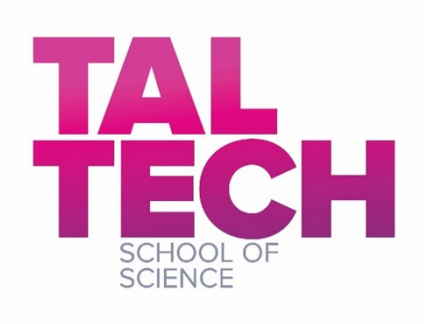

Flux balance analysis of continuous cultivation for investigating the lipogenesis metabolism in *Rhodotorula toruloides*

Bachelor thesis

Student: Maive Hanni
Supervisor: Alina Rekena, Department of Chemistry and Biotechnology, Early Stage Researcher
Study program: Applied Chemistry and Gene Technology

Tallinn 2024

Rhodotorula toruloides lipogeneesi uurimine, kasutades metaboolsete voogude analüüsi

Bakalaureusetöö

Üliõpilane: Maive Hanni
Juhendaja: Alina Rekena, Keemia ja biotehnoloogia instituut, Doktorant-nooremteadur
Õppekava: Rakenduskeemia ja geenitehnoloogia

Tallinn 2024

Declaration

Hereby I declare that I have compiled the paper independently and all works, important standpoints and data by other authors have been properly referenced and the same paper has not been previously been presented for grading.

Author: Maive Hanni
Signature, date

The paper conforms to requirements in force.
Supervisor: name
Signature, date

Permitted to the defence. Chairman of the Defence Committee: Name
Signature, date

Contents

Abstract

Annotatsioon

Abbreviations

Introduction

1. Theoretical background

1.1. *Rhodotorula toruloides*

Rhodotorula toruloides (previously *Rhodosporidium toruloides*) is a non-conventional, oleaginous yeast, that can naturally accumulate high amounts of lipids [1]. This yeast occurs naturally in leaves, soil, sea water, etc. It has a broad substrate range, which has made this yeast a popular for producing biological oils from inedible substrates such as pentose sugars and crude glycerol.*R. toruloides* lipid fraction contains also carotenoid pigments, omega-3 linolenic acid and heptadecenoic acid, which makes this yeast a promising organism for production of pharma- and nutraceuticals [2]. It has been found that nitrogen limitation further induces lipid accumulation, 65% of lipids of dry cell weight were reached in a batch cultivation regime [1].

One of the major determinants of the oleaginous phenotype of *R. toruloides* is its capacity for acetylCoA production. Unlike non-oleaginous yeasts such as *Saccharomyces cerevisiae*, *R. toruloides* possesses the enzyme ATP-citrate lyase, which is the main source of acetyl-CoA for lipid synthesis. Mitochondrial beta-oxidation pathway provides additional source of acetylCoA in this yeast. [2] Synthesis of acetyl-CoA from xylulose 5-phosphate by phosphoketolase (XPK), and the conversion of S-malate into pyruvate by malic enzyme (ME) that provides for NADPH enzymatic pathways also differ from the model yeast *Saccharomyces cerevisiae* and which specifically facilitate the generation of lipid precursors.[1] Lipid biosynthetic reactions downstream of acetyl-CoA synthesis do not differ between oleaginous and non-oleaginous yeast species [2].

In *R. toruloides* metabolic pathways producing acetyl-CoA and a cofactor NADPH have been the main focus of metabolic studies due to their central role in lipid biosynthesis. Fatty acids mainly accumulate as triacylglycerols (TAGs), and they are produced via four enzymatic reactions that require 1 ATP and 2 NADPH molecules for adding 1 acetyl-CoA to the fatty acid chain. Proteomics analysis has suggested that NADPH is primarily regenerated through the pentose phosphate pathway, when grown on xylose and glucose, but the role of malic enzyme is not clearly understood. The role of XPK in the generation of acetyl-CoA has not been acknowledged previously, whereas ACL has been demonstrated to be upregulated during lipid accumulation, especially in presence of xylose.[1]

1.2. Overview of cellular growth laws

1.3. Genome-scale metabolic modeling

A microorganism's phenotype can be described by its pattern of metabolic fluxes [3]. Metabolic flux is the rate of turnover of molecules through a metabolic pathway. Flux is regulated by the enzymes involved in this pathway. Within cells, regulation of flux is vital for all metabolic pathways to regulate the pathway's activity under different conditions. [4]

In biotechnology the aim is often to increase the capacity of specific fluxes. For this, metabolic engineering methods have been developed and many of these rely on balancing of intracellular metabolites using genome-scale models (GEMs) that in combination with appropriate objective functions and constraints can be used to predict potential gene targets for obtaining a preferred flux distribution. [3]

Genome-scale metabolic models are comprehensive summaries of metabolic reactions of a organism, which are derived from the organism's genome sequence [2]. GEMs are organism specific and they account for genes, proteins, and biochemical reactions, which enable systematic analysis of metabolism, where typically the objective is to achieve a global picture of possible flux patterns [3], [5]. GEMs can be a powerful and helpful tool in metabolic studies, if their predictive power is good [1]. They can be used for identification of potential metabolic engineering targets, or select the best performing metabolic network among alternatives. The first genome-scale model of *R. toruloides* metabolism (rhto-GEM), includes 4869 genes, 897 reactions, and 3334 metabolites. The simulation results confirmed that the *R. toruloides* model provides valid growth predictions on glucose, xylose and glycerol. This makes rtho-GEM useful for future studies to improve the production of other industrially important oleochemicals including both value-added fatty acids and carotenoids. [2]

The reconstructed networks can be converted into mathematical stoichiometric matrices, where rows represent metabolites and columns represent individual reactions. This matrix will determine the solution space of the metabolic network. GEMs are constrained by (1) the stoichiometry of the network; (2) preset upper and lower boundaries for selected reactions (usually substrate uptake reactions, but can also describe thermodynamics); and (3) the assumption of a steady state. [3] Reconstruction of GEMs is shown in the figure 1.1.

Steady state assumption presumes that all the flux rates and metabolite concentrations are constant over time. In experiments this state can be reached in chemostat cultivation and, in theory, during controlled batch logarithmic growth phase when organisms grow at their maximum specific growth rate and the change of neither substrate concentration nor product formation will affect the growth regulation. [3]

Constraint-based modeling

Metabolism operates under countless of constraints. These include constraints that remain

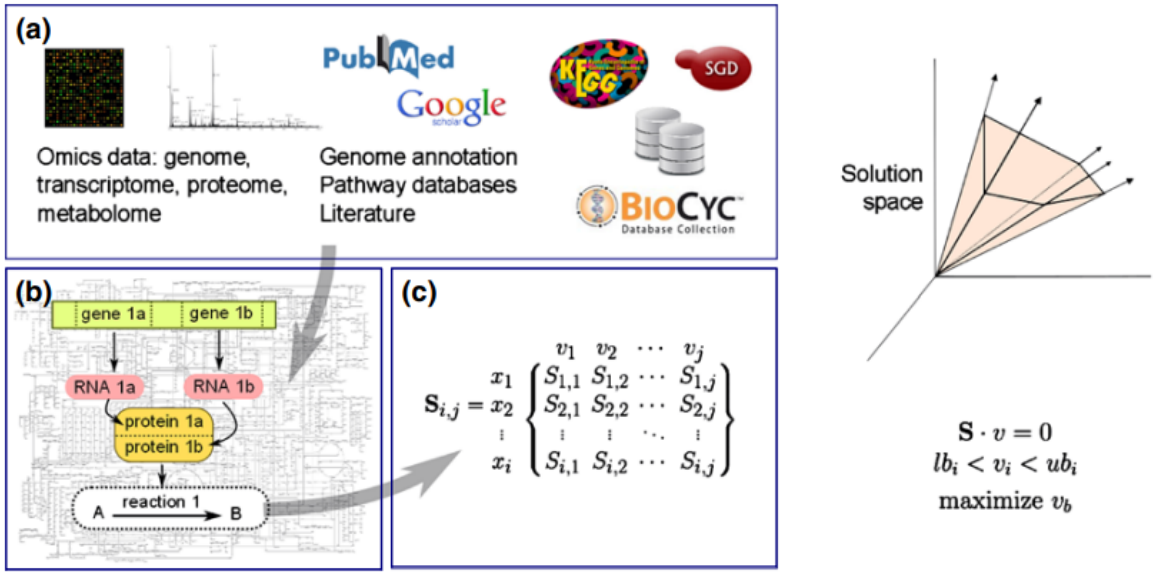

Figure 1.1: Reconstruction of GEMs. (a) The genome annotation can be used to reconstruct the metabolic network. (b) Gene-protein-reaction relationships are defined for the metabolic model. (c). A solution space is defined from the constraints applied to the model and subsequent analysis provides instructions for metabolic engineering strategies. This figure has been taken from article [3].

unchanged as they abide by the laws of physics (e.g. conservation of mass and energy, thermodynamics), and constraints that can vary by organism, environmental condition, the state of the cell (e.g. nutrient uptake rate, biomass composition), and constraints that can also change through evolution or changes in gene expression [Citation].

A major limitation in the predictive power of conventional GEMs is that they ignore the capacity of a cell to support a metabolic flux is constrained by its resource allocation, as most metabolic reactions are catalyzed by enzymes. The synthesis of enzymes is resource- and energy-expensive, their catalytic capacities are limited by their kinetics, also the quantity of enzymes is space-constraint, such that stringency in resource allocation is vital for optimal cell growth. The modeling problem of resource allocation can be narrowly defined by only considering protein allocation: "given a certain budget, what is the best way to distribute it", where budget refers to the total cellular protein level that is distributed over all its constituent proteins. [Citation]

This means that an increase in the requirement of an enzyme or a pathway would be a trade-off for other functions. Experimental evidence indicates that for various organisms resource re-allocation could be an effective strategy in response to nutrient and growth shift, which demonstrates the biological significance of proteome constraints. [5] Applying such constraints in a metabolism model reduces simulated flux distributions to those that are most economic and also limits the phenotypes that the model can simulate. These both contribute to more realistic results. Such models have already found numerous applications in e.g. unraveling the underlying mechanisms for observed metabolic phenotypes and the prediction of strain optimization strategies.[Citation] This suggests that proteome constraints could be a valuable

addition to GEMs to improve model predictions [5].

The first approach that allows for a direct incorporation of proteomics data to account for enzyme limitation was the GECKO framework. The GECKO method is built on the principle that any metabolic reaction flux has a biologically natural constraint equal to the intracellular enzyme concentration multiplied by the enzyme's turnover number (k_{cat}). In metabolic network, the enzyme constraint is defined as maximum rate of enzymatic reaction (v_{max}) that the metabolic flux cannot exceed. [Citation]

Enzyme-constrained GEMs integrate additional constraints on enzyme capacity and their total abundances. Phenomenological constraint is imposed on metabolic flux (v; mmol/gDCW/h), formulated as enzyme kinetics: $v <= E \cdot k_{cat}$, where E is protein abundance (mmol/gDCW) and k_{cat} is the enzyme's turnover number (1/s), provided with an upper limit on individual or total protein abundances. The integration of enzymatic constraints in *S. cerevisiae* has significantly improved phenotype prediction. [1]

Flux balance analysis

Flux balance analysis (FBA) is commonly used for adding additional constraints, where an assumed biological objective is applied in the form of maximizing (or minimizing) a certain flux. This results in a solution in the solution space that satisfies the presumed objective. The most commonly used objective function is maximization of the specific growth rate, ATP generation or product formation. FBA is often used for estimating the biotechnological potential of microorganisms and pinpoint genetic manipulations that could improve the performance of a cell. The main applications of FBA are: (1) Instructions for metabolic engineering purposes; (2) Biological interpretation and discovery through contextualizing high-throughput data; (3) Development of a computational framework; (4) Evolutionary elucidation; (5) Description of multispecies communities. [3]

1.4. Overview of microbial cultivation methods

2. Aims of the thesis

3. Materials and methods

3.1. Criteria for selecting experimental data

3.2. Biomass integration to the model

3.3. Flux balance analysis and sampling of the solution space

4. Results

5. Conclusion

Summary

Acknowledgements

References

[1] A. Reķēna, M. J. Pinheiro, N. Bonturi, *et al.*, "Genome-scale metabolic modeling reveals metabolic trade-offs associated with lipid production in rhodotorula toruloides," *PLOS Computational Biology*, vol. 19, no. 4, R. Mahadevan, Ed., e1011009, Apr. 2023, ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1011009.

[2] I. A. Tiukova, S. Prigent, J. Nielsen, M. Sandgren, and E. J. Kerkhoven, "Genome-scale model of rhodotorula toruloides metabolism," *Biotechnology and Bioengineering*, vol. 116, no. 12, pp. 3396–3408, Sep. 2019, ISSN: 1097-0290. DOI: 10.1002/bit.27162.

[3] E. J. Kerkhoven, P.-J. Lahtvee, and J. Nielsen, "Applications of computational modeling in metabolic engineering of yeast," *FEMS Yeast Research*, n/a–n/a, Sep. 2014, ISSN: 1567-1356. DOI: 10.1111/1567-1364.12199.

[4] D. Voet and J. G. Voet, *Biochemistry (2nd ed.) Donald*, 2. rev. ed. New York [u.a.]: Wiley, 1995, vol. [Hauptbd.] 1361 pp., ISBN: 047158651X.

[5] M. Chen, T. Xie, H. Li, Y. Zhuang, J. Xia, and J. Nielsen, "Yeast increases glycolytic flux to support higher growth rates accompanied by decreased metabolite regulation and lower protein phosphorylation," *Proceedings of the National Academy of Sciences*, vol. 120, no. 25, Jun. 2023, ISSN: 1091-6490. DOI: 10.1073/pnas.2302779120.

Supplementary