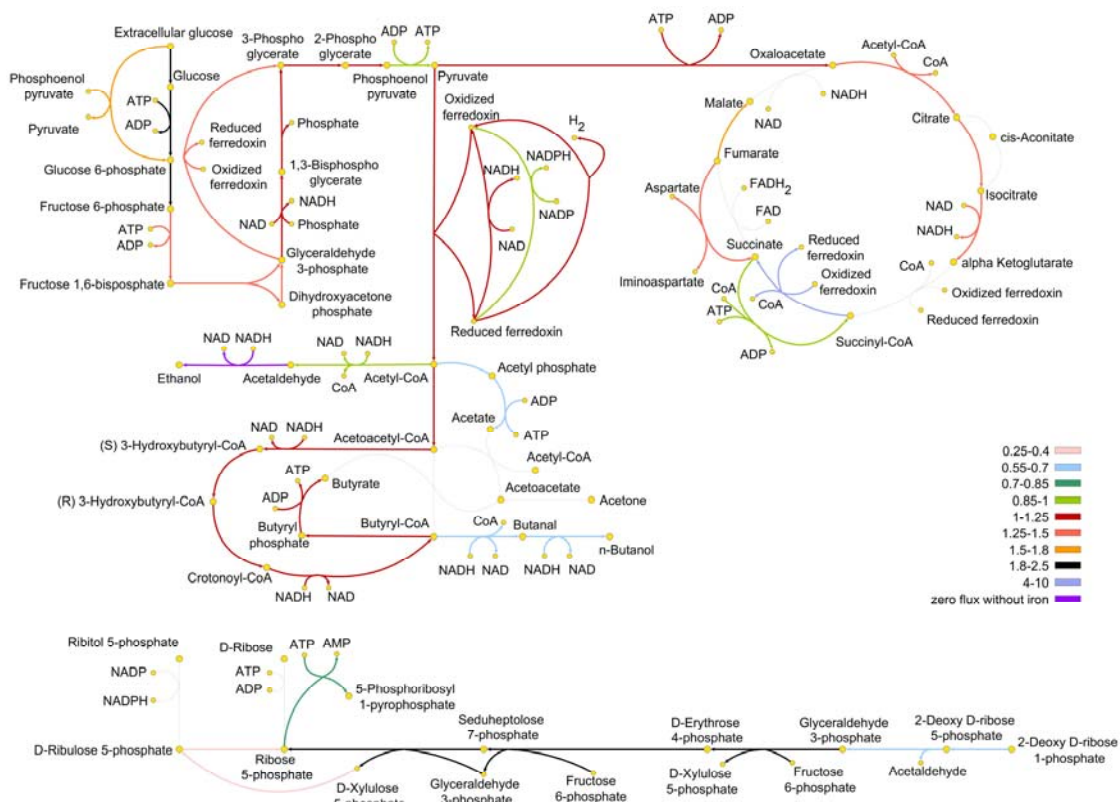


Figure 5-4. The ratio of mean value of fluxes with and without iron reduction using reduced ferredoxin. The mean of flux values (in mmol/gDW.h) from model-based samples was used to calculate the ratio of fermentation with iron reduction divided by fermentation without iron reduction for the reactions of glycolysis, fermentation, citric acid cycle and the pentose phosphate pathway. Grey lines show reactions where no flux was present in the model and gaps in the values within the figure legend indicate that these ranges were not represented. Colours representing values greater than 1 show reactions with higher flux in fermentation with iron reduction and those less than 1 show reactions with higher flux in fermentation without iron reduction.

5.4.4 Minimal reaction subnetwork for fermentation with and without iron reduction

Minimal reaction subnetworks provide the minimum number of reactions that are necessary to allow the transformation of given substrates into given products. There can be more than one minimal subnetwork for a given substrate/product combination. In this section, we imposed all experimental constraints including a lower bound on biomass flux before calculating the minimal reaction subnetworks. Reactions that are present in some, but not all of the minimal networks are termed as 'substitutive reactions'. Minimal networks were obtained for fermentation with and without presence of iron either with NAD(P)H or with ferredoxin as the physiological electron donor for iron reduction. Figure 5-5 and Figure 5-6 show the glycolysis and solventogenesis parts of these minimal subnetworks. It should be

noted that for iron reduction with ferredoxin, experimental constraint for pyruvate production is not imposed on the model, because it will cause model infeasibility. As it can be seen in figures 5-5 and 5-6, no additional reactions are needed for the reduction of Fe(III) relative to fermentation alone (apart from the reaction describing the reduction of Fe(III) by NAD(P)H or reduced ferredoxin). If the iron reduction process uses NAD(P)H, the reactions from butyryl-CoA to form butyrate via butyryl phosphate are not necessary for all minimal networks, whereas they are necessary for fermentation alone or when reduced ferredoxin serves as the physiological electron donor. These can be seen as the reactions that are green in Figure 5-5 and black in Figure 5-6. Red reactions in Figure 5-5, which are gray in Figure 5-6 show the reactions that are substitutive in minimal networks with NAD(P)H but absent in minimal networks of ferredoxin, and the condition with no iron reduction. On the other hand, glycolysis is more flexible with ferredoxin used as the cofactor compared to NAD(P)H or no iron reduction. Reactions that are red and green in Figure 5-6 and black and gray in Figure 5-5 indicate this difference in glycolysis between subnetworks. For fermentation with iron reduction modelled with either reduced ferredoxin or NAD(P)H as the physiological electron donor, more alternative minimal subnetworks can be found compared to fermentation without iron reduction. In conclusion, the higher number of alternative networks and more possible pathways in both glycolysis and solventogenesis, suggest that the network including reactions for iron reduction have higher flexibility compared to the network that cannot reduce iron.

5.4.5 Excess ATP production calculation

The maximal excess ATP production rate was calculated to be 13 mmol/gDW.h for fermentation without iron and 15.6 mmol/gDW.h during iron reduction by using NAD(P)H as a physiological electron donor. Reduced ferredoxin as a physiological electron donor for iron reduction gives a maximum flux for ATP of 9.6 mmol/gDW.h. Since the experimental results confirm that energy generation is increased when iron is provided for the cells, our results suggest that iron reduction is more probably reduced via NAD(P)H as cofactor (148).

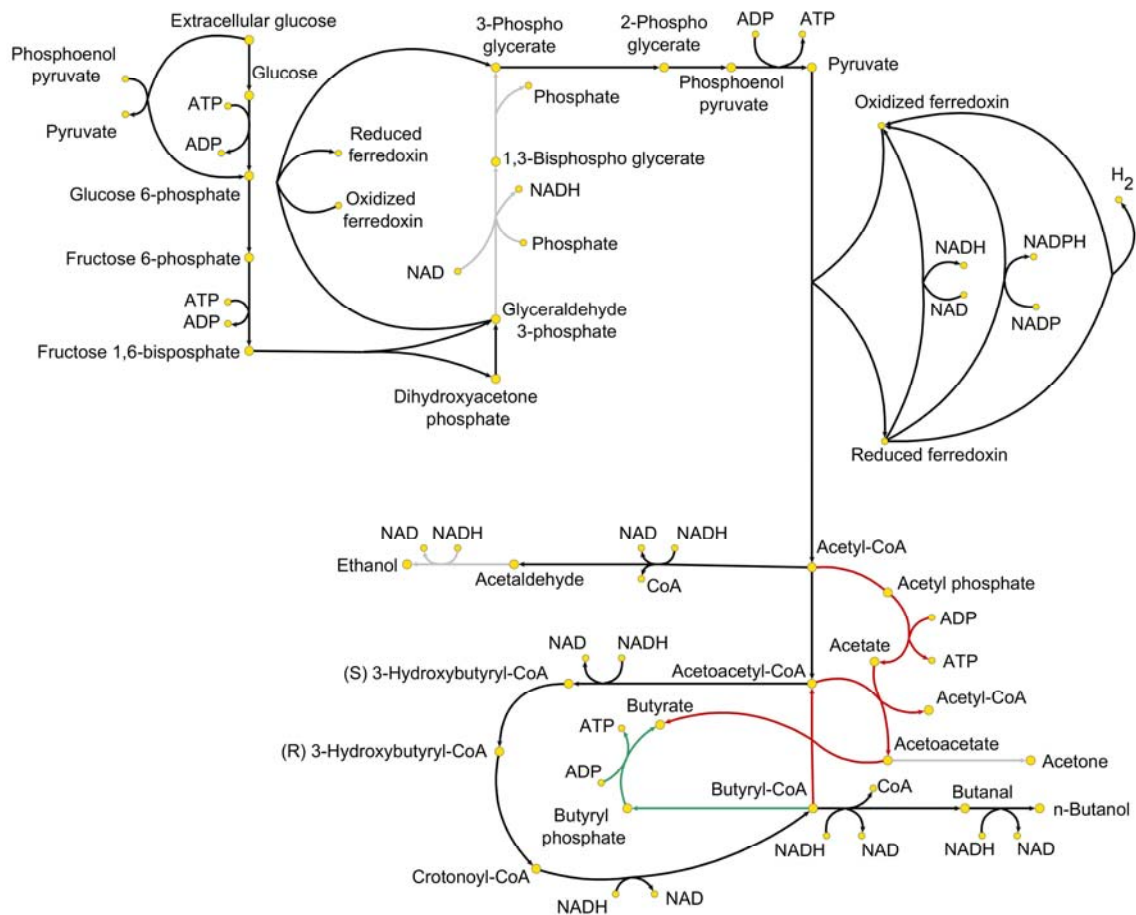


Figure 5-5. Minimal reaction subnetwork for fermentation with and without iron reduction with NAD(P)H as electron donor (only glycolysis and solventogenesis reactions are shown). Black arrows show reactions present in all minimal networks (necessary); green arrows show reactions present in some minimal networks (substitutive) with iron provided and present in all minimal networks of fermentation without iron; and red arrows show reactions present in some minimal networks of fermentation with iron, but not present in any minimal networks of fermentation without iron reduction. Grey arrows are reactions not present in any minimal subnetwork.

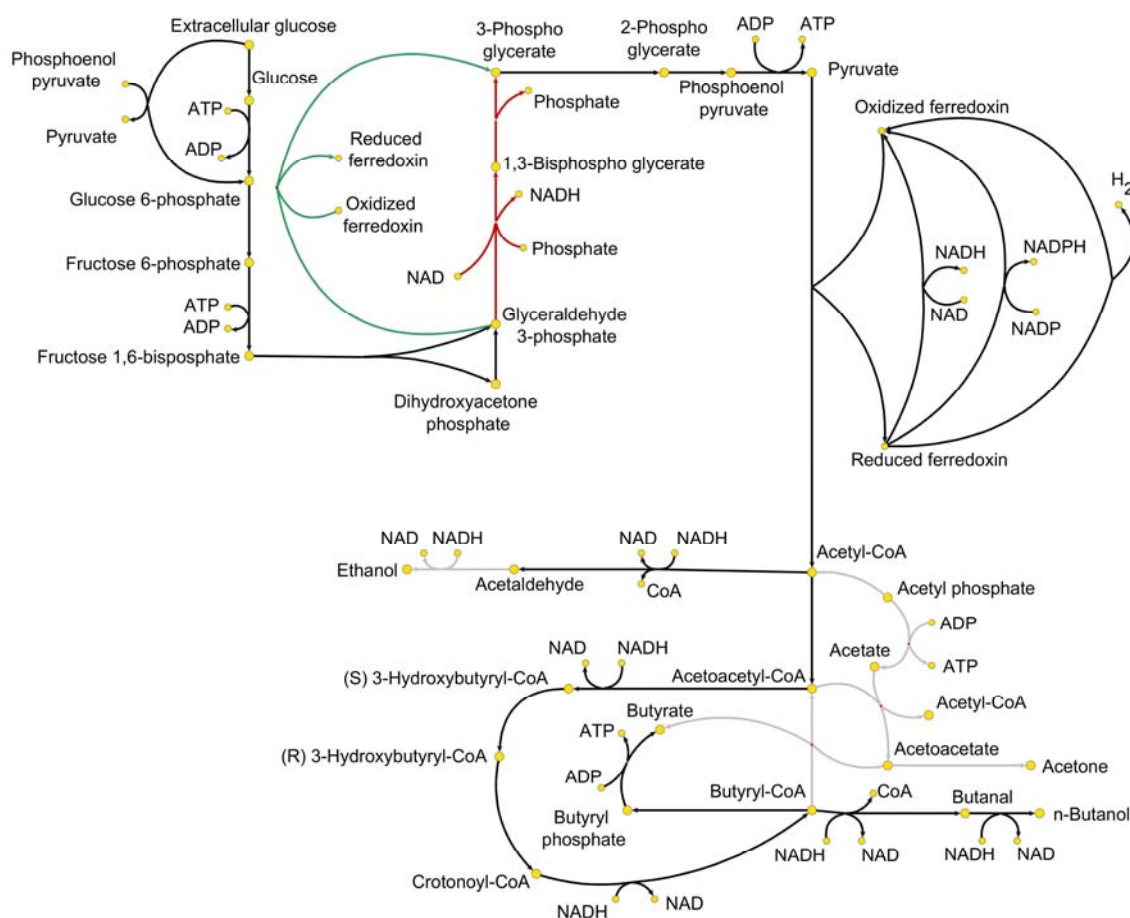


Figure 5-6. Minimal reaction subnetwork for fermentation with and without iron reduction with ferredoxin as electron donor (only glycolysis and solventogenesis reactions are shown). Black arrows show reactions present in all minimal networks (necessary); green arrows show reactions present in some minimal networks (substitutive) with iron provided and present in all minimal networks of fermentation without iron; and red arrows show reactions present in some minimal networks of fermentation with iron, but not present in any minimal networks of fermentation without iron reduction. Grey arrows are reactions not present in any minimal subnetwork.

5.4.6 Enrichment analysis with MiNEA

We performed MiNEA to see if certain metabolic subnetworks are significantly up- or downregulated according to the transcriptomic data (19). We have three sets of transcriptomics data for solid and soluble Fe(III) and for condition when no Fe(III) is provided. From these data list of upregulated and downregulated genes are obtained for solid and soluble Fe(III) compared to no Fe(III) condition. Therefore, we have in total six combinations of conditions that is formed from three different physiological electron donors (NADH, NADPH and reduced ferredoxin) and two different relative gene expression data from solid or soluble iron. For all these six combinations, MiNEA was done for the following tasks: biosynthesis of all biomass building blocks, NADH, Fe(II), lactate, hydrogen, butanol and

reduced ferredoxin. A list of the metabolites for them significantly enriched subnetworks were identified is shown in Table 5-1. It mainly includes amino acids, RNA and cross-linked peptidoglycan. Interestingly, with all the three potential electron donors (NADH, NADPH and ferredoxin) we can find enriched subnetworks for FAD and ATP synthesis in the presence of solid or soluble iron. Besides of the synthesis of some amino acids, there is no difference between the three potential electron donors. With solid iron, the production of NAD(P) and NAD(P)H was also enriched with deregulated genes, showing that NAD(P)H is involved in iron reduction and suggesting that it can be the potential physiological electron donor for Fe(III) rather than ferredoxin.

Table 5-1. List of enriched subnetworks by MiNEA. MiNEA was done by maximizing the formation of certain metabolites using three potential electron donors (NADH, NADPH and ferredoxin) for Fe(III) reduction. Underlined are the differences between the potential electron donors within either solid iron reduction or soluble iron reduction. The model was constrained with glucose and Fe(III) fluxes

Physiological electron donor for Fe(III) reduction	Metabolic tasks for them enriched subnetworks were identified
Solid Fe(III) as electron acceptor	
NADH	NAD(H), NADP(H), acetyl-CoA, CoA, FAD, ATP, dATP, dGTP, RNA, crosslinked peptidoglycan, histidine, isoleucine, leucine, threonine, tryptophan, valine
NADPH	NAD(H), NADP(H), acetyl-CoA, CoA, FAD, ATP, dATP, dGTP, RNA, crosslinked peptidoglycan, histidine, isoleucine, leucine, threonine, tryptophan, valine, <u>phenylalanine, thymine</u>
Ferredoxin	NAD(H), NADP(H), acetyl-CoA, CoA, FAD, ATP, dATP, dGTP, RNA, cross-linked peptidoglycan, histidine, isoleucine, leucine, threonine, tryptophan, valine, <u>phenylalanine, thymine</u>
Soluble Fe(III) as electron acceptor	
NADH	FAD, ATP, dATP, dGTP, RNA, cross-linked peptidoglycan, threonine, tryptophan, valine, menaquinone, arginine, proline, serine, glutamate, glycine, citrulline, <u>thymine</u>
NADPH	FAD, ATP, dATP, dGTP, RNA, cross-linked peptidoglycan, threonine, tryptophan, valine, menaquinone, arginine, proline, serine, glutamate, glycine, citrulline, <u>methionine</u>
Ferredoxin	FAD, ATP, dATP, dGTP, RNA, cross-linked peptidoglycan, threonine, tryptophan, valine, menaquinone, arginine, proline, serine, glutamate, glycine, citrulline, <u>glutamine, leucine</u>

We analyzed the enrichment of NADH, Fe(II), lactate, hydrogen, butanol and reduced ferredoxin subnetworks in more details. We enumerated all the alternative minimal

subnetworks for maximum production of each of these metabolites. Figure 5-7 shows number of alternative subnetworks in different conditions. As it can be seen, NADH has the most flexible biosynthesis among these metabolites due to the large number of alternative subnetworks. It should be noted that the difference in number of alternative subnetworks between two conditions is only because of the difference in the imposed lower and upper bound on uptake fluxes of glucose and Fe(III) and upper bound on the uptake flux of by-products. Figure 5-8 indicates sizes of minimal subnetworks. More parts of the system are involved in NADH synthesis, and lactate, H₂ and Fe(II) need smallest number of reactions for their biosynthesis. In Figure 5-9, number of core reactions (common reactions among all alternative subnetworks) are shown for all metabolites in different conditions. we can see that reduced ferredoxin has only one alternative subnetwork and all of the reactions are core reactions. For almost all other metabolites, when solid iron is provided the number of core reactions are less compared to when soluble iron is provided. It shows that biosynthesis of these metabolites can be more divergent with solid iron. Finally, figure 5-10 indicates the percentage of deregulated core reactions. As it can be seen in Table 5-1, only NADH biosynthesis among these six metabolites is significantly deregulated. However, single reactions within subnetworks are deregulated and it can give a hint for the importance of these metabolites in Fe(III) reduction. The highest amount of upregulated core reactions is in the case of maximized NADH production (Figure 5-10). That is the case for assuming NADH, NADPH or ferredoxin as electron donor. Hence, this result gives more evidence that NADH is the most probable physiological electron donor of Fe(III) reduction. As stated before, other metabolites have also deregulated reactions in their subnetwork. For instance, we can see in Figure 5-10 that Fe(II) and H₂ have more deregulated reaction with solid iron compared to soluble iron. We show the deregulated core genes and corresponding reactions for Fe(II), H₂, butanol and lactate biosynthesis with NADH as electron donor in Table 5-2. These deregulated genes show what parts of production of these important by-products and cofactors are affected by providing Fe(III) for the cell.

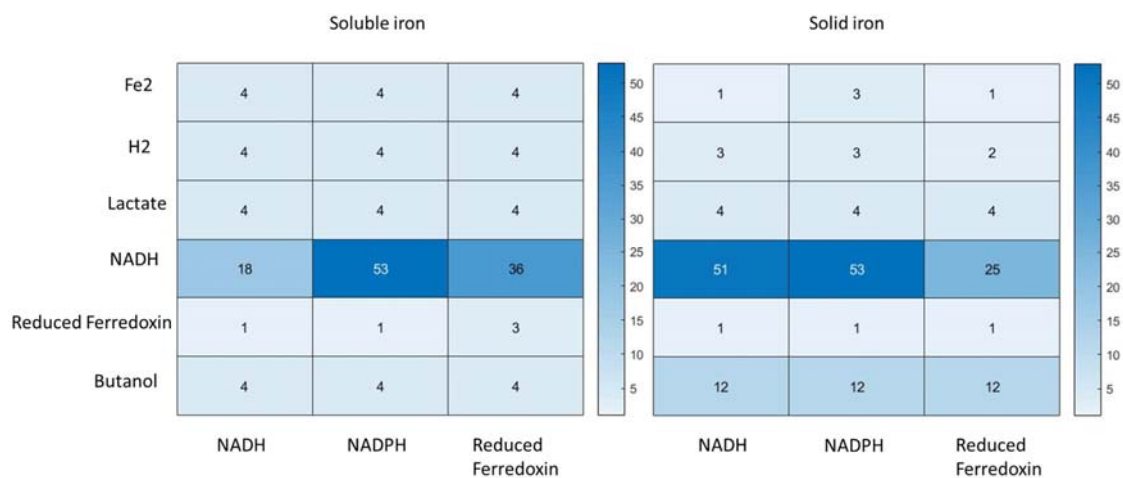


Figure 5-7. Number of alternative minimal subnetworks in different conditions.

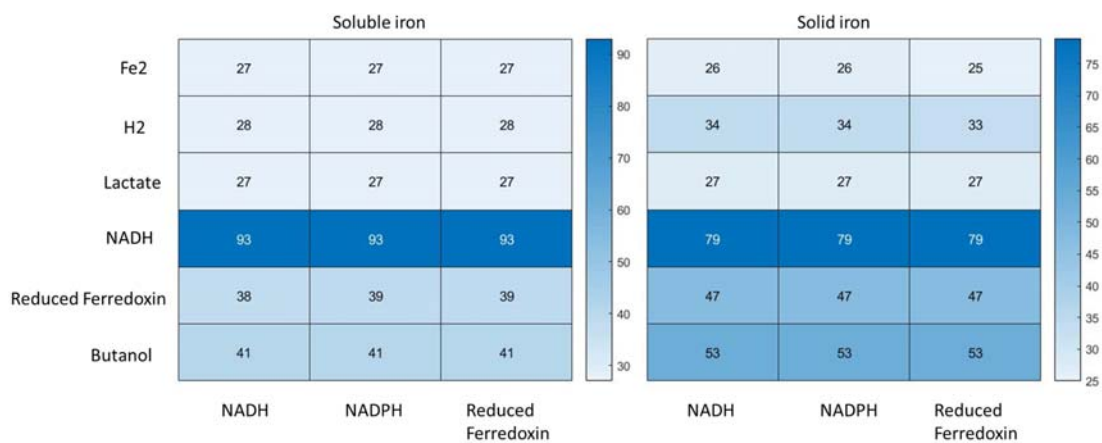


Figure 5-8. Sizes of alternative minimal subnetworks in different conditions.

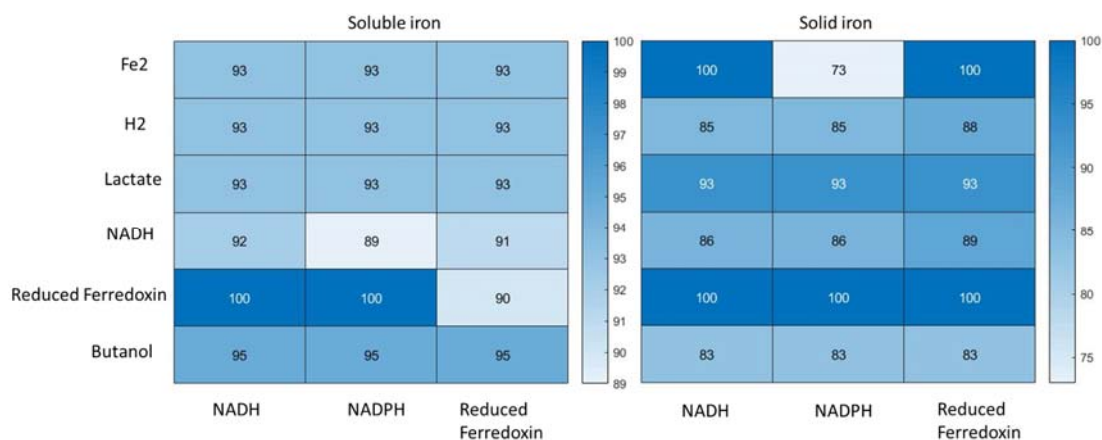


Figure 5-9. Percentage of common reactions between alternative minimal subnetworks in different conditions.

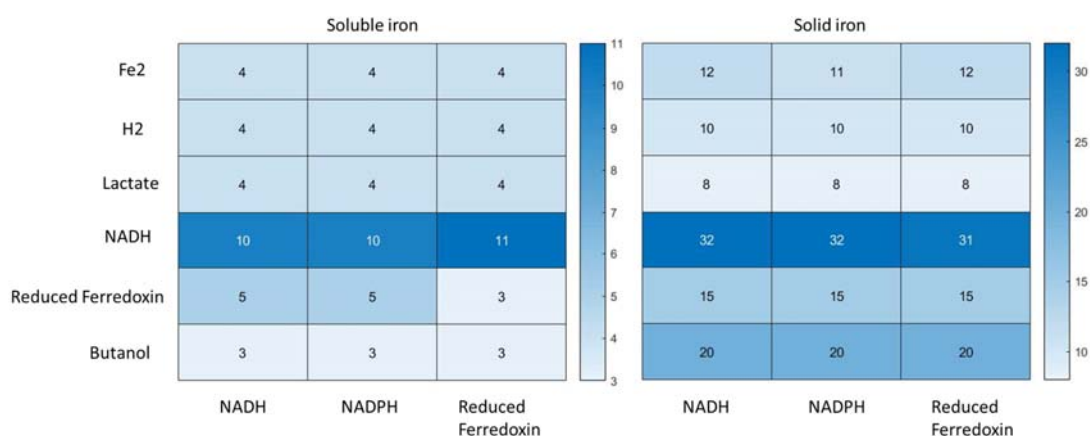


Figure 5-10. Percentage of deregulated common reactions in different conditions.

Table 5-2. Up and down regulated genes and corresponding reactions for the biosynthesis of metabolites that are enriched with deregulated genes when NADH is used as electron donor. Essential genes are those genes that are present in all alternative minimal subnetworks and substitutive genes are those genes that are present only in a subset of alternative minimal subnetworks. e and c after the names of some metabolites show that they are extracellular or cytoplasmic, respectively

Deregulated genes	Associated reactions
Fe(II) synthesis with Solid Fe(III) as electron acceptor	
CAC0682 (essential)	$\text{NH}_4\text{e} \rightleftharpoons \text{NH}_4\text{c}$
CAC0711 (essential)	glyceraldehyde triphosphate \rightleftharpoons dehydroxy acetone phosphate
CAC2018 (essential)	$\text{H}_2\text{O} + \text{oxidized ferredoxin} + \text{glyceraldehyde triphosphate} \rightleftharpoons 3\text{phosphoglycerate} + 3\text{h} + \text{reduced ferredoxin}$
H ₂ synthesis with Solid Fe(III) as electron acceptor	
CAC0682 (essential)	$\text{NH}_4\text{e} \rightleftharpoons \text{NH}_4\text{c}$
CAC0711 (essential)	glyceraldehyde-3-phosphate \rightleftharpoons dehydroxy acetone phosphate
CAC2018 (essential)	$\text{H}_2\text{O} + \text{oxidized ferredoxin} + \text{glyceraldehyde-3-phosphate} \rightleftharpoons 3\text{phosphoglycerate} + 3\text{h} + \text{reduced ferredoxin}$
CAC1673 and CAC1674 (substitutive)	$\text{NADP} + 2\text{L-glutamate} \rightleftharpoons \text{H}^+ + \text{NADPH} + \text{L-glutamine} + 2\text{-oxoglutarate}$
CAC2658 (substitutive)	$\text{nh}_4 + \text{ATP} + \text{L-glutamate} \rightleftharpoons \text{H}^+ + \text{phosphate} + \text{ADP} + \text{L-glutamine}$
CAC3112 (substitutive)	$\text{ATP} + \text{AMP} \rightleftharpoons 2\text{ADP}$
Butanol synthesis with Solid Fe(III) as electron acceptor	
CAC0682 (essential)	$\text{NH}_4\text{e} \rightleftharpoons \text{NH}_4\text{c}$
CAC0711 (essential)	glyceraldehyde-3-phosphate \rightleftharpoons dehydroxy acetone phosphate

CAC1348 (essential)	D-fructose 6-phosphate + glyceraldehyde-3-phosphate <=> D-erythrose4-phosphate + D-xylulose 5-phosphate sedoheptulose7-phosphate + glyceraldehyde-3-phosphate <=> D-xylulose 5-phosphate + D-ribose-5-phosphate
CAC1347 (essential)	sedoheptulose7-phosphate + glyceraldehyde-3-phosphate <=> D-fructose 6-phosphate + D-erythrose4-phosphate
CAC2658 (essential)	NH ₄ + ATP + L-glutamate <=> H ⁺ + phosphate + ADP + L-glutamine
CAC2708 (essential)	NAD + (S)-3-hydroxybutyryl-CoA <=> NADH + H ⁺ + acetoacetyl-CoA
CAC3112 (essential)	ATP + AMP <=> 2 ADP
CAP0035 (essential)	NAD + CoA + acetaldehyde <=> NADH + H ⁺ + acetyl-CoA
Lactate synthesis with Solid Fe(III) as electron acceptor	
CAC0711 (essential)	glyceraldehyde-3-phosphate <=> dehydroxy acetone phosphate
CAC2018 (essential)	H ₂ O + oxidized ferredoxin + glyceraldehyde-3-phosphate <=> 3phosphoglycerate + 3 H ⁺ + reduced ferredoxin
Fe(II) synthesis with Soluble Fe(III) as electron acceptor	
CAC0532 and CAC0570 (essential)	phosphoenolpyruvate + glucose_e <=> pyruvate + glucose 6-phosphate
H ₂ synthesis with Soluble Fe(III) as electron acceptor	
CAC0532 and CAC0570 (essential)	phosphoenolpyruvate + glucose_e <=> pyruvate + glucose 6-phosphate
Butanol synthesis with Soluble Fe(III) as electron acceptor	
CAC0532 and CAC0570 (essential)	phosphoenolpyruvate + glucose_e <=> pyruvate + glucose 6-phosphate
Lactate synthesis with Soluble Fe(III) as electron acceptor	
CAC0532 and CAC0570 (essential)	phosphoenolpyruvate + glucose_e <=> pyruvate + glucose 6-phosphate

5.5 Conclusion

In this chapter, we used a reconstructed genome-scale metabolic network of *C. acetobutylicum* to study the mechanisms behind iron reduction in this model organism. Our computational studies were validated by comparing our results with experimental growth rate and metabolite concentrations data. Our work shows the importance of using computational models as a platform to integrate experimental metabolomics and transcriptomics data to better understand the behavior of cells in different conditions. The

usage of experimental data helped us to understand important reactions, genes and subnetworks that are involved in iron reduction. In addition, we indicated that providing iron for the cell could affect different aspects of metabolism, even those parts that are not obviously related to iron reduction or fermentation. We showed that synthesis of ATP and FAD are significantly affected in all conditions. The mechanism of iron reduction is not perfectly known in *C. acetobutylicum*. Our results showed that with NAD(P)H as cofactor for iron reduction, ATP synthesis is more in agreement with experiments. In addition, biosynthesis of NADH is highly affected when Fe(III) is provided for the cells. Therefore, our analysis suggests that NAD(P)H is the potential physiological electron donor in process of iron reduction. In conclusion, this systematic evaluation of metabolism using genome-scale metabolic model helped us to have a better understanding of metabolic changes during iron transfer and reduction by *Clostridium acetobutylicum*.

6 Conclusions and perspectives

In this thesis, we had mainly three perspectives i) investigating the computational methods and objective functions that have been used in the study of metabolism in both biological and mathematical points of view; ii) developing a novel method for analysing complex cellular objectives for metabolic models with integer variables; and iii) use known and novel computational methods for analysing metabolic models of specific organisms for different purposes.

In the second chapter of this thesis, a repository of objective functions and optimization methods has been built based on the well-known or recent research studies of metabolic network modelling. We made a classification system based on the type or purpose of studies. Our work shows how optimization methods have been evolved to better capture the complexities of biological systems. For instance, metabolic engineering methods started from simple single gene knock-out studies, to systematic evaluation of effect of single or multiple gene knock outs on the production of desired metabolites. It goes further to assess not only the effect of gene knock outs but also gene over or under expressions on production yield. In addition, the advancement in computational strategies made it possible to integrate different kind of experimental data into metabolic network models and hence have the models that encapsulate more aspects of metabolism and can make more accurate predictions. In addition to reviewing biological aspects of strategies for analysis of metabolic networks, we have also proposed several mathematical reformulations to transform complex optimization methods to simpler problems that can be solved to global optimality. Particularly, we analysed the conversion of fractional programming to linear programming both with and without integer variables and bi-level programming to single-level programming. Moreover, we provided some strategies for solving multi-objective problems. Our work made a comprehensive platform to formulate and solve more complicated problems with several simple or complex objectives in the future. This is especially necessary to analyse more complicated metabolic activities to better understand the metabolic behaviour of more

complex organisms such as eukaryotes or less known ones such as bacteria or parasites in dormant phase.

One of the important features that make the study of metabolic networks difficult is the interconnection of different elements of these networks. In other words, one metabolite can be involved in so many reactions and therefore the activities of reactions are not independent of each other. Flux coupling analysis is a computational method to study these flux dependencies in metabolic network models (8). It is originally formulated for FBA-based models with no integer variables. However, integer variables are now an important part of optimization methods in metabolic network modelling. They are mainly used for turning on and off different variables such as metabolic fluxes and are extremely important in metabolic engineering studies. In chapter three of this thesis, we developed a new general formulation for metabolic models that consists of integer variables and have a fractional objective function. We then, used this general formulation for the specific problem of calculating flux coupling analysis in thermodynamically constrained metabolic network models (5), which we called TFCA. We studied the effect of introducing thermodynamic properties on flux coupling relationships on a reduced *E. coli* model. We introduced the concept of conditionally directionally coupled reaction pairs to emphasize the effect of having different physiologies on reaction dependencies. We showed that number and therefore, combination of BDRs is reduced because of having reaction pairs that are conditionally coupled to each other. By making FDP trees based on conditionally coupled reaction pairs, one can realize which reactions have the most important role in dictating directionality of other reactions. In future studies, this feature of TFCA can be used for guiding researchers to more realistically select for the suitable FDP in the study of metabolism specially for kinetic models. This is particularly important, since kinetic models of metabolism already involve large number of parameters and reducing the complexity of metabolic model under study can facilitate the kinetic analysis. A very important follow-up study for TFCA is proposing genetic modification strategies to couple the production of a metabolite of interest to a target flux in the network such as biomass flux. This is previously done for Flux-balanced models of metabolism (109). However, the proposed method does not capture the thermodynamic properties of reactions. In order to get thermodynamically feasible and therefore more reliable metabolic engineering strategies one can use TFCA to couple desired products with the flux of interest in the model by identifying proper gene knock-outs, over-expressions or under-expressions. In general,

based on the proposed strategies in chapter two and three of this thesis, one can extend any method that is only beneficial for stoichiometrically balanced models to be used for the study of thermodynamically constrained models. One of the interesting ideas for doing so is to have Optknock (20) framework for TFA models. This is mathematically challenging because of the bi-level nature of the problem and the fact that both inner and outer levels of the problem include integer variables. However, it can give us promising engineering strategies, since thermodynamic feasibilities of the solutions are ensured in advance with the way the problem is formulated. Since these problems are examining the metabolism from different aspects, the comparison between the result of this study with the study of forcing coupling relations in TFA models can give an idea of the performance, efficiency and usefulness of different engineering methods.

The aim of fourth chapter of this thesis, is to analyse malaria parasites in the dormancy phase. It is a poorly characterized stage in parasite life cycle and the biological goal of parasite in this stage is not known. However, some information about dormancy are proposed before which are known as hallmarks of dormancy. We used these information in addition to several potential objective functions in order to study the cellular objectives of dormancy which leads to better understanding of metabolic behaviour of parasite in this phase. Our results suggested that among all the objective functions that we tested, dormant parasites aim to minimize their energy dissipation. Minimization of energy dissipation could better capture up-regulation and down-regulation of metabolic genes in the dormancy phase compared to other objective functions. Some of the current drugs for malaria parasite are targeting protein synthesis pathways. However, our analysis suggest that these genes are already down-regulated in the dormancy phase. Hence, for future drug targeting it is important to take this into account and prevent using drugs that may trigger entering the dormancy phase. Instead, combination of drugs that can attack appropriate genes in both non-dormant and dormant parasites can be most effective. For future follow-up study of chapter four, one can analyse the essential genes toward the objective function that can best describe this phase. Since the objective function is not growth, essentiality here does not mean growth reduction and lethality. It shows the increase in total dissipated energy if a single gene is knocked out. One of the most important steps toward improvement of this study is the reconstruction *P. vivax* genome-scale model, which is one of the *plasmodium* species that is actually going through the dormancy phase and the experimental data that is used here belongs also to this parasite.

Moreover, from the repository of objective functions from chapter two of this thesis, additional single or combinations of objective functions can be examined for the dormant parasite. Finally, since little information is available for the dormancy phase in general, the framework that is presented in chapter four can be used for the study of dormancy in other organisms such as *Mycobacterium tuberculosis*.

In chapter five of this thesis, we aim to characterize the mechanism of iron reduction in *Clostridium acetobutylicum* using experimental gene expression and metabolite concentration data. It is shown that this anaerobic gram positive bacterium is able to reduce Fe(III) in solid or soluble phase, but the involved enzymes and active pathways are not determined. Our results suggest that NAD(P)/NAD(P)H play a role in iron reduction and most probably are physiological electron donors for Fe(III). In addition, we identified up-regulated biosynthesis pathways for several metabolic tasks using MiNEA framework (19). For future studies, one can analyse essential genes and reactions for iron reduction process and suggest gene-knock out strategies to increase iron reduction. In addition, redGEM (131) and lumpGEM (132) can be used to focus the study of *C. acetobutylicum* metabolism around the iron reduction, solventogenesis and other subsystems of interest.

In this thesis, we discussed known and novel computational methods for the study of cell metabolism and evaluated the performance of these methods on metabolic network models of several organisms. The increasing availability of experimental data elevate the need for having appropriate methods to integrate these data to genome-scale models and improve the obtained computational results. In this thesis, we used gene expression and metabolite concentration data for different purposes. Strategies and methods proposed in this thesis can be applied on any metabolic network model for the same or different research purposes.

7 Appendix

Table 7.1

	Objective Function	Type of Objective							Type of Problem								Goal of Study			Consistency with This Formulation		References
		LP	MILP	LFP	QP	MIQP	QFP	MO ⁸	LP	MILP	LFP	QP	MIQP	QFP	MO	BL ⁹	Phys ¹	App ¹	DI ¹	NRN ¹	RN ¹	
1	Max. biomass (growth rate)	*							*	*							*	*	*	*		(5,28,84,159–184)
2	Max. ATP yield (ATP flux divided by substrate flux)			*							*						*	*			*	(185)
3	Max. ATP per flux unit (ATP flux divided by sum of fluxes)			*							*						*				*	(36)
4	Max. ATP production	*							*								*	*	*	*		(32,47,48,186–199)
5	Max. number of reactions whose activity is		*							*									*	*		(17,35,200–202)

⁸ Multi-objective

⁹ Bi-level

¹ Physiology ⁰

¹ Application ¹

¹ Data Integration ²

¹ No reformulation is needed ³

¹ Reformulation is needed ⁴

	consistent with their expression state																					
6	Min. the inconsistency between gene expression and flux values	*							*									*	*			(81,201,203–206)
7	Min. the number of reactions that can carry flux and produce a specific set of metabolites		*							*								*	*			(19)
8	Max. the consistency between relative experimentally observed changes in gene expression and metabolite changes with the flux levels		*							*								*	*			(83)
9	Outer: Max. bioengineering objective Inner: Max. biomass	*								*					*		*			*		(20,68,207–214)
10	Outer: Max. bioengineering objective and Min. (weighted) number of gene	*	*							*					*		*			*		(99)

	deletions or over-expressions Inner: Max. biomass																					
11	Min. distance from a given flux vector	*			*				*			*					*	*	*	*		(7,56,63,64,78,89,101,102,121,156,168,198,215–226)
12	Min. the number of significant flux changes		*							*							*	*	*	*		(33,222,227)
13	Max. sum of a subset internal fluxes	*							*								*	*		*		(199)
14	Min. the difference between fluxes (whole flux vector or parts of it) and Min. the total sum of square of fluxes				*							*					*		*		*	(79,228)
15	Min. the total (weighted) sum of absolute fluxes	*							*								*		*	*		(38,51,78,229–234)
16	Min. the number of active reactions		*							*							*			*		(49,50,54,235)
17	Min. the set of possible exchanged metabolites between two organisms		*							*							*			*		(55)

18	Outer: Max. the similarity with fluxomics and metabolomics data Inner: Min. the squared sum of fluxes				*						*			*		*	*		*	(106)
19	Max. biomass divided by squared sum of fluxes					*					*		*	*	*			*	(185)	
20	Outer: Max. community growth (or any community-level objective) Inner: Max. species growth (or any species-level objective)	*										*	*	*	*			*	(44)	
21	Max. ATP yield per unit of flux (sum of squared fluxes)					*					*		*	*	*			*	(185)	
22	Outer: Max. bioengineering objective Inner: weighted Max. of biomass and Min. of sum of fluxes	*							*			*	*	*	*			*	(80)	

23	Min. the metabolite turnover	*							*								*	*		*		(236)
24	Max. the weighted sum for patterns of gene activation and inactivation		*							*									*	*		(82)
25	Max. the production of a specific metabolite	*							*								*	*	*	*		(31,65–69,77,104,185,191,229,237–260)
26	Min. the total sum of square of fluxes				*							*					*			*		(53,261)
27	Max. the rate of protein translation	*							*								*		*	*		(262,263)
28	Max. biomass divided by a weighted sum of square of fluxes and square of ATP production flux						*						*				*			*		(264)
29	Outer: Min. the Euclidean distance between flux predictions and experimentally observed fluxes Inner: Max. the weighted sum of	*			*										*		*		*	*		(15)

	multiple linear objectives																				
30	Min. (weighted) number of reactions to be added to the model		*						*								*	*	*		(21,90,108,265–268)
31	Min. the number of reactions that must be removed from the model		*						*								*	*	*		(90)
32	Max. the presence of high-score reactions and Min. the presence of low-score reactions in an organism specific model		*						*					*			*	*	*		(267)
33	Outer: Max. minimum bioengineering objective Inner: max. biomass	*							*					*			*		*		(91)
34	Outer: min number of knock ups/downs/outs Inner: max. bioengineering product	*	*						*					*			*		*		(94,269,270)

35	Max. growth Max. production of byproducts	*	*											*		*	*	*		*	(30,66,92,93,100,2 10,224,225,234,27 1–287)
36	Max. biomass Min. substrate consumption Min. undesired byproduct production	*												*		*	*			*	(288)
37	Min. uptake of a specific substrate Max. production of a specific substrate	*												*			*			*	(289)
38	Max. the uptake of a specific substrate	*						*								*		*	*		(67,290,291)
39	Min. absolute sum of fluxes Min. deviation of log concentrations of metabolites from their experimentally measured values						*							*		*		*		*	(292)
40	Min. the absolute difference between measurements and predicted fluxes multiplied						*							*		*		*		*	(293)

	by a scaling variable Min. the L1-norm distance between relative estimates of metabolite secretion/uptake and predicted fluxes Min. absolute sum of fluxes																				
41	Min. (squared) sum of fluxes Max. biomass	*			*						*			*		*		*		*	(18,64,72,76,198,22,232,239,251,280,294–303)
42	Max. growth, Max. ATP production Min. the total abundance of metabolic enzymes Min. the carbon uptake	*												*		*				*	(73)
43	Outer: Max. bioengineering objective Inner: Min. bioengineering objective	*								*					*		*			*	(95,96)
44	Inner: Max. the cellular objective, Outer: Min. the squared sum of	*				*						*		*	*			*		*	(304)

	differences between predictions and measured fluxes and the number of reactants and products in the inferred reaction																				
45	Inner: Min. the uptake rate of a specific metabolite Outer: Max. the uptake rate of the similar metabolite	*												*		*			*		(305)
46	Max. growth, Max. demand flux for metabolites with increased concentration, Min. demand flux for metabolites with decreased concentration	*											*			*			*		(88)
47	Min. the number of reactions that connect an extracellular metabolite to the core		*						*						*		*	*			(306)
48	Max. production of virulence factors (definition	*						*							*			*			(307)

	similar to biomass)																				
49	Max. flux through symbiosis reaction (definition similar to biomass)	*						*								*			*		(275)
50	Max. weighted sum of secretion of biomass building blocks and Min. weighted sum of number of added reactions to the model	*	*						*					*			*		*		(308)
51	Max. the non-growth associated maintenance	*						*								*			*		(68,309–311)
52	Min. the sum of deviation from mass action kinetics and deviation from reference fluxes due to SNPs		*						*									*	*		(85)
53	Inner: Max. biomass Middle: Min. the upper bound of pyruvate kinase,	*												*		*		*		*	(312)

	Outer: Min. flux through lactate dehydrogenase ¹																				
54	Inner: Min. the production of the specific product Outer: Min. the number of reaction knockouts	*	*						*					*		*		*			(96)
55	Max. the usage of reactions in host to simulate the maximum metabolic exploitation of pathogen	*							*						*			*			(235)
56	Min. the L1 norm of slack variables for the reactions (lower and upper bounds calculated based on gene expression data)	*							*							*	*	*			(313)
57	Min. photon usage Min. sum of fluxes Max. biomass	*											*		*			*			(314)
58	Min. number of reaction		*						*							*		*			(315)

¹ This is a tri-level optimization.

	knockouts (static) and regulated valves (dynamic) to allow switching between two distinct metabolic phenotypes																				
59	Min. weighted sum of fluxes and violation of reaction bounds from experimental measurements	*			*						*			*			*		*		(86)
60	Outer: Max. growth Inner: Max. ATP maintenance	*												*	*		*		*		(316)
61	Max. the minimum lower bound of driving force of all reactions in a pathway		*						*						*			*			(317)
62	Max. the absolute value of reaction fluxes		*						*						*			*			(318)
63	Outer: max. turnover rate of bioengineering product	*							*					*		*		*			(318)

	Inner: max. biomass																					
64	Max. (combination of all) metabolic tasks (specific to a tissue)	*							*								*	*		*		(22,319–322)
65	Min. variation between flux prediction and the flux calculated from proteomics data				*						*								*	*		(37)
66	Max. a sink reaction of measured metabolites with their concentrations as stoichiometric coefficient	*							*								*		*	*		(323)
67	Min. the difference between experimentally constrained and unconstrained model and Min. the number of constraints and Max. a biologically	*	*		*					*		*		*					*		*	(87)

	meaningful objective																				
68	Max. biomass Min. sum of absolute values of all fluxes Min. flux differences between two conditions	*												*		*		*		*	(324)
69	Min. difference between predicted and estimated flux values Max. flux of ATP maintenance reaction Min. squared sum of fluxes	*			*						*			*		*		*		*	(325)
70	Min. number of reactions in the media		*							*						*		*	*	*	(34,55,235,306,326)
71	Max. consistency score with expression data		*							*							*	*	*	*	(326,327)
72	Max. (community) biomass Min. substrate uptake	*												*		*	*			*	(46,56–59,240,274,310,328)
73	Max. biomass	*	*							*				*		*			*	*	(34)

	Min. number of uptakes Max. consumption of specific metabolite																				
74	Max. biomass Min. number of uptakes Max. secretion of specific metabolite	*	*							*					*		*			*	(34)
75	Min. substrate uptake	*							*							*	*	*	*		(45,60,196,329)
76	Max. biomass Max. ATP production	*												*		*	*		*	*	(239,261,330)
77	Max. biomass Max. NAD(P)H production	*												*		*			*	*	(239)
78	Max. the ratio of biomass production rate to the substrate uptake rate			*							*					*	*		*	*	(29,185,331)
79	Max. biomass Min. NAD(P)H production or consumption	*												*		*			*	*	(61)
80	Outer: Min. the difference between kinetic	*			*						*			*	*	*			*	*	(332)

	and FBA growth rates Inner: Min. squared sum of fluxes and maximizing biomass																					
81	Max. biomass Max. ATP production Max. production of specific metabolites Min. sum of absolute fluxes	*												*		*					*	(71)
82	Max. biomass Max. production of a specific metabolites Min. (squared) sum of fluxes	*			*						*			*		*	*				*	(333,334)
83	Max. biomass Max (or min). specific metabolite secretion Max. specific substrate uptake	*												*		*	*				*	(335)
84	Outer: combination of Min. maximum biomass	*	*							*				*	*		*				*	(116)

	Max. maximum production of specific metabolite Min. number of knockouts Inner: Max. biomass																					
85	Min and Max. flux ratios			*						*						*					*	(8)
86	Outer: Min. biomass Inner: Max. biomass	*												*		*	*				*	(21,70)
87	Max. weighted sum of biomass fluxes Min. square of growth rate of each species	*			*						*			*		*					*	(46)
88	Max. ATP hydrolysis Min. weighted sum of substrate uptakes	*												*		*					*	(336)
89	Max. biomass Max. ATP production Max. production of specific product	*												*		*					*	(337)

90	Max. combination of internal fluxes Max. biomass Max. uptake of specific substrates Max. Secretion of specific products	*												*		*				*	(338)
91	Max. ATP production Min. total sum of fluxes	*												*		*				*	(74)
92	Max. ATP consumption	*							*							*			*		(339,340)
93	Min. number of metabolite turnovers		*							*						*			*		(222)
94	Min. flux through drug affected reaction Max. biomass Max. sum of fluxes	*												*		*	*			*	(294)
95	Max. ATP production Max. NADP to NADPH conversion	*												*		*				*	(341)
96	Max. total amino acid content of the whole community	*							*							*			*		(342)

97	Max. sum of production of specific metabolites in community	*							*								*			*		(247)
98	Max. biomass of different microbial species Min. total sum of fluxes	*												*			*	*			*	(197,343,344)
99	Max. biomass of different microbial species	*												*			*		*		*	(42,43,345–357)
100	Max. NADPH production rate	*							*									*		*		(105)
101	Max. NADPH generation from one pathway and min. NADPH generation from another pathway	*												*				*			*	(289)
102	Max. product yield Min. number of added reactions to have that yield		*	*						*	*			*				*			*	(98)
103	Max. bioengineering objective Min. number of added reactions to have that production	*	*							*				*				*			*	(97)

8 References

1. King ZA, Lloyd CJ, Feist AM, Palsson BO. Next-generation genome-scale models for metabolic engineering. *Current opinion in biotechnology*. 2015;35:23–9.
2. Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews Genetics*. 2014;15(2):107–20.
3. Yurkovich JT, Palsson BO. Solving puzzles with missing pieces: the power of systems biology. *Proceedings of the IEEE*. 2015;104(1):2–7.
4. Ataman M, Hatzimanikatis V. Heading in the right direction: thermodynamics-based network analysis and pathway engineering. *Current opinion in biotechnology*. 2015;36:176–82.
5. Henry CS, Broadbelt LJ, Hatzimanikatis V. Thermodynamics-based metabolic flux analysis. *Biophysical journal*. 2007;92(5):1792–805.
6. Schuetz R, Kuepfer L, Sauer U. Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Molecular systems biology*. 2007;3(1):119.
7. Segre D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences*. 2002;99(23):15112–7.
8. Burgard AP, Nikolaev EV, Schilling CH, Maranas CD. Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome research*. 2004;14(2):301–12.
9. Chiappino-Pepe A, Tymoshenko S, Ataman M, Soldati-Favre D, Hatzimanikatis V. Bioenergetics-based modeling of *Plasmodium falciparum* metabolism reveals its essential genes, nutritional requirements, and thermodynamic bottlenecks. *PLoS computational biology*. 2017;13(3):e1005397.
10. Dash S, Mueller TJ, Venkataramanan KP, Papoutsakis ET, Maranas CD. Capturing the response of *Clostridium acetobutylicum* to chemical stressors using a regulated genome-scale metabolic model. *Biotechnology for biofuels*. 2014;7(1):144.
11. Floudas CA. *Deterministic global optimization: theory, methods and applications*. Vol. 37. Springer Science & Business Media; 2013.
12. Gu C, Kim GB, Kim WJ, Kim HU, Lee SY. Current status and applications of genome-scale metabolic models. *Genome biology*. 2019;20(1):121.

13. Nielsen J. Systems biology of metabolism. *Annual review of biochemistry*. 2017;86:245–75.
14. Schuster S, Pfeiffer T, Fell DA. Is maximization of molar yield in metabolic networks favoured by evolution? *Journal of theoretical biology*. 2008;252(3):497–504.
15. Burgard AP, Maranas CD. Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnology and bioengineering*. 2003;82(6):670–7.
16. Gianchandani EP, Oberhardt MA, Burgard AP, Maranas CD, Papin JA. Predicting biological system objectives de novo from internal state measurements. *BMC bioinformatics*. 2008;9(1):43.
17. Zur H, Ruppin E, Shlomi T. iMAT: an integrative metabolic analysis tool. *Bioinformatics*. 2010;26(24):3140–2.
18. Correia DM, Sargo CR, Silva AJ, Santos ST, Giordano RC, Ferreira EC, et al. Mapping *Salmonella typhimurium* pathways using ¹³C metabolic flux analysis. *Metabolic engineering*. 2019;52:303–14.
19. Pandey V, Hatzimanikatis V. Investigating the deregulation of metabolic tasks via Minimum Network Enrichment Analysis (MiNEA) as applied to nonalcoholic fatty liver disease using mouse and human omics data. *PLoS computational biology*. 2019;15(4):e1006760.
20. Burgard AP, Pharkya P, Maranas CD. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and bioengineering*. 2003;84(6):647–57.
21. Kumar VS, Maranas CD. GrowMatch: an automated method for reconciling in silico/in vivo growth predictions. *PLoS Comput Biol*. 2009;5(3):e1000308.
22. Larsson I, Uhlén M, Zhang C, Mardinoglu A. Genome-Scale Metabolic Modeling of Glioblastoma Reveals Promising Targets for Drug Development. *Frontiers in Genetics*. 2020;11:381.
23. Charnes A, Cooper WW. Programming with linear fractional functionals. *Naval Research logistics quarterly*. 1962;9(3-4):181–6.
24. Glover F. Improved linear integer programming formulations of nonlinear integer problems. *Management Science*. 1975;22(4):455–60.
25. D’Huys P-J, Lule I, Vercammen D, Anné J, Van Impe JF, Bernaerts K. Genome-scale metabolic flux analysis of *Streptomyces lividans* growing on a complex medium. *Journal of biotechnology*. 2012;161(1):1–13.
26. Adadi R, Volkmer B, Milo R, Heinemann M, Shlomi T. Prediction of microbial growth rate versus biomass yield by a metabolic network with kinetic parameters. *PLoS Comput Biol*. 2012;8(7):e1002575.

27. Varma A, Palsson BO. Metabolic capabilities of *Escherichia coli* II. Optimal growth patterns. *Journal of Theoretical Biology*. 1993;165(4):503–22.
28. Edwards JS, Palsson BO. Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *Journal of Biological Chemistry*. 1999;274(25):17410–6.
29. Georgiev GY, Aho T, Kesseli J, Yli-Harja O, Kauffman SA. Action and power efficiency in self-organization: the case for growth efficiency as a cellular objective in *Escherichia coli*. In: *Evolution, Development and Complexity*. Springer; 2019. p. 229–44.
30. Pereira B, Miguel J, Vilaça P, Soares S, Rocha I, Carneiro S. Reconstruction of a genome-scale metabolic model for *Actinobacillus succinogenes* 130Z. *BMC systems biology*. 2018;12(1):61.
31. Dias O, Saraiva J, Faria C, Ramirez M, Pinto F, Rocha I. iDS372, a phenotypically reconciled model for the metabolism *Streptococcus pneumoniae* strain R6. *Frontiers in microbiology*. 2019;10:1283.
32. Ramakrishna R, Edwards JS, McCulloch A, Palsson BO. Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*. 2001;280(3):R695–704.
33. Shlomi T, Berkman O, Ruppin E. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proceedings of the national academy of sciences*. 2005;102(21):7695–700.
34. Rosario D, Benfeitas R, Bidkhorji G, Zhang C, Uhlen M, Shoaie S, et al. Understanding the representative gut microbiota dysbiosis in metformin-treated type 2 diabetes patients using genome-scale metabolic modeling. *Frontiers in physiology*. 2018;9:775.
35. Shlomi T, Cabili MN, Herrgård MJ, Palsson BØ, Ruppin E. Network-based prediction of human tissue-specific metabolism. *Nature biotechnology*. 2008;26(9):1003–10.
36. Dauner M, Sauer U. Stoichiometric growth model for riboflavin-producing *Bacillus subtilis*. *Biotechnology and bioengineering*. 2001;76(2):132–43.
37. Yizhak K, Benyamini T, Liebermeister W, Ruppin E, Shlomi T. Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. *Bioinformatics*. 2010;26(12):i255–60.
38. Bonarius HP, Hatzimanikatis V, Meesters KP, de Gooijer CD, Schmid G, Tramper J. Metabolic flux analysis of hybridoma cells in different culture media using mass balances. *Biotechnology and bioengineering*. 1996;50(3):299–318.
39. Edwards JS, Palsson BO. The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proceedings of the National Academy of Sciences*. 2000;97(10):5528–33.

40. Edwards JS, Palsson BO. Metabolic flux balance analysis and the in silico analysis of *Escherichia coli* K-12 gene deletions. *BMC bioinformatics*. 2000;1(1):1.
41. O'Brien EJ, Lerman JA, Chang RL, Hyduke DR, Palsson BØ. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Molecular systems biology*. 2013;9(1):693.
42. Stolýar S, Van Dien S, Hillesland KL, Pinel N, Lie TJ, Leigh JA, et al. Metabolic modeling of a mutualistic microbial community. *Molecular systems biology*. 2007;3(1):92.
43. Freilich S, Zarecki R, Eilam O, Segal ES, Henry CS, Kupiec M, et al. Competitive and cooperative metabolic interactions in bacterial communities. *Nature communications*. 2011;2(1):1–7.
44. Zomorodi AR, Maranas CD. OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Comput Biol*. 2012;8(2):e1002363.
45. Shoaie S, Karlsson F, Mardinoglu A, Nookaew I, Bordel S, Nielsen J. Understanding the interactions between bacteria in the human gut through metabolic modeling. *Scientific reports*. 2013;3:2532.
46. Diener C, Gibbons SM, Resendis-Antonio O. MICOM: metagenome-scale modeling to infer metabolic interactions in the gut microbiota. *Msystems*. 2020;5(1).
47. Majewski RA, Domach MM. Simple constrained-optimization view of acetate overflow in *E. coli*. *Biotechnology and bioengineering*. 1990;35(7):732–8.
48. Torres P, Saa PA, Albiol J, Ferrer P, Agosin E. Contextualized genome-scale model unveils high-order metabolic effects of the specific growth rate and oxygenation level in recombinant *Pichia pastoris*. *Metabolic engineering communications*. 2019;9:e00103.
49. Burgard AP, Vaidyaraman S, Maranas CD. Minimal reaction sets for *Escherichia coli* metabolism under different growth requirements and uptake environments. *Biotechnology progress*. 2001;17(5):791–7.
50. Puchałka J, Oberhardt MA, Godinho M, Bielecka A, Regenhardt D, Timmis KN, et al. Genome-Scale Reconstruction and Analysis of the *Pseudomonas putida* KT2440 Metabolic Network Facilitates Applications in Biotechnology. *PLoS Comput Biol*. 2008;4(10):e1000210.
51. Holzhütter H-G. The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. *European journal of biochemistry*. 2004;271(14):2905–22.
52. Blank LM, Kuepfer L, Sauer U. Large-scale ¹³C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome biology*. 2005;6(6):R49.

53. Pfau T, Christian N, Masakapalli SK, Sweetlove LJ, Poolman MG, Ebenhöf O. The intertwined metabolism during symbiotic nitrogen fixation elucidated by metabolic modelling. *Scientific reports*. 2018;8(1):1–11.
54. De Figueiredo LF, Podhorski A, Rubio A, Kaleta C, Beasley JE, Schuster S, et al. Computing the shortest elementary flux modes in genome-scale metabolic networks. *Bioinformatics*. 2009;25(23):3158–65.
55. Klitgord N, Segrè D. Environments that induce synthetic microbial ecosystems. *PLoS Comput Biol*. 2010;6(11):e1001002.
56. Oliveira AP, Nielsen J, Förster J. Modeling *Lactococcus lactis* using a genome-scale flux model. *BMC microbiology*. 2005;5(1):39.
57. Tomàs-Gamisans M, Ferrer P, Albiol J. Fine-tuning the *P. pastoris* iMT1026 genome-scale metabolic model for improved prediction of growth on methanol or glycerol as sole carbon sources. *Microbial biotechnology*. 2018;11(1):224–37.
58. Li C-T, Yelsky J, Chen Y, Zuñiga C, Eng R, Jiang L, et al. Utilizing genome-scale models to optimize nutrient supply for sustained algal growth and lipid productivity. *NPJ systems biology and applications*. 2019;5(1):1–11.
59. Koch S, Kohrs F, Lahmann P, Bissinger T, Wendschuh S, Benndorf D, et al. RedCom: A strategy for reduced metabolic modeling of complex microbial communities and its application for analyzing experimental datasets from anaerobic digestion. *PLoS computational biology*. 2019;15(2):e1006759.
60. Chen Y, McConnell BO, Dhara VG, Naik HM, Li C-T, Antoniewicz MR, et al. An unconventional uptake rate objective function approach enhances applicability of genome-scale models for mammalian cells. *NPJ systems biology and applications*. 2019;5(1):1–11.
61. Carvalho M, Nikdel A, Riesberg J, Lyons D, Budman H. Identification of a Dynamic Metabolic Flux Model for a Mammalian Cell Culture. *IFAC-PapersOnLine*. 2019;52(1):88–93.
62. Lewis JE, Costantini F, Mims J, Chen X, Furdui CM, Boothman DA, et al. Genome-scale modeling of NADPH-driven β -lapachone sensitization in head and neck squamous cell carcinoma. *Antioxidants & redox signaling*. 2018;29(10):937–52.
63. Lee J, Yun H, Feist AM, Palsson BØ, Lee SY. Genome-scale reconstruction and in silico analysis of the *Clostridium acetobutylicum* ATCC 824 metabolic network. *Applied microbiology and biotechnology*. 2008;80(5):849–62.
64. Succurro A, Segrè D, Ebenhöf O. Emergent subpopulation behavior uncovered with a community dynamic metabolic model of *Escherichia coli* diauxic growth. *Msystems*. 2019;4(1).

65. Rafieenia R, Pivato A, Schievano A, Lavagnolo MC. Dark fermentation metabolic models to study strategies for hydrogen consumers inhibition. *Bioresource technology*. 2018;267:445–57.
66. Oyetunde T, Liu D, Martin HG, Tang YJ. Machine learning framework for assessment of microbial factory performance. *PloS one*. 2019;14(1):e0210558.
67. Pramanik J, Keasling JD. Stoichiometric model of *Escherichia coli* metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnology and bioengineering*. 1997;56(4):398–421.
68. Castañeda MT, Nuñez S, Garelli F, Voget C, De Battista H. Comprehensive analysis of a metabolic model for lipid production in *Rhodospiridium toruloides*. *Journal of biotechnology*. 2018;280:11–8.
69. Bjerkelund Røkke G, Hohmann-Marriott MF, Almaas E. An adjustable algal chloroplast plug-and-play model for genome-scale metabolic models. *PloS one*. 2020;15(2):e0229408.
70. Suthers PF, Zomorodi A, Maranas CD. Genome-scale gene/reaction essentiality and synthetic lethality analysis. *Molecular systems biology*. 2009;5(1):301.
71. Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U. Multidimensional optimality of microbial metabolism. *Science*. 2012;336(6081):601–4.
72. Mori M, Marinari E, De Martino A. A yield-cost tradeoff governs *Escherichia coli*'s decision between fermentation and respiration in carbon-limited growth. *NPJ systems biology and applications*. 2019;5(1):1–9.
73. Dai Z, Yang S, Xu L, Hu H, Liao K, Wang J, et al. Identification of Cancer-associated metabolic vulnerabilities by modeling multi-objective optimality in metabolism. *Cell Communication and Signaling*. 2019;17(1):1–15.
74. Küken A, Eloundou-Mbebi JM, Basler G, Nikoloski Z. Cellular determinants of metabolite concentration ranges. *PLoS computational biology*. 2019;15(1):e1006687.
75. Grafahrend-Belau E, Schreiber F, Koschützki D, Junker BH. Flux balance analysis of barley seeds: a computational approach to study systemic properties of central metabolism. *Plant physiology*. 2009;149(1):585–98.
76. Alsiyabi A, Immethun CM, Saha R. Modeling the Interplay between Photosynthesis, CO₂ Fixation, and the Quinone Pool in a Purple Non-Sulfur Bacterium. *Scientific reports*. 2019;9(1):1–9.
77. Van Gulik WM, Heijnen JJ. A metabolic network stoichiometry analysis of microbial growth and product formation. *Biotechnology and bioengineering*. 1995;48(6):681–98.
78. Kuepfer L, Sauer U, Blank LM. Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*. *Genome research*. 2005;15(10):1421–30.

79. Scheunemann M, Brady SM, Nikoloski Z. Integration of large-scale data for extraction of integrated Arabidopsis root cell-type specific models. *Scientific reports*. 2018;8(1):1–15.
80. Pharkya P, Maranas CD. An optimization framework for identifying reaction activation/inhibition or elimination candidates for overproduction in microbial systems. *Metabolic engineering*. 2006;8(1):1–13.
81. Becker SA, Palsson BO. Context-specific metabolic networks are consistent with experiments. *PLoS Comput Biol*. 2008;4(5):e1000082.
82. Jensen PA, Papin JA. Functional integration of a metabolic network model and expression data without arbitrary thresholding. *Bioinformatics*. 2011;27(4):541–7.
83. Pandey V, Hadadi N, Hatzimanikatis V. Enhanced flux prediction by integrating relative expression and relative metabolite abundance into thermodynamically consistent metabolic models. *PLoS computational biology*. 2019;15(5):e1007036.
84. Liang X-L, Liang Z-M, Wang S, Chen X-H, Ruan Y, Zhang Q-Y, et al. An analysis of the mechanism underlying photocatalytic disinfection based on integrated metabolic networks and transcriptional data. *Journal of Environmental Sciences*. 2020;92:28–37.
85. Sarkar D, Maranas CD. SNPeffect: identifying functional roles of SNPs using metabolic networks. *The Plant Journal*. 2020;512–31.
86. Tian M, Reed JL. Integrating proteomic or transcriptomic data into metabolic models using linear bound flux balance analysis. *Bioinformatics*. 2018;34(22):3882–8.
87. Nikdel A, Braatz RD, Budman HM. A systematic approach for finding the objective function and active constraints for dynamic flux balance analysis. *Bioprocess and biosystems engineering*. 2018;41(5):641–55.
88. Hastings J, Mains A, Virk B, Rodriguez N, Murdoch S, Pearce J, et al. Multi-omics and genome-scale modeling reveal a metabolic shift during *C. elegans* aging. *Frontiers in molecular biosciences*. 2019;6:2.
89. Mo ML, Palsson BØ, Herrgård MJ. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC systems biology*. 2009;3(1):37.
90. Henry CS, Zinner JF, Cohoon MP, Stevens RL. iBsu1103: A new genome-scale metabolic model of *Bacillus subtilis* based on SEED annotations. *Genome biology*. 2009;10(6):R69.
91. Tepper N, Shlomi T. Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways. *Bioinformatics*. 2010;26(4):536–43.
92. Daud KM, Mohamad MS, Zakaria Z, Hassan R, Shah ZA, Deris S, et al. A non-dominated sorting Differential Search Algorithm Flux Balance Analysis (ndsDSAFBA) for in silico multiobjective optimization in identifying reactions knockout. *Computers in biology and medicine*. 2019;113:103390.

93. Choi HS, Lee SY, Kim TY, Woo HM. In silico identification of gene amplification targets for improvement of lycopene production. *Applied and environmental microbiology*. 2010;76(10):3097–105.
94. Ranganathan S, Suthers PF, Maranas CD. OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. *PLoS Comput Biol*. 2010;6(4):e1000744.
95. Alter TB, Ebert BE. Determination of growth-coupling strategies and their underlying principles. *BMC bioinformatics*. 2019;20(1):447.
96. Lasry Testa R, Delpino C, Estrada V, Diaz SM. In silico strategies to couple production of bioethanol with growth in cyanobacteria. *Biotechnology and bioengineering*. 2019;116(8):2061–73.
97. Pharkya P, Burgard AP, Maranas CD. OptStrain: a computational framework for redesign of microbial production systems. *Genome research*. 2004;14(11):2367–76.
98. Wu W, Long MR, Zhang X, Reed JL, Maravelias CT. A framework for the identification of promising bio-based chemicals. *Biotechnology and bioengineering*. 2018;115(9):2328–40.
99. Kim J, Reed JL. OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC systems biology*. 2010;4(1):53.
100. Lee SJ, Lee D-Y, Kim TY, Kim BH, Lee J, Lee SY. Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and in silico gene knockout simulation. *Applied and environmental microbiology*. 2005;71(12):7880–7.
101. Park JH, Lee KH, Kim TY, Lee SY. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proceedings of the national academy of sciences*. 2007;104(19):7797–802.
102. Jung YK, Kim TY, Park SJ, Lee SY. Metabolic engineering of *Escherichia coli* for the production of polylactic acid and its copolymers. *Biotechnology and bioengineering*. 2010;105(1):161–71.
103. Aslan S, Noor E, Vaquerizo SB, Lindner SN, Bar-Even A. Design and engineering of *E. coli* metabolic sensor strains with a wide sensitivity range for glycerate. *Metabolic Engineering*. 2020;57:96–109.
104. Koduru L, Lakshmanan M, Lee D-Y. In silico model-guided identification of transcriptional regulator targets for efficient strain design. *Microbial cell factories*. 2018;17(1):167.
105. Chemler JA, Fowler ZL, McHugh KP, Koffas MA. Improving NADPH availability for natural product biosynthesis in *Escherichia coli* by metabolic engineering. *Metabolic engineering*. 2010;12(2):96–104.

106. Wu W-H, Chien C-Y, Wu Y-H, Wu H-H, Lai J-M, Chang PM-H, et al. Inferring oncoenzymes in a genome-scale metabolic network for hepatocytes using bilevel optimization framework. *Journal of the Taiwan Institute of Chemical Engineers*. 2018;91:97–104.
107. Raškevičius V, Mikalayeva V, Antanavičiūtė I, Ceslevičienė I, Skeberdis VA, Kairys V, et al. Genome scale metabolic models as tools for drug design and personalized medicine. *PloS one*. 2018;13(1):e0190636.
108. Kumar VS, Dasika MS, Maranas CD. Optimization based automated curation of metabolic reconstructions. *BMC bioinformatics*. 2007;8(1):212.
109. Tervo CJ, Reed JL. FOCAL: an experimental design tool for systematizing metabolic discoveries and model development. *Genome biology*. 2012;13(12):R116.
110. Chowdhury A, Zomorodi AR, Maranas CD. Bilevel optimization techniques in computational strain design. *Computers & Chemical Engineering*. 2015;72:363–72.
111. Marler RT, Arora JS. The weighted sum method for multi-objective optimization: new insights. *Structural and multidisciplinary optimization*. 2010;41(6):853–62.
112. Haimes Y. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE transactions on systems, man, and cybernetics*. 1971;1(3):296–7.
113. Charnes A, Cooper WW. Goal programming and multiple objective optimizations: Part 1. *European journal of operational research*. 1977;1(1):39–54.
114. Pike-Burke C. Multi-objective optimization.
115. Marler RT, Arora JS. Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*. 2004;26(6):369–95.
116. Torres M, Jiang S, Pelta D, Kaiser M, Krasnogor N. Strain design as multiobjective network interdiction problem: a preliminary approach. In: *Conference of the Spanish Association for Artificial Intelligence*. Springer; 2018. p. 273–82.
117. Nikolaev EV, Burgard AP, Maranas CD. Elucidation and structural analysis of conserved pools for genome-scale metabolic reconstructions. *Biophysical journal*. 2005;88(1):37–49.
118. David L, Marashi S-A, Larhlimi A, Mieth B, Bockmayr A. FFCA: a feasibility-based method for flux coupling analysis of metabolic networks. *BMC bioinformatics*. 2011;12(1):236.
119. Larhlimi A, David L, Selbig J, Bockmayr A. F2C2: a fast tool for the computation of flux coupling in genome-scale metabolic networks. *BMC bioinformatics*. 2012;13(1):57.
120. Pál C, Papp B, Lercher MJ. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nature genetics*. 2005;37(12):1372–5.

121. Szappanos B, Kovács K, Szamecz B, Honti F, Costanzo M, Baryshnikova A, et al. An integrated approach to characterize genetic interaction networks in yeast metabolism. *Nature genetics*. 2011;43(7):656–62.
122. Notebaart RA, Kensche PR, Huynen MA, Dutilh BE. Asymmetric relationships between proteins shape genome evolution. *Genome biology*. 2009;10(2):R19.
123. Hosseini Z, Marashi S-A. Discovering missing reactions of metabolic networks by using gene co-expression data. *Scientific reports*. 2017;7(1):1–12.
124. Soh KC, Miskovic L, Hatzimanikatis V. From network models to network responses: integration of thermodynamic and kinetic properties of yeast genome-scale metabolic networks. *FEMS yeast research*. 2012;12(2):129–43.
125. Soh KC, Hatzimanikatis V. Constraining the flux space using thermodynamics and integration of metabolomics data. In: *Metabolic Flux Analysis*. Springer; 2014. p. 49–63.
126. Miskovic L, Tokic M, Fengos G, Hatzimanikatis V. Rites of passage: requirements and standards for building kinetic models of metabolic phenotypes. *Current opinion in biotechnology*. 2015;36:146–53.
127. Tymoshenko S, Oppenheim RD, Soldati-Favre D, Hatzimanikatis V. Functional genomics of *Plasmodium falciparum* using metabolic modelling and analysis. *Briefings in functional genomics*. 2013;12(4):316–27.
128. Yue D, Guillén-Gosálbez G, You F. Global optimization of large-scale mixed-integer linear fractional programming problems: a reformulation-linearization method and process scheduling applications. *AIChE Journal*. 2013;59(11):4255–72.
129. McCloskey D, Gangoiti JA, King ZA, Naviaux RK, Barshop BA, Palsson BO, et al. A model-driven quantitative metabolomics analysis of aerobic and anaerobic metabolism in *E. coli* K-12 MG1655 that is biochemically and thermodynamically consistent. *Biotechnology and bioengineering*. 2014;111(4):803–15.
130. Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, et al. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. *Molecular systems biology*. 2011;7(1):535.
131. Ataman M, Hernandez Gardiol DF, Fengos G, Hatzimanikatis V. redGEM: Systematic reduction and analysis of genome-scale metabolic reconstructions for development of consistent core metabolic models. *PLoS computational biology*. 2017;13(7):e1005444.
132. Ataman M, Hatzimanikatis V. lumpGEM: Systematic generation of subnetworks and elementally balanced lumped reactions for the biosynthesis of target metabolites. *PLoS computational biology*. 2017;13(7):e1005513.

133. Graewe S, Stanway RR, Rennenberg A, Heussler VT. Chronicle of a death foretold: Plasmodium liver stage parasites decide on the fate of the host cell. *FEMS microbiology reviews*. 2012;36(1):111–30.
134. Hulden L, Hulden L. Activation of the hypnozoite: a part of Plasmodium vivax life cycle and survival. *Malaria journal*. 2011;10(1):90.
135. Paredes CJ, Alsaker KV, Papoutsakis ET. A comparative genomic view of clostridial sporulation and physiology. *Nature Reviews Microbiology*. 2005;3(12):969–78.
136. Mancio-Silva L, Slavic K, Ruivo MTG, Grosso AR, Modrzynska KK, Vera IM, et al. Nutrient sensing modulates malaria parasite virulence. *Nature*. 2017;547(7662):213–6.
137. Gural N, Mancio-Silva L, Miller AB, Galstian A, Butty VL, Levine SS, et al. In vitro culture, drug sensitivity, and transcriptome of Plasmodium vivax hypnozoites. *Cell host & microbe*. 2018;23(3):395-406. e4.
138. Agren R, Liu L, Shoaie S, Vongsangnak W, Nookaew I, Nielsen J. The RAVEN toolbox and its use for generating a genome-scale metabolic model for Penicillium chrysogenum. *PLoS Comput Biol*. 2013;9(3):e1002980.
139. Wells TN, Burrows JN, Baird JK. Targeting the hypnozoite reservoir of Plasmodium vivax: the hidden obstacle to malaria elimination. *Trends in parasitology*. 2010;26(3):145–51.
140. Voorberg-van der Wel A, Roma G, Gupta DK, Schuierer S, Nigsch F, Carbone W, et al. A comparative transcriptomic analysis of replicating and dormant liver stages of the relapsing malaria parasite Plasmodium cynomolgi. *Elife*. 2017;6:e29605.
141. Saint-Léger A, Sinadinou C, Ribas de Pouplana L. The growing pipeline of natural aminoacyl-tRNA synthetase inhibitors for malaria treatment. *Bioengineered*. 2016;7(2):60–4.
142. Mailu BM, Li L, Arthur J, Nelson TM, Ramasamy G, Fritz-Wolf K, et al. Plasmodium apicoplast Gln-tRNA^{Gln} biosynthesis utilizes a unique GatAB amidotransferase essential for erythrocytic stage parasites. *Journal of Biological Chemistry*. 2015;290(49):29629–41.
143. Yoo M, Bestel-Corre G, Croux C, Riviere A, Meynial-Salles I, Soucaille P. A quantitative system-scale characterization of the metabolism of Clostridium acetobutylicum. *MBio*. 2015;6(6).
144. Jones DT, Van der Westhuizen A, Long S, Allcock ER, Reid SJ, Woods DR. Solvent production and morphological changes in Clostridium acetobutylicum. *Applied and Environmental Microbiology*. 1982;43(6):1434–9.
145. Jones DT, Woods DR. Acetone-butanol fermentation revisited. *Microbiological reviews*. 1986;50(4):484.

146. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *Journal of microbiological methods*. 2007;70(3):452–64.
147. Shao L, Hu S, Yang Y, Gu Y, Chen J, Yang Y, et al. Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*. *Cell research*. 2007;17(11):963–5.
148. List C, Hosseini Z, Lederballe Meibom K, Hatzimanikatis V, Bernier-Latmani R. Impact of iron reduction on the metabolism of *Clostridium acetobutylicum*. *Environmental microbiology*. 2019;21(10):3548–63.
149. Pollock J, Weber KA, Lack J, Achenbach LA, Mormile MR, Coates JD. Alkaline iron (III) reduction by a novel alkaliphilic, halotolerant, *Bacillus* sp. isolated from salt flat sediments of Soap Lake. *Applied microbiology and biotechnology*. 2007;77(4):927–34.
150. Dong Y, Sanford RA, Chang Y, McInerney MJ, Fouke BW. Hematite reduction buffers acid generation and enhances nutrient uptake by a fermentative iron reducing bacterium, *Orenia metallireducens* strain Z6. *Environmental science & technology*. 2017;51(1):232–42.
151. Lovley DR, Holmes DE, Nevin KP. Dissimilatory Fe (iii) and Mn (iv) reduction. *Advances in microbial physiology*. 2004;49(2):219–86.
152. Lehours A-C, Rabiet M, Morel-Desrosiers N, Morel J-P, Jouve L, Arbeille B, et al. Ferric iron reduction by fermentative strain BS2 isolated from an iron-rich anoxic environment (Lake Pavin, France). *Geomicrobiology Journal*. 2010;27(8):714–22.
153. Dalla Vecchia E, Suvorova EI, Maillard J, Bernier-Latmani R. Fe (III) reduction during pyruvate fermentation by *D. esultotomaculum* *reducens* strain MI-1. *Geobiology*. 2014;12(1):48–61.
154. Senger RS, Papoutsakis ET. Genome-scale model for *Clostridium acetobutylicum*: Part I. Metabolic network resolution and analysis. *Biotechnology and bioengineering*. 2008;101(5):1036–52.
155. McAnulty MJ, Yen JY, Freedman BG, Senger RS. Genome-scale modeling using flux ratio constraints to enable metabolic engineering of clostridial metabolism in silico. *BMC systems biology*. 2012;6(1):42.
156. Mahadevan R, Schilling CH. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic engineering*. 2003;5(4):264–76.
157. Wiback SJ, Famili I, Greenberg HJ, Palsson BØ. Monte Carlo sampling can be used to determine the size and shape of the steady-state flux space. *Journal of theoretical biology*. 2004;228(4):437–47.
158. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? *Nature biotechnology*. 2010;28(3):245–8.

159. Covert MW, Xiao N, Chen TJ, Karr JR. Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics*. 2008;24(18):2044–50.
160. Shlomi T, Eisenberg Y, Sharan R, Ruppin E. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Molecular systems biology*. 2007;3(1):101.
161. Lu H, Cao W, Liu X, Sui Y, Ouyang L, Xia J, et al. Multi-omics integrative analysis with genome-scale metabolic model simulation reveals global cellular adaptation of *Aspergillus niger* under industrial enzyme production condition. *Scientific reports*. 2018;8(1):1–15.
162. Botero D, Valdés I, Rodríguez M-J, Henao D, Danies G, González AF, et al. A genome-scale metabolic reconstruction of *Phytophthora infestans* with the integration of transcriptional data reveals the key metabolic patterns involved in the interaction of its host. *Frontiers in genetics*. 2018;9:244.
163. Covert MW, Schilling CH, Palsson B. Regulation of gene expression in flux balance models of metabolism. *Journal of theoretical biology*. 2001;213(1):73–88.
164. Zhuang K, Vemuri GN, Mahadevan R. Economics of membrane occupancy and respiration. *Molecular systems biology*. 2011;7(1):500.
165. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. *Science*. 2017;357(6352).
166. Åkesson M, Förster J, Nielsen J. Integration of gene expression data into genome-scale metabolic models. *Metabolic engineering*. 2004;6(4):285–93.
167. Covert MW, Palsson BO. Constraints-based models: regulation of gene expression reduces the steady-state solution space. *Journal of theoretical biology*. 2003;221(3):309–25.
168. Khodae S, Asgari Y, Totonchi M, Karimi-Jafari MH. iMM1865: A New Reconstruction of Mouse Genome-Scale Metabolic Model. *Scientific reports*. 2020;10(1):1–13.
169. Yeo HC, Hong J, Lakshmanan M, Lee D-Y. Enzyme capacity-based genome scale modelling of CHO cells. *Metabolic Engineering*. 2020;138–47.
170. Tsouka S, Hatzimanikatis V. redLips: a comprehensive mechanistic model of the lipid metabolic network of yeast. *FEMS Yeast Research*. 2020;20(2):foaa006.
171. Molina L, La Rosa R, Nogales J, Rojo F. Influence of the Crc global regulator on substrate uptake rates and the distribution of metabolic fluxes in *Pseudomonas putida* KT2440 growing in a complete medium. *Environmental microbiology*. 2019;21(11):4446–59.
172. Marcišauskas S, Ji B, Nielsen J. Reconstruction and analysis of a *Kluyveromyces marxianus* genome-scale metabolic model. *BMC bioinformatics*. 2019;20(1):551.

173. Yousefi M, Marashi S-A, Sharifi-Zarchi A, Taleahmad S. The metabolic network model of primed/naive human embryonic stem cells underlines the importance of oxidation-reduction potential and tryptophan metabolism in primed pluripotency. *Cell & Bioscience*. 2019;9(1):71.
174. Covert MW, Palsson BØ. Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *Journal of Biological Chemistry*. 2002;277(31):28058–64.
175. Botero K, Restrepo S, Pinzón A. A genome-scale metabolic model of potato late blight suggests a photosynthesis suppression mechanism. *BMC genomics*. 2018;19(8):863.
176. Persi E, Duran-Frigola M, Damaghi M, Roush WR, Aloy P, Cleveland JL, et al. Systems analysis of intracellular pH vulnerabilities for cancer therapy. *Nature communications*. 2018;9(1):1–11.
177. Varma A, Boesch BW, Palsson BO. Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. *Applied and environmental microbiology*. 1993;59(8):2465–73.
178. Varma A, Palsson BO. Metabolic capabilities of *Escherichia coli* II. Optimal growth patterns. *Journal of Theoretical Biology*. 1993;165(4):503–22.
179. Reed JL, Palsson BØ. Genome-scale in silico models of *E. coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. *Genome research*. 2004;14(9):1797–805.
180. Zhu Y, Lu J, Zhao J, Zhang X, Heidi HY, Velkov T, et al. Complete genome sequence and genome-scale metabolic modelling of *Acinetobacter baumannii* type strain ATCC 19606. *International Journal of Medical Microbiology*. 2020;151412.
181. Lu H, Li F, Sánchez BJ, Zhu Z, Li G, Domenzain I, et al. A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. *Nature communications*. 2019;10(1):1–13.
182. Shaw R, Cheung CY. A dynamic multi-tissue flux balance model captures carbon and nitrogen metabolism and optimal resource partitioning during *Arabidopsis* growth. *Frontiers in plant science*. 2018;9:884.
183. Bordel S. Constraint based modeling of metabolism allows finding metabolic cancer hallmarks and identifying personalized therapeutic windows. *Oncotarget*. 2018;9(28):19716.
184. Agren R, Mardinoglu A, Asplund A, Kampf C, Uhlen M, Nielsen J. Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Molecular systems biology*. 2014;10(3):721.
185. Toro L, Pinilla L, Avignone-Rossa C, Ríos-Esteva R. An enhanced genome-scale metabolic reconstruction of *Streptomyces clavuligerus* identifies novel strain improvement strategies. *Bioprocess and biosystems engineering*. 2018;41(5):657–69.

186. Yu H, Blair RH. Integration of probabilistic regulatory networks into constraint-based models of metabolism with applications to Alzheimer's disease. *BMC bioinformatics*. 2019;20(1):386.
187. Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles ED. Metabolic network structure determines key aspects of functionality and regulation. *Nature*. 2002;420(6912):190–3.
188. Teusink B, Wiersma A, Molenaar D, Francke C, de Vos WM, Siezen RJ, et al. Analysis of growth of *Lactobacillus plantarum* WCFS1 on a complex medium using a genome-scale metabolic model. *Journal of Biological Chemistry*. 2006;281(52):40041–8.
189. Bordel S, Rojas A, Muñoz R. Reconstruction of a Genome Scale Metabolic Model of the polyhydroxybutyrate producing methanotroph *Methylocystis parvus* OBBP. *Microbial cell factories*. 2019;18(1):1–11.
190. Mazat J-P, Ransac S. The Fate of Glutamine in Human Metabolism. The interplay with glucose in proliferating cells. *Metabolites*. 2019;9(5):81.
191. Li F, Xie W, Yuan Q, Luo H, Li P, Chen T, et al. Genome-scale metabolic model analysis indicates low energy production efficiency in marine ammonia-oxidizing archaea. *AMB Express*. 2018;8(1):106.
192. Jóhannsson F, Guðmundsson S, Paglia G, Guðmundsson S, Palsson B, Sigurjónsson ÓE, et al. Systems analysis of metabolism in platelet concentrates during storage in platelet additive solution. *Biochemical Journal*. 2018;475(13):2225–40.
193. Sinha N, Suarez-Diez M, Hooiveld GJ, Keijer J, Martin dos Santos V, van Schothorst EM. A constraint-based model analysis of enterocyte mitochondrial adaptation to dietary interventions of lipid type and lipid load. *Frontiers in physiology*. 2018;9:749.
194. Valgepea K, Lemgruber R de SP, Abdalla T, Binos S, Takemori N, Takemori A, et al. H₂ drives metabolic rearrangements in gas-fermenting *Clostridium autoethanogenum*. *Biotechnology for biofuels*. 2018;11(1):1–15.
195. Abdel-Haleem AM, Hefzi H, Mineta K, Gao X, Gojobori T, Palsson BO, et al. Functional interrogation of *Plasmodium* genus metabolism identifies species- and stage-specific differences in nutrient essentiality and drug targeting. *PLoS computational biology*. 2018;14(1):e1005895.
196. Olavarria K, Fina A, Velasco MI, van Loosdrecht MC, Wahl SA. Metabolism of sucrose in a non-fermentative *Escherichia coli* under oxygen limitation. *Applied microbiology and biotechnology*. 2019;103(15):6245–56.
197. Pacheco MP, Bintener T, Ternes D, Kulms D, Haan S, Letellier E, et al. Identifying and targeting cancer-specific metabolism with network-based drug target prediction. *EBioMedicine*. 2019;43:98–106.
198. Massaiu I, Pasotti L, Sonnenschein N, Rama E, Cavaletti M, Magni P, et al. Integration of enzymatic data in *Bacillus subtilis* genome-scale metabolic model improves phenotype

- predictions and enables in silico design of poly- γ -glutamic acid production strains. *Microbial Cell Factories*. 2019;18(1):1–20.
199. Wegrzyn AB, Stolle S, Rienksma RA, dos Santos VAM, Bakker BM, Suarez-Diez M. Cofactors revisited–Predicting the impact of flavoprotein-related diseases on a genome scale. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*. 2019;1865(2):360–70.
 200. Katzir R, Polat IH, Harel M, Katz S, Foguet C, Selivanov VA, et al. The landscape of tiered regulation of breast cancer cell metabolism. *Scientific reports*. 2019;9(1):1–12.
 201. de Mas IM, Torrents L, Bedia C, Nielsen LK, Cascante M, Tauler R. Stoichiometric gene-to-reaction associations enhance model-driven analysis performance: Metabolic response to chronic exposure to Aldrin in prostate cancer. *BMC genomics*. 2019;20(1):1–12.
 202. Yan J, Estanbouli H, Liao C, Kim W, Monk JM, Rahman R, et al. Systems-level analysis of NalD mutation, a recurrent driver of rapid drug resistance in acute *Pseudomonas aeruginosa* infection. *PLoS computational biology*. 2019;15(12):e1007562.
 203. Siriwach R, Matsuda F, Yano K, Hirai MY. Drought stress responses in context-specific genome-scale metabolic models of *Arabidopsis thaliana*. *Metabolites*. 2020;10(4):159.
 204. D'Arrigo I, Cardoso JG, Rennig M, Sonnenschein N, Herrgård MJ, Long KS. Analysis of *Pseudomonas putida* growth on non-trivial carbon sources using transcriptomics and genome-scale modelling. *Environmental microbiology reports*. 2019;11(2):87–97.
 205. Huang Z, Yoon S. Integration of Time-Series Transcriptomic Data with Genome-Scale CHO Metabolic Models for mAb Engineering. *Processes*. 2020;8(3):331.
 206. Baloni P, Sangar V, Yurkovich JT, Robinson M, Taylor S, Karbowski CM, et al. Genome-scale metabolic model of the rat liver predicts effects of diet restriction. *Scientific reports*. 2019;9(1):1–12.
 207. Lun DS, Rockwell G, Guido NJ, Baym M, Kelner JA, Berger B, et al. Large-scale identification of genetic design strategies using local search. *molecular systems biology*. 2009;5(1):296.
 208. Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, et al. Metabolic engineering of *Escherichia coli* for direct production of 1, 4-butanediol. *Nature chemical biology*. 2011;7(7):445–52.
 209. Fong SS, Burgard AP, Herring CD, Knight EM, Blattner FR, Maranas CD, et al. In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnology and bioengineering*. 2005;91(5):643–8.
 210. Aminian-Dehkordi J, Mousavi SM, Marashi S-A, Jafari A, Mijakovic I. A Systems-Based Approach for Cyanide Overproduction by *Bacillus megaterium* for Gold Bioleaching Enhancement. *Frontiers in Bioengineering and Biotechnology*. 2020;8:528.

211. Kumelj T, Sulheim S, Wentzel A, Almaas E. Predicting Strain Engineering Strategies Using iKS1317: A Genome-Scale Metabolic Model of *Streptomyces coelicolor*. *Biotechnology journal*. 2019;14(4):1800180.
212. Bodor Z, Lanyi S, Albert B, Bodor K, Nechifor AC, Miklossy I. Model Driven Analysis of the Biosynthesis of 1, 4-butanediol from Renewable Feedstocks in *Escherichia coli*. *chemistry*. 2019;1:13.
213. Navone L, McCubbin T, Gonzalez-Garcia RA, Nielsen LK, Marcellin E. Genome-scale model guided design of *Propionibacterium* for enhanced propionic acid production. *Metabolic engineering communications*. 2018;6:1–12.
214. Du W, Jongbloets JA, van Bortel C, Hernández HP, Lips D, Oliver BG, et al. Alignment of microbial fitness with engineered product formation: obligatory coupling between acetate production and photoautotrophic growth. *Biotechnology for biofuels*. 2018;11(1):38.
215. Winternute EH, Silver PA. Emergent cooperation in microbial metabolism. *Molecular systems biology*. 2010;6(1):407.
216. Cordes H, Thiel C, Baier V, Blank LM, Kuepfer L. Integration of genome-scale metabolic networks into whole-body PBPK models shows phenotype-specific cases of drug-induced metabolic perturbation. *NPJ systems biology and applications*. 2018;4(1):1–11.
217. Selvarasu S, Ho YS, Chong WP, Wong NS, Yusufi FN, Lee YY, et al. Combined in silico modeling and metabolomics analysis to characterize fed-batch CHO cell culture. *Biotechnology and bioengineering*. 2012;109(6):1415–29.
218. Tong H, Küken A, Nikoloski Z. Integrating molecular markers into metabolic models improves genomic selection for *Arabidopsis* growth. *Nature communications*. 2020;11(1):1–9.
219. Tibocha-Bonilla JD, Kumar M, Richelle A, Godoy-Silva RD, Zengler K, Zuñiga C. Dynamic resource allocation drives growth under nitrogen starvation in eukaryotes. *NPJ systems biology and applications*. 2020;6(1):1–9.
220. Mintz-Oron S, Meir S, Malitsky S, Ruppin E, Aharoni A, Shlomi T. Reconstruction of *Arabidopsis* metabolic network models accounting for subcellular compartmentalization and tissue-specificity. *Proceedings of the National Academy of Sciences*. 2012;109(1):339–44.
221. Papp B, Pál C, Hurst LD. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature*. 2004;429(6992):661–4.
222. Costa RS, Vinga S. Assessing *Escherichia coli* metabolism models and simulation approaches in phenotype predictions: Validation against experimental data. *Biotechnology progress*. 2018;34(6):1344–54.

223. Ma Z, Ye C, Deng W, Xu M, Wang Q, Liu G, et al. Reconstruction and analysis of a genome-scale metabolic model of *Ganoderma lucidum* for improved extracellular polysaccharide production. *Frontiers in microbiology*. 2018;9:3076.
224. Alper H, Jin Y-S, Moxley JF, Stephanopoulos G. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metabolic engineering*. 2005;7(3):155–64.
225. Sundararaghavan A, Mukherjee A, Sahoo S, Suraishkumar GK. Mechanism of the oxidative stress-mediated increase in lipid accumulation by the bacterium, *R. opacus* PD630: Experimental analysis and genome-scale metabolic modeling. *Biotechnology and Bioengineering*. 2020;117(6):1779–88.
226. Mienda BS. *Escherichia coli* genome-scale metabolic gene knockout of lactate dehydrogenase (ldhA), increases succinate production from glycerol. *Journal of Biomolecular Structure and Dynamics*. 2018;36(14):3680–6.
227. Chandrasekaran S, Price ND. Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences*. 2010;107(41):17845–50.
228. Fan J, Kamphorst JJ, Mathew R, Chung MK, White E, Shlomi T, et al. Glutamine-driven oxidative phosphorylation is a major ATP source in transformed mammalian cells in both normoxia and hypoxia. *Molecular systems biology*. 2013;9(1):712.
229. Mardinoglu A, Shoaie S, Bergentall M, Ghaffari P, Zhang C, Larsson E, et al. The gut microbiota modulates host amino acid and glutathione metabolism in mice. *Molecular systems biology*. 2015;11(10):834.
230. Pannala VR, Estes SK, Rahim M, Trenary I, O'Brien TP, Shiota C, et al. Mechanism-based identification of plasma metabolites associated with liver toxicity. *Toxicology*. 2020;152493.
231. Blank LM, Kuepfer L, Sauer U. Large-scale ¹³C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome biology*. 2005;6(6):R49.
232. Calmels C, Arnoult S, Yahia BB, Malphettes L, Andersen MR. Application of a genome-scale model in tandem with enzyme assays for identification of metabolic signatures of high and low CHO cell producers. *Metabolic engineering communications*. 2019;9:e00097.
233. Moreira TB, Shaw R, Luo X, Ganguly O, Kim H-S, Coelho LGF, et al. A genome-scale metabolic model of soybean (*Glycine max*) highlights metabolic fluxes in seedlings. *Plant physiology*. 2019;180(4):1912–29.

234. Ahmad A, Pathania R, Srivastava S. Biochemical Characteristics and a Genome-Scale Metabolic Model of an Indian Euryhaline Cyanobacterium with High Polyglucan Content. *Metabolites*. 2020;10(5):177.
235. Rodenburg SY, Seidl MF, Judelson HS, Vu AL, Govers F, de Ridder D. Metabolic model of the *Phytophthora infestans*-tomato interaction reveals metabolic switches during host colonization. *Mbio*. 2019;10(4):e00454-19.
236. Brochado AR, Matos C, Møller BL, Hansen J, Mortensen UH, Patil KR. Improved vanillin production in baker's yeast through in silico design. *Microbial cell factories*. 2010;9(1):84.
237. Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, Farhat MR, et al. Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production. *PLoS Comput Biol*. 2009;5(8):e1000489.
238. Bordbar A, Lewis NE, Schellenberger J, Palsson BØ, Jamshidi N. Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Molecular systems biology*. 2010;6(1):422.
239. Roell GW, Carr RR, Campbell T, Shang Z, Henson WR, Czajka JJ, et al. A concerted systems biology analysis of phenol metabolism in *Rhodococcus opacus* PD630. *Metabolic engineering*. 2019;55:120–30.
240. Borodina I, Krabben P, Nielsen J. Genome-scale analysis of *Streptomyces coelicolor* A3 (2) metabolism. *Genome research*. 2005;15(6):820–9.
241. Andersen MR, Nielsen ML, Nielsen J. Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. *Molecular Systems Biology*. 2008;4(1):178.
242. Nouri H, Fouladiha H, Moghimi H, Marashi S-A. A reconciliation of genome-scale metabolic network model of *Zymomonas mobilis* ZM4. *Scientific Reports*. 2020;10(1):1–11.
243. Campos DT, Zuñiga C, Passi A, Del Toro J, Tibocha-Bonilla JD, Zepeda A, et al. Modeling of nitrogen fixation and polymer production in the heterotrophic diazotroph *Azotobacter vinelandii* DJ. *Metabolic Engineering Communications*. 2020;e00132.
244. Bordel S, Rodríguez Y, Hakobyan A, Rodríguez E, Lebrero R, Muñoz R. Genome scale metabolic modeling reveals the metabolic potential of three Type II methanotrophs of the genus *Methylocystis*. *Metabolic engineering*. 2019;54:191–9.
245. Mekanik M, Motamedian E, Fotovat R, Jafarian V. Reconstruction of a genome-scale metabolic model for *Auxenochlorella protothecoides* to study hydrogen production under anaerobiosis using multiple optimal solutions. *International Journal of Hydrogen Energy*. 2019;44(5):2580–91.

246. Prigent S, Nielsen JC, Frisvad JC, Nielsen J. Reconstruction of 24 *Penicillium* genome-scale metabolic models shows diversity based on their secondary metabolism. *Biotechnology and Bioengineering*. 2018;115(10):2604–12.
247. Kumar M, Ji B, Babaei P, Das P, Lappa D, Ramakrishnan G, et al. Gut microbiota dysbiosis is associated with malnutrition and reduced plasma amino acid levels: lessons from genome-scale metabolic modeling. *Metabolic engineering*. 2018;49:128–42.
248. Souza SS de, Castro J de V, Porto LM. Modeling the core metabolism of *Komagataeibacter hansenii* ATCC 23769 to evaluate nanocellulose biosynthesis. *Brazilian Journal of Chemical Engineering*. 2018;35(3):869–86.
249. Yaşar Yildiz S, Nikerel E, Toksoy Öner E. Genome-Scale Metabolic Model of a Microbial Cell Factory (*Brevibacillus thermoruber* 423) with Multi-Industry Potentials for Exopolysaccharide Production. *OMICS: A Journal of Integrative Biology*. 2019;23(4):237–46.
250. Parichehreh R, Gheshlaghi R, Mahdavi MA, Elkamel A. Optimization of lipid production in *Chlorella vulgaris* for biodiesel production using flux balance analysis. *Biochemical Engineering Journal*. 2019;141:131–45.
251. Kristjansdottir T, Bosma EF, Dos Santos FB, Özdemir E, Herrgård MJ, França L, et al. A metabolic reconstruction of *Lactobacillus reuteri* JCM 1112 and analysis of its potential as a cell factory. *Microbial cell factories*. 2019;18(1):186.
252. Bauer E, Thiele I. From metagenomic data to personalized in silico microbiotas: predicting dietary supplements for Crohn's disease. *NPJ systems biology and applications*. 2018;4(1):1–9.
253. Klanchui A, Dulsawat S, Chaloeamngam K, Cheevadhanarak S, Prommeenate P, Meechai A. An Improved Genome-Scale Metabolic Model of *Arthrospira platensis* C1 (iAK888) and Its Application in Glycogen Overproduction. *Metabolites*. 2018;8(4):84.
254. Banerjee D, Raghunathan A. Constraints-based analysis identifies NAD⁺ recycling through metabolic reprogramming in antibiotic resistant *Chromobacterium violaceum*. *PloS one*. 2019;14(1):e0210008.
255. Hon MK, Mohamad MS, Salleh AHM, Choon YW, Daud KM, Remli MA, et al. Identifying a gene knockout strategy using a hybrid of simple constrained artificial bee colony algorithm and flux balance analysis to enhance the production of succinate and lactate in *Escherichia coli*. *Interdisciplinary Sciences: Computational Life Sciences*. 2019;11(1):33–44.
256. Lee KH, Park JH, Kim TY, Kim HU, Lee SY. Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Molecular systems biology*. 2007;3(1):149.
257. Koivuranta K, Castillo S, Jouhten P, Ruohonen L, Penttilä M, Wiebe MG. Enhanced triacylglycerol production with genetically modified *Trichosporon oleaginosus*. *Frontiers in microbiology*. 2018;9:1337.

258. Averagesch NJ, Martínez VS, Nielsen LK, Krömer JO. Toward synthetic biology strategies for adipic acid production: an in silico tool for combined thermodynamics and stoichiometric analysis of metabolic networks. *ACS synthetic biology*. 2018;7(2):490–509.
259. Li J, Sun R, Ning X, Wang X, Wang Z. Genome-scale metabolic model of *Actinosynnema pretiosum* ATCC 31280 and its application for ansamitocin P-3 production improvement. *Genes*. 2018;9(7):364.
260. Occhipinti A, Eyassu F, Rahman TJ, Rahman PK, Angione C. In silico engineering of *Pseudomonas* metabolism reveals new biomarkers for increased biosurfactant production. *PeerJ*. 2018;6:e6046.
261. Wang F-S, Wu W-H, Hsiu W-S, Liu Y-J, Chuang K-W. Genome-Scale Metabolic Modeling with Protein Expressions of Normal and Cancerous Colorectal Tissues for Oncogene Inference. *Metabolites*. 2020;10(1):16.
262. Vilkhovoy M, Horvath N, Shih C-H, Wayman JA, Calhoun K, Swartz J, et al. Sequence specific modeling of *E. coli* cell-free protein synthesis. *ACS synthetic biology*. 2018;7(8):1844–57.
263. Dai D, Horvath N, Varner J. Dynamic sequence specific constraint-based modeling of cell-free protein synthesis. *Processes*. 2018;6(8):132.
264. Serrano-Bermúdez LM, González Barrios AF, Montoya D. *Clostridium butyricum* population balance model: Predicting dynamic metabolic flux distributions using an objective function related to extracellular glycerol content. *PloS one*. 2018;13(12):e0209447.
265. Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature biotechnology*. 2010;28(9):977–82.
266. Reed JL, Patel TR, Chen KH, Joyce AR, Applebee MK, Herring CD, et al. Systems approach to refining genome annotation. *Proceedings of the National Academy of Sciences*. 2006;103(46):17480–4.
267. Machado D, Andrejev S, Tramontano M, Patil KR. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic acids research*. 2018;46(15):7542–53.
268. Tramontano M, Andrejev S, Pruteanu M, Klünemann M, Kuhn M, Galardini M, et al. Nutritional preferences of human gut bacteria reveal their metabolic idiosyncrasies. *Nature microbiology*. 2018;3(4):514–22.
269. Xu P, Ranganathan S, Fowler ZL, Maranas CD, Koffas MA. Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. *Metabolic engineering*. 2011;13(5):578–87.

270. Cheng F, Yu H, Stephanopoulos G. Engineering *Corynebacterium glutamicum* for high-titer biosynthesis of hyaluronic acid. *Metabolic engineering*. 2019;55:276–89.
271. Gómez-Cerón S, Galindo-Betancur D, Ramírez-Malule H. Data set of in silico simulation for the production of clavulanic acid and cephamycin C by *Streptomyces clavuligerus* using a genome scale metabolic model. *Data in brief*. 2019;24:103992.
272. Seif Y, Kavvas E, Lachance J-C, Yurkovich JT, Nuccio S-P, Fang X, et al. Genome-scale metabolic reconstructions of multiple *Salmonella* strains reveal serovar-specific metabolic traits. *Nature communications*. 2018;9(1):1–12.
273. Nogales J, Palsson BØ, Thiele I. A genome-scale metabolic reconstruction of *Pseudomonas putida* KT2440: i JN746 as a cell factory. *BMC systems biology*. 2008;2(1):79.
274. Sheikh K, Förster J, Nielsen LK. Modeling hybridoma cell metabolism using a generic genome-scale metabolic model of *Mus musculus*. *Biotechnology progress*. 2005;21(1):112–21.
275. Contador CA, Lo S-K, Chan SH, Lam H-M. Metabolic Analyses of Nitrogen Fixation in the Soybean Microsymbiont *Sinorhizobium fredii* Using Constraint-Based Modeling. *Msystems*. 2020;5(1).
276. Zou W, Xiong X, Zhang J, Zhang K, Zhao X, Zhao C. Reconstruction and analysis of a genome-scale metabolic model of *Methylovorus* sp. MP688, a high-level pyrroloquinolone quinone producer. *Biosystems*. 2018;172:37–42.
277. Widiastuti H, Lee N-R, Karimi IA, Lee D-Y. Genome-scale in silico analysis for enhanced production of succinic acid in *Zymomonas mobilis*. *Processes*. 2018;6(4):30.
278. Bekiaris PS, Klamt S. Automatic construction of metabolic models with enzyme constraints. *BMC bioinformatics*. 2020;21(1):1–13.
279. Kjeldsen KR, Nielsen J. In silico genome-scale reconstruction and validation of the *Corynebacterium glutamicum* metabolic network. *Biotechnology and bioengineering*. 2009;102(2):583–97.
280. Koduru L, Kim HY, Lakshmanan M, Mohanty B, Lee YQ, Lee CH, et al. Genome-scale metabolic reconstruction and in silico analysis of the rice leaf blight pathogen, *Xanthomonas oryzae*. *Molecular Plant Pathology*. 2020;21(4):527–40.
281. Aminian-Dehkordi J, Mousavi SM, Jafari A, Mijakovic I, Marashi S-A. Manually curated genome-scale reconstruction of the metabolic network of *Bacillus megaterium* DSM319. *Scientific reports*. 2019;9(1):1–14.
282. Feng J, Yang J, Yang W, Chen J, Jiang M, Zou X. Metabolome-and genome-scale model analyses for engineering of *Aureobasidium pullulans* to enhance polymalic acid and malic acid production from sugarcane molasses. *Biotechnology for biofuels*. 2018;11(1):94.

283. Tzamali E, Tzedakis G, Sakkalis V. A framework linking glycolytic metabolic capabilities and tumor dynamics. *IEEE journal of biomedical and health informatics*. 2019;23(5):1844–54.
284. Raethong N, Wang H, Nielsen J, Vongsangnak W. Optimizing cultivation of *Cordyceps militaris* for fast growth and cordycepin overproduction using rational design of synthetic media. *Computational and structural biotechnology journal*. 2020;18:1–8.
285. Rezazadeh M, Babaeipour V, Motamedian E. Reconstruction, verification and in-silico analysis of a genome-scale metabolic model of bacterial cellulose producing *Komagataeibacter xylinus*. *Bioprocess and Biosystems Engineering*. 2020;1–10.
286. Kuriya Y, Araki M. Dynamic Flux Balance Analysis to Evaluate the Strain Production Performance on Shikimic Acid Production in *Escherichia coli*. *Metabolites*. 2020;10(5):198.
287. Ding S, Cai P, Yuan L, Tian Y, Tu W, Zhang D, et al. CF-targeter: a rational biological cell factory targeting platform for biosynthetic target chemicals. *ACS Synthetic Biology*. 2019;8(10):2280–6.
288. Beck AE. Metabolic Efficiency of Sugar Co-Metabolism and Phenol Degradation in *Alicyclobacillus acidocaldarius* for Improved Lignocellulose Processing. *Processes*. 2020;8(5):502.
289. Fell DA, Small JR. Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochemical Journal*. 1986;238(3):781–6.
290. Wang Y, Eddy JA, Price ND. Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. *BMC systems biology*. 2012;6(1):153.
291. Mellbye BL, Giguere AT, Murthy GS, Bottomley PJ, Sayavedra-Soto LA, Chaplen FW. Genome-scale, constraint-based modeling of nitrogen oxide fluxes during coculture of *Nitrosomonas europaea* and *Nitrobacter winogradskyi*. *Msystems*. 2018;3(3).
292. Hoppe A, Hoffmann S, Holzhütter H-G. Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks. *BMC systems biology*. 2007;1(1):23.
293. Zampieri M, Hörl M, Hotz F, Müller NF, Sauer U. Regulatory mechanisms underlying coordination of amino acid and glucose catabolism in *Escherichia coli*. *Nature communications*. 2019;10(1):1–13.
294. Rienksma RA, Schaap PJ, Martins dos Santos VA, Suarez-Diez M. Modeling the metabolic state of *Mycobacterium tuberculosis* upon infection. *Frontiers in cellular and infection microbiology*. 2018;8:264.
295. Zhuangrong H, Seongkyu Y. Identifying metabolic features and engineering targets for productivity improvement in CHO cells by integrated transcriptomics and genome-scale metabolic model. *Biochemical Engineering Journal*. 2020;107624.

296. Grafahrend-Belau E, Schreiber F, Koschützki D, Junker BH. Flux balance analysis of barley seeds: a computational approach to study systemic properties of central metabolism. *Plant physiology*. 2009;149(1):585–98.
297. Magnúsdóttir S, Heinken A, Kutt L, Ravcheev DA, Bauer E, Noronha A, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nature biotechnology*. 2017;35(1):81.
298. Poolman MG, Miguët L, Sweetlove LJ, Fell DA. A genome-scale metabolic model of *Arabidopsis* and some of its properties. *Plant physiology*. 2009;151(3):1570–81.
299. Naizabekov S, Lee EY. Genome-Scale Metabolic Model Reconstruction and in Silico Investigations of Methane Metabolism in *Methylosinus trichosporium* OB3b. *Microorganisms*. 2020;8(3):437.
300. Zhang Y, Lin Y-H. Metabolic flux analysis of *Saccharomyces cerevisiae* during redox potential-controlled very high-gravity ethanol fermentation. *Biotechnology and Applied Biochemistry*. 2020;67(1):140–7.
301. Martínez-Monge I, Comas P, Triquell J, Casablanco A, Lecina M, Paredes CJ, et al. Concomitant consumption of glucose and lactate: A novel batch production process for CHO cells. *Biochemical Engineering Journal*. 2019;151:107358.
302. Noecker C, Chiu H-C, McNally CP, Borenstein E. Defining and evaluating microbial contributions to metabolite variation in microbiome-metabolome association studies. *Msystems*. 2019;4(6).
303. Mesquita TJB, Sargo CR, Fuzer JR, Paredes SAH, de Campos Giordano R, Horta ACL, et al. Metabolic fluxes-oriented control of bioreactors: a novel approach to tune micro-aeration and substrate feeding in fermentations. *Microbial cell factories*. 2019;18(1):1–17.
304. Yang L, Saunders MA, Lachance J-C, Palsson BO, Bento J. Estimating Cellular Goals from High-Dimensional Biological Data. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2019. p. 2202–11.
305. Lloyd CJ, King ZA, Sandberg TE, Hefner Y, Olson CA, Phaneuf PV, et al. The genetic basis for adaptation of model-designed syntrophic co-cultures. *PLoS computational biology*. 2019;15(3):e1006213.
306. Masid M, Ataman M, Hatzimanikatis V. Analysis of human metabolism by reducing the complexity of the genome-scale models using redHUMAN. *Nature Communications*. 2020;11(1):1–12.
307. Gerlin L, Cottret L, Cesbron S, Taghouti G, Jacques M-A, Genin S, et al. Genome-Scale Investigation of the Metabolic Determinants Generating Bacterial Fastidious Growth. *Msystems*. 2020;5(2).

308. Ong WK, Midford PE, Karp PD. Taxonomic weighting improves the accuracy of a gap-filling algorithm for metabolic models. *Bioinformatics*. 2020;36(6):1823–30.
309. Lopes HJS, Bonturi N, Kerkhoven EJ, Miranda EA, Lahtvee P-J. C/N ratio and carbon source-dependent lipid production profiling in *Rhodotorula toruloides*. *Applied microbiology and biotechnology*. 2020;1–11.
310. Seif Y, Monk JM, Mih N, Tsunemoto H, Poudel S, Zuniga C, et al. A computational knowledge-base elucidates the response of *Staphylococcus aureus* to different media types. *PLoS computational biology*. 2019;15(1):e1006644.
311. Tokuyama K, Toya Y, Horinouchi T, Furusawa C, Matsuda F, Shimizu H. Application of adaptive laboratory evolution to overcome a flux limitation in an *Escherichia coli* production strain. *Biotechnology and bioengineering*. 2018;115(6):1542–51.
312. Angione C. Integrating splice-isoform expression into genome-scale models characterizes breast cancer metabolism. *Bioinformatics*. 2018;34(3):494–501.
313. Ben Guebila M, Thiele I. Predicting gastrointestinal drug effects using contextualized metabolic models. *PLoS computational biology*. 2019;15(6):e1007100.
314. Shaw R, Cheung CM. A mass and charge balanced metabolic model of *Setaria viridis* revealed mechanisms of proton balancing in C4 plants. *BMC bioinformatics*. 2019;20(1):1–11.
315. Venayak N, von Kamp A, Klamt S, Mahadevan R. MoVE identifies metabolic valves to switch between phenotypic states. *Nature communications*. 2018;9(1):1–9.
316. Yaneske E, Angione C. The poly-omics of ageing through individual-based metabolic modelling. *BMC bioinformatics*. 2018;19(14):83–96.
317. Hädicke O, von Kamp A, Aydogan T, Klamt S. OptMDFpathway: Identification of metabolic pathways with maximal thermodynamic driving force and its application for analyzing the endogenous CO₂ fixation potential of *Escherichia coli*. *PLoS computational biology*. 2018;14(9):e1006492.
318. Mishra P, Lee N-R, Lakshmanan M, Kim M, Kim B-G, Lee D-Y. Genome-scale model-driven strain design for dicarboxylic acid production in *Yarrowia lipolytica*. *BMC systems biology*. 2018;12(2):9–20.
319. Robinson JL, Kocabaş P, Wang H, Cholley P-E, Cook D, Nilsson A, et al. An atlas of human metabolism. *Science Signaling*. 2020;13(624).
320. Chang RL, Xie L, Xie L, Bourne PE, Palsson BØ. Drug off-target effects predicted using structural analysis in the context of a metabolic network model. *PLoS Comput Biol*. 2010;6(9):e1000938.
321. Gatto F, Ferreira R, Nielsen J. Pan-cancer analysis of the metabolic reaction network. *Metabolic engineering*. 2020;57:51–62.

322. He Y, Wang Y, Zhang B, Li Y, Diao L, Lu L, et al. Revealing the metabolic characteristics of human embryonic stem cells by genome-scale metabolic modeling. *FEBS letters*. 2018;592(22):3670–82.
323. Rizvi A, Shankar A, Chatterjee A, More TH, Bose T, Dutta A, et al. Rewiring of metabolic network in *Mycobacterium tuberculosis* during adaptation to different stresses. *Frontiers in microbiology*. 2019;10:2417.
324. Tajparast M, Frigon D. Predicting the accumulation of storage compounds by *Rhodococcus jostii* RHA1 in the feast-famine growth cycles using genome-scale flux balance analysis. *Plos one*. 2018;13(3):e0191835.
325. Heckmann D, Lloyd CJ, Mih N, Ha Y, Zielinski DC, Haiman ZB, et al. Machine learning applied to enzyme turnover numbers reveals protein structural correlates and improves metabolic models. *Nature communications*. 2018;9(1):1–10.
326. Stanway RR, Bushell E, Chiappino-Pepe A, Roques M, Sanderson T, Franke-Fayard B, et al. Genome-scale identification of essential metabolic processes for targeting the *Plasmodium* liver stage. *Cell*. 2019;179(5):1112–1128. e26.
327. Krishnan A, Kloehn J, Lunghi M, Chiappino-Pepe A, Waldman BS, Nicolas D, et al. Functional and computational genomics reveal unprecedented flexibility in stage-specific *Toxoplasma* metabolism. *Cell Host & Microbe*. 2020;27(2):290–306. e11.
328. Bro C, Regenbreg B, Förster J, Nielsen J. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metabolic engineering*. 2006;8(2):102–11.
329. Huang Y-Y, Jian X-X, Lv Y-B, Nian K-Q, Gao Q, Chen J, et al. Enhanced squalene biosynthesis in *Yarrowia lipolytica* based on metabolically engineered acetyl-CoA metabolism. *Journal of biotechnology*. 2018;281:106–14.
330. Özcan E, Selvi SS, Nikerel E, Teusink B, Öner ET, Çakır T. A genome-scale metabolic network of the aroma bacterium *Leuconostoc mesenteroides* subsp. *cremoris*. *Applied microbiology and biotechnology*. 2019;103(7):3153–65.
331. Klamt S, Müller S, Regensburger G, Zanghellini J. A mathematical framework for yield (vs. rate) optimization in constraint-based modeling and applications in metabolic engineering. *Metabolic engineering*. 2018;47:153–69.
332. Pan D-T, Wang X-D, Shi H-Y, Yuan D-C, Xiu Z-L. Dynamic flux balance analysis for microbial conversion of glycerol into 1, 3-propanediol by *Klebsiella pneumoniae*. *Bioprocess and biosystems engineering*. 2018;41(12):1793–805.
333. Ferreira R, Skrekas C, Hedin A, Sánchez BJ, Siewers V, Nielsen J, et al. Model-assisted fine-tuning of central carbon metabolism in yeast through dCas9-based regulation. *ACS Synthetic Biology*. 2019;8(11):2457–63.

334. Ramirez-Malule H, Junne S, Cruz-Bournazou MN, Neubauer P, Ríos-Esteva R. *Streptomyces clavuligerus* shows a strong association between TCA cycle intermediate accumulation and clavulanic acid biosynthesis. *Applied microbiology and biotechnology*. 2018;102(9):4009–23.
335. Patel A, Carlson RP, Henson MA. In Silico Metabolic Design of Two-Strain Biofilm Systems Predicts Enhanced Biomass Production and Biochemical Synthesis. *Biotechnology journal*. 2019;14(7):1800511.
336. Nilsson A, Björnson E, Flockhart M, Larsen FJ, Nielsen J. Complex I is bypassed during high intensity exercise. *Nature communications*. 2019;10(1):1–11.
337. Lemgruber R de SP, Valgepea K, Tappel R, Behrendorff JB, Palfreyman RW, Plan M, et al. Systems-level engineering and characterisation of *Clostridium autoethanogenum* through heterologous production of poly-3-hydroxybutyrate (PHB). *Metabolic engineering*. 2019;53:14–23.
338. De Winter K, Ghesquière J, Teughels W, Waldherr S, Bernaerts K. Towards the construction of GSMN-based community model for an oral biofilm. *IFAC-PapersOnLine*. 2019;52(26):193–9.
339. Bidkhor G, Benfeitas R, Klevstig M, Zhang C, Nielsen J, Uhlen M, et al. Metabolic network-based stratification of hepatocellular carcinoma reveals three distinct tumor subtypes. *Proceedings of the National Academy of Sciences*. 2018;115(50):E11874–83.
340. Bidkhor G, Benfeitas R, Elmas E, Kararoudi MN, Arif M, Uhlen M, et al. Metabolic network-based identification and prioritization of anticancer targets based on expression data in hepatocellular carcinoma. *Frontiers in physiology*. 2018;9:916.
341. Lewis JE, Costantini F, Mims J, Chen X, Furdui CM, Boothman DA, et al. Genome-scale modeling of NADPH-driven β -lapachone sensitization in head and neck squamous cell carcinoma. *Antioxidants & redox signaling*. 2018;29(10):937–52.
342. Ankrah NY, Chouaia B, Douglas AE. The cost of metabolic interactions in symbioses between insects and bacteria with reduced genomes. *MBio*. 2018;9(5).
343. Schwalm ND, Mojadedi W, Gerlach ES, Benyamin M, Perisin MA, Akingbade KL. Developing a microbial consortium for enhanced metabolite production from simulated food waste. *Fermentation*. 2019;5(4):98.
344. Perisin MA, Sund CJ. Human gut microbe co-cultures have greater potential than monocultures for food waste remediation to commodity chemicals. *Scientific reports*. 2018;8(1):1–10.
345. Malek Shahkouhi A, Motamedian E. Reconstruction of a regulated two-cell metabolic model to study biohydrogen production in a diazotrophic cyanobacterium *Anabaena variabilis* ATCC 29413. *PloS one*. 2020;15(1):e0227977.

346. Ye C, Zou W, Xu N, Liu L. Metabolic model reconstruction and analysis of an artificial microbial ecosystem for vitamin C production. *Journal of biotechnology*. 2014;182:61–7.
347. Heinken A, Sahoo S, Fleming RM, Thiele I. Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut microbes*. 2013;4(1):28–40.
348. Hu B, Wang M, Geng S, Wen L, Wu M, Nie Y, et al. Metabolic exchange with non-alkane-consuming *Pseudomonas stutzeri* SLG510A3-8 improves n-alkane biodegradation by the alkane degrader *Dietzia* sp. strain DQ12-45-1b. *Applied and environmental microbiology*. 2020;86(8).
349. Zimmermann J, Obeng N, Yang W, Pees B, Petersen C, Waschina S, et al. The functional repertoire contained within the native microbiota of the model nematode *Caenorhabditis elegans*. *The ISME journal*. 2020;14(1):26–38.
350. Phalak P, Henson MA. Metabolic modelling of chronic wound microbiota predicts mutualistic interactions that drive community composition. *Journal of applied microbiology*. 2019;127(5):1576–93.
351. Chan SH, Friedman ES, Wu GD, Maranas CD. Predicting the longitudinally and radially varying gut microbiota composition using multi-scale microbial metabolic modeling. *Processes*. 2019;7(7):394.
352. Xu X, Zarecki R, Medina S, Ofaim S, Liu X, Chen C, et al. Modeling microbial communities from atrazine contaminated soils promotes the development of biostimulation solutions. *The ISME journal*. 2019;13(2):494–508.
353. Henson MA, Orazi G, Phalak P, O’Toole GA. Metabolic modeling of cystic fibrosis airway communities predicts mechanisms of pathogen dominance. *MSystems*. 2019;4(2).
354. Di Stefano A, Scatà M, Vijayakumar S, Angione C, La Corte A, Liò P. Social dynamics modeling of chrono-nutrition. *PLoS computational biology*. 2019;15(1):e1006714.
355. Alzoubi D, Desouki AA, Lercher MJ. Alleles of a gene differ in pleiotropy, often mediated through currency metabolite production, in *E. coli* and yeast metabolic simulations. *Scientific reports*. 2018;8(1):1–9.
356. Hale VL, Jeraldo P, Chen J, Mundy M, Yao J, Priya S, et al. Distinct microbes, metabolites, and ecologies define the microbiome in deficient and proficient mismatch repair colorectal cancers. *Genome medicine*. 2018;10(1):1–13.
357. Hunt KA, Jennings RM, Inskeep WP, Carlson RP. Multiscale analysis of autotroph-heterotroph interactions in a high-temperature microbial community. *PLoS computational biology*. 2018;14(9):e1006431.



ZHALEH HOSSEINI

Bioinformatician and Computational biologist

Birthday: 11 March 1991

Address: Rue de Verdeaux 7D, 1020, Renens, Switzerland

Phone: +41 (78) 6411857

Email: zhaleh.hosseini@epfl.ch

Linkedin: <https://www.linkedin.com/in/zhaleh-hosseini-7b450159/>

Education

Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

June 2016-Oct. 2020

Ph.D. in Computational Systems Biotechnology

Thesis title: “*Optimization methods and objective functions for analysis of metabolic network models*”

Selected courses: Optimization methods and models, Principles and applications of systems biology

University of Tehran (UT), Tehran, Iran

Feb. 2013-May 2016

M.Sc. in a Direct Ph.D. program of Biotechnology, GPA: 18.48/20

Thesis title: “*Gap filling of metabolic networks using flux coupling analysis and expression profiles of metabolic genes*”

Selected courses: Advanced Bioinformatics, Molecular Pharmacogenetics, Advanced Molecular Biology, Molecular and cellular biophysics

University of Tehran (UT), Tehran, Iran

Sep. 2009- Feb. 2013

B.Sc. in Direct Ph.D. program of Biotechnology, GPA: 18.36/20

Thesis title: “*Developing a novel FCA-based method for hierarchical organization of fluxes in metabolic networks*”

Selected courses: Advanced Programming in C++, Immunology, Cellular & Molecular Biology, Molecular Genetics, Advanced Statistics

Profile Summary

Over the past years, I mastered diverse set of technical skills in the field of **computational and systems biology**. My main technical expertise are: experience in **constraint-based** and kinetic modeling, machine learning, data visualization, data analysis using different software and programming languages, **omics data analysis**, and developing and implementing different optimization algorithms. I have worked with different experimental researchers to analyse their data and use it for modeling and prediction.

Selected Publications

In Preparation:

- **Z. Hosseini**, V. Hatzimanikatis (2020) “*A Flux coupling analysis framework for the study of metabolic network models that include integer variables*”.
- **Z. Hosseini**, O. Oftadeh, V. Hatzimanikatis (2020) “*Global optimization problem to describe different objective functions used in studying cellular metabolism*”.
- A. Chiappino-pepe, E. Vayena, **Z. Hosseini**, H. Frammery, V. Hatzimanikatis (2020) “*Objective function, essentiality, and metabolic hallmarks of dormant malaria parasites*”.
- J. Hafner, B. Lopes, A. Chiappino-Pepe, **Z. Hosseini**, V. Hatzimanikatis (2020) “*Atom-level constraint-based model of the Escherichia coli metabolism*”.

Peer-Reviewed Journal Papers:

- C. List, **Z. Hosseini**, K. Lederballe Meibom, V. Hatzimanikatis, R. Bernier-Latmani (2019) “*Impact of iron reduction on the metabolism of Clostridium acetobutylicum*”, Published in Environmental Microbiology.
- **Z. Hosseini**, SA. Marashi (2017) “*Discovering missing reactions of metabolic networks by using gene co-expression data*”, Scientific Reports, 7: 41774.
- **Z. Hosseini**, SA. Marashi (2015) “*Hierarchical organization of fluxes in Escherichia coli metabolic network: using flux coupling analysis for understanding the physiological properties of metabolic genes*”, Gene, 561(2):199-208.
- SA. Marashi, **Z. Hosseini** (2015) “*On correlated reaction sets and coupled reaction sets in metabolic networks*”, Journal of Bioinformatics and Computational Biology, 13 (4): 1571003.

Peer-Reviewed Conference Papers:

- **Z. Hosseini**, V. Hatzimanikatis: “*Coupling analysis for evaluating thermodynamically constrained metabolic network models*”, From Functional Genomics to Systems Biology, Germany, 2018.

- **Z. Hosseini**, V. Hatzamanikatis: “*Flux coupling analysis for studying the dependencies and degrees of freedom in thermodynamically constrained metabolic network models*”, 253rd American Chemical Society National meeting and Exposition, Advanced Materials, Technologies, Systems and Processes, US, 2017.
- **Z. Hosseini**, SA. Marashi “*Gap Filling of Metabolic Networks Using Flux Coupling Analysis and Expression Profiles of Metabolic Genes*”, 4th Conference on Constraint-Based Reconstruction and Analysis (COBRA), Germany, 2015.

Research Experience

Research Assistant, Laboratory of Computational Systems Biotechnology, EPFL, Switzerland, Jun. 2016-present

- Developed a novel MILP-based method for computing flux couplings in thermodynamically constrained metabolic network models in order to study degrees of freedom and impact of thermodynamics (Gibbs free energy and metabolite concentrations) in these networks (MATLAB).
- Collaborative developing of a method for analysis and integration of C13 labelling data into metabolic models to study the impact of different physiological states on the distribution of atom labels in the central carbon metabolism (MATLAB)
- Developed a pipeline for integration and evaluation of novel pathways in metabolic network models and collaborated with researchers in Nestle and SilicoLife institutes to evaluate novel pathways for production of different target metabolites (MATLAB)
- Worked on a thorough review on optimization methods and objective functions used in metabolic network analysis. Building a general platform for each class of optimization problems is in progress (MATLAB and Python)
- Collaborated with environmental microbiology researchers to study iron reduction and transfer in *Clostridium Acetobutylicum* for the first time using time-series experimental data on metabolite concentrations and RNA sequencing data (MATLAB)
- Analysed malaria parasite metabolism using different experimental data and metabolic modelling and hypothesized and validated possible objective functions for the latency phase of this organism (MATLAB)

Research Assistant, Systems Biology Lab, University of Tehran, Iran, Jul. 2012-May 2016

- Designed a novel MILP-based method for gap filling of metabolic networks based on flux coupling analysis, co-expression of metabolic genes and KEGG database for selection and validation of candidate gap filling reactions (MATLAB).
- Collaborated with phytochemistry researchers to find the over-represented pathways in rice plants under diazinon stress using GC-MS metabolomics data of rice (*Oryza sativa L.*) and bioinformatics/metabolic network-based methods (MetaboAnalyst online tool, MetaboNetworks toolbox in MATLAB, analysis of reporter reactions using COBRA toolbox in MATLAB, and R).
- Built a novel MILP-based method for linking metabolomics data to the activity of metabolic pathways by using metabolic network models and elementary flux mode analysis (MATLAB).
- Worked on comparing the *in silico* and *in vitro* capabilities of two *Bacillus* species (*Bacillus subtilis* and *Bacillus megaterium*) in order to investigate the accuracy of corresponding genome-scale models (MATLAB).
- Developed a method for hierarchical organization of metabolic networks based on flux coupling analysis and the association between different features of reactions with their level in hierarchical flux coupling graph (MATLAB and R).
- Worked on comparative analysis of flux coupling analysis and flux correlation analysis to explain the inherent difference of the two concepts in the field of metabolic network analysis (using MATLAB and R).

Research Assistant, Molecular Biotechnology Lab, Iran, Jul. 2012-Nov. 2012

- Worked on extracting lipase producing bacteria from different sources of soil in a metagenomics study.

Summer Intern, Kowsar Medical Genetics Lab, Iran, Jul. 2011-Sep. 2011

- Worked on DNA extraction, PCR, Gel electrophoresis, Analysis of DNA sequencing results, Primer design, etc.

Teaching Experience

- 3 years teaching assistant in Bioreactor modeling and simulation course in EPFL (used MATLAB to design different projects based on ODE systems for modeling different types of bioreactors)
- 1 year teaching assistant in numerical methods in chemistry and introduction to chemical engineering courses in EPFL

Language

English: Fluent
Native

French: Basic

German: Basic

Persian:

Technical Skills

- **MATLAB**, R, Python, C++
- Statistical Analysis, Optimization Programming, **Bioinformatics (handling and analysis of “omics” data)**
- **Constrained-Based Modeling (FBA, FVA, FCA, Monte Carlo Sampling of metabolic models)**, Kinetic Modeling