

Biotech Beer Brewing

Final report

Dominik Schmidt
Jakob Wittmann

Abstract—Bacterial contamination in alcoholic fermentation lowers the productivity of yeast and has a negative influence on process costs and product quality. Finding new ways to enhance the process is necessary. Simulation methods using genome-scale models has gained popularity as they provide deep insight in internal cell processes and are much cheaper than using real bacteria cultures. Latest implementations uses dynamic flux balance analyses (DFBA) methods to simulate competitive co-cultures. One successfully applied framework for Matlab is *Dynamic Multispecies Metabolic Modeling* (DMMM). In this work the implementation of the DMMM is discussed and implemented in python using the COBRApy package. As the contamination of *Lactobacillus plantarum* is a major issue in alcoholic fermentation in food industry, the simulator is demonstrated in a co-simulation of *L. plantarum* and *Saccharomyces cerevisiae*. The results show...

I. INTRODUCTION

Alcoholic fermentation is applied in industrial scale in different markets like food industry e.g. for wine, beer or whisky production but also other market segments like bio fuel production. Reducing production costs while increasing the process efficiency and product quality is a major goal in industry. In the fermentation process contamination from tanks, pipes, centrifuges, heat exchangers or ingredients can have a great impact on process parameters and must be controlled. Contamination can lead to a loss in process productivity of up to 30%. The main contaminants are lactic acid bacteria and wild yeasts, lowering the performance of the applied yeast. *Lactobacillus* bacteria contribute to induce yeast flocculation what in turn reduces the viability of the cells, the fermentation rate and finally the process yield. Competition between yeasts and microbial contaminants affects the process additionally [6] [19].

In beer, wine or whisky production additional product quality aspects like color, viscosity and flavor plays a big role and has a great influence on its price. On the other hand the process itself underlies strict laws to guarantee transparency for the customers [2] [3]. Additionally the growing organic food market calls for products with even more strict processing standards [4].

Chemicals like acids, antibiotics and antifoam agents as they are used in other industries can not be applied here. These constraints increases the effort of finding optimal ways for controlling product quality dramatically. As experimental development in laboratories is very costly and not affordable by smaller breweries, simulation methods are gaining popularity.

A well established group of methods uses Genome-Scale Models (GEM) to model metabolic fluxes in bacteria [29]. GEMs are created by analysing the information in the genome of a bacteria about its enzymatical processes and extrapolating

a network of metabolites and metabolite reactions which represents the internal fluxes of this bacteria. Based on this model possible flux distributions can be predicted using linear optimization methods, known as flux balance analysis (FBA) [21].

In its basic form describes FBA the behavior of only one bacteria in a constant environment. Zomorri et al. summarizes some approaches to extend FBA to multispecies simulations (*steady-state models*) in time variant environments (*dynamic models*) and even considering spacial variations (*spatio-temporal models*) [29].

The simulation of co-cultures using GEMs is still an emerging field. Every approach has its strengths and weaknesses and must be chosen dependent on the application. In *Steady-state models* like compartmentalized community-level metabolic modeling a common objective function for all involved species must be formulated. This approach has a relatively low computational effort but shows drawbacks in competitive co-cultures where a common objective can not easily be formulated. The spacial dimension of *Spatio-temporal models* results in very high computational effort and it has to be evaluated if this enhances the accuracy of the estimated behavior in that case.

The class of *dynamic models* described by Zomorrodi et al. uses a system of ordinary differential equations (ODE) to extend the static simulation approach of FBA with external system dynamics like the exchange of metabolites and bacteria growth and is called dynamic flux balance analysis (DFBA) [29].

Mehadevan et al. describes two basic categories of DFBA approaches: *dynamic optimization approach* (DOA) and *static optimization approach* (SOA) [18].

In DOA the linear programming (LP) problem to predict the behavior of the bacteria and the differential equations are reformulized to a non-linear programming problem (NLP). This approach has a very high computational effort [13] compared to SOA and has only been used for relatively small GEMs with up to 13 modeled fluxes and 8 metabolites [17] [16].

In SOA the simulation time is discretised by a defined time interval and a FBA is executed for each time point. The results are used to iteratively solve the discretised differential equations to update the environmental conditions at each time point. The described integration method in SOA is similar to Euler-Cauchy. In this approach it is assumed that the cell internal dynamics are much faster than the external dynamics.

Höffner et al. basically generalizes SOA to arbitrary integration methods for solving the system of ODEs by using a third-party solver and names it *direct approach* (DA) [13].

Henson et al. introduces a further method which shows similarities to SOA and DA but enhances the efficiency of the algorithm [13] by reformulating the LP and ODE to a differential-algebraic equation system [11].

The DA described by Höffner et al. has been successfully applied to simulate alcohol fermentation in a wine approximating setup [24] [23] and to predict growth of *Escherichia coli* and *Saccharomyces cerevisiae* bacteria in a co-culture [8]. Zhuang et al. implemented the method in the *Dynamic Multispecies Metabolic Modeling* (DMMM) framework which is publicly accessible and written for Matlab [27]. It uses the Matlab toolbox *COBRA* [9] to implement FBA. The method developed by Henson et al. compared to DMMM has the potential to increase the accuracy by fixed calculation runtime but since the implementation of DMMM is publicly accessible and relatively simple to implement, this method is used in the simulator in this work.

The next chapter will give a overview over the applied models and algorithms followed by the description of a simulation setup including *Saccharomyces cerevisiae* and *Lactobacillus plantarum* which shall be serve as a use-case. As a first step *Saccharomyces cerevisiae* is simulated with different densities of lactic acid to tune the inhibition of the yeast model by the acid. In a second step *S. cerevisiae* and *L. plantarum* are simulated in a co-culture. The results are compared to the outcomes of *in vivo* experiments and discussed in chapter III. Chapter IV gives an outlook on further improvements of the implementation.

II. METHODS

A. Genome-Scale Models

A genome-scale model (GEM) is an analytical reconstruction of ideally all cellular process of a cell. All available biochemical, genetic and genomic (BiGG) information of the processes are analyzed to identify gene-to-protein-to-reaction (GPR) associations and involved compounds. Based on this data a network composed of compounds (nodes) and reactions (links) is created and added to the reconstruction. [22]

Tools like *Pathway Tools* [15], *SEED* [10] or *AUTOGRAPH* [20] are used to generate a reconstruction based on a (annotated) genome. The produced models contain many wrong assignments, gaps and inconsistencies due to many organism-specific choices and must be refined by hand. A biomass function to model growth and input/output reactions are added to the network [25].

The reconstruction is the basis to create geno- and phenotypes of a cell and to transform it into mathematical models for further analysis. An important mathematical representation is the stoichiometric matrix. It uniquely represents the metabolic network of the cell, based on the stoichiometric constants of its reactions [22].

B. Flux Balance Analysis

Flux Balance Analysis (FBA) is a mathematical tool which analysis possible flux distributions through the metabolic network based on the stoichiometric matrix of a GEM.

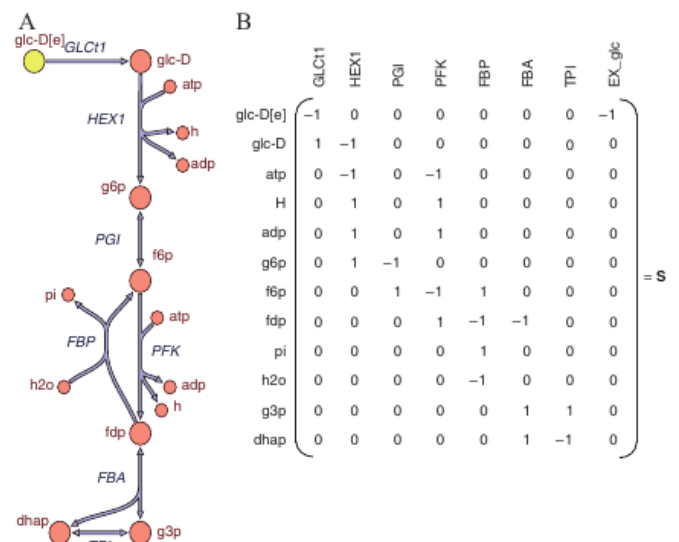


Figure 1: (create simplified version here, combine with figure 2)

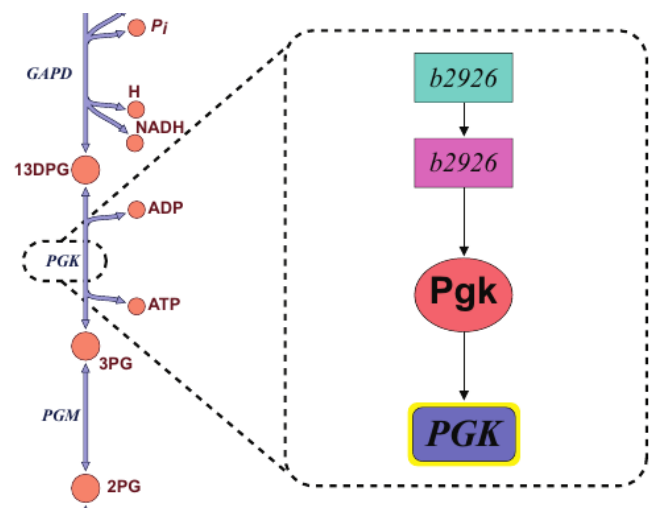


Figure 2: (create simplified version here, combine with figure 1)

The system of linear equations built by the stoichiometric matrix S of a GEM is usually underdetermined. To choose one flux distribution from the solution space spanned by S , it is assumed the organism adapted its metabolic fluxes to a certain environment and is now in a steady state where the fluxes are optimal distributed recording to some goal. In FBA this flux distribution is chosen by solving a optimization problem as in equation 1 using linear programming (LP) [21].

$$\begin{aligned} \max_v \quad & \mu = \mathbf{w}^T \mathbf{v} \\ \text{s.t.} \quad & \mathbf{S} \mathbf{v} = \mathbf{0}, \\ & v_{min} \leq v \leq v_{max} \end{aligned} \quad (1)$$

The fluxes in the metabolic network v are optimized so that μ is maximized where w is a weighting vector which defines the adaption goal of the cell. A often used goal is to maximize the growth of the organism. The optimization is done

using mostly two constraints, (1) the stoichiometric constants S for each reaction which is defined by the used GEM and (2) lower and upper bounds for each flux, v_{min} and v_{max} , in the metabolic network but further constraints can be added as well.

The lower and upper bounds can be used to simulate gene know-outs by setting the upper bound of reactions dependent on this gene to zero or to define environmental conditions like the availability of certain metabolites.

The chosen solution by the LP solver represents a phenotype of the given organism due to adaption to its environment but FBA has its limitations. Usually GEMs does not model regulatory structures as activation of enzymes or regulation of gene expressions. As the GEM does not contain kinetic parameters of the involved compounds, their densities can not be considered and must be modeled externally [21].

C. Simulation Algorithm

Table I: Overview of implemented features compared to DMMM

Feature	DMMM	This project
Model		
arbitrary many GEMs	yes	yes
arbitrary many metabolites in environment	yes	yes
mortality of bacteria	yes (in output flux)	yes
input/output flux of bacteria and metabolites	yes	no
parameterized initial state of environment composition	yes	yes
Michaelis-Menten kinetics	yes	yes
Algorithm		
ODE solver	yes	yes
different ODE solvers	yes	no
analytical solver	yes	no

As described by Zhuang et al. in [28] the algorithm uses a ODE solver with embedded FBA. A FBA is solved for each GEM in the model and for each time step in the discretised simulation time interval considering the changed metabolite and bacteria densities in the shared environment. The results of the FBAs are used by the ODE solver to solve the differential equations

$$\frac{dx_j}{dt} = \mu_j x_j \quad (2)$$

$$\frac{ds_i}{dt} = \sum_{j=1}^N v_{i,j} x_j \quad (3)$$

which models the dynamics of the bacteria's environment [27] where $i = 1 \dots N$ is the index of metabolites in the shared environment and $j = 1 \dots M$ is the index of bacteria. The bacteria density is modeled in x_j with $[x_j] = \frac{g}{l}$ and μ_j is the bacteria's growth rate with $[\mu_j] = \frac{mmol}{g_{DW}h}$. Input and output fluxes of the bacteria's models are modeled in $v_{i,j}$ with $[v_{i,j}] = \frac{mmol}{g_{DW}h}$, the densities of metabolites in the shared environment in s_i with $[s_i] = \frac{mmol}{l}$.

In each time step each bacteria's metabolite intake must be changed dependent on the densities of the metabolites in

the shared environment. To model saturation of metabolite intake for high metabolite densities Zhuang et al. implemented Michaelis-Menten kinetics [14]

$$v_{max,i,j} = \frac{v_{mm,i,j} s_i}{s_i + k_{mm,i,j}} \frac{1}{1 + \frac{s_a}{k_{a,i,j}}} \quad (4)$$

This formula describes the upper bound of the input flux $v_{max,i,j}$ for metabolite i of bacteria j dependent on the metabolite density s_i . The formula is characterized by to constants $[v_{mm,i,j}] = \frac{mmol}{g_{DW}h}$ and $[k_{mm,i,j}] = \frac{mmol}{l}$ for each bacteria and metabolite. s_a is the metabolite density of metabolite a , $k_{a,i,j}$ is the inhibition constant of metabolite a and describes the inhibition of the intake of metabolite i by metabolite a at species j .

Mortality is considered using a constant $[\mu_{mort,j}] = \frac{mmol}{g_{DW}h}$ for each bacteria j in this implementation while Zhuang et al. modeled this using the output flux of bacteria out of the system.

Algorithm 1 shows a basic implementation of the differential equations solved by an ODE solver during the simulation similar to DMMM [28].

The algorithm expects a list of bacteria models consisting of

- GEM of this bacteria: A , v_{min} , v_{max} , w_{growth}
- v_{mm} (Michaelis-Menten V_{max}) for each exchange metabolite and species
- k_{mm} (Michaelis-Menten K) for each exchange metabolite and species
- mortality μ_{mort}
- inhibition constants k_a

Furthermore a list of all exchange metabolites in the environment, the bacteria and metabolite densities.

Algorithm 1: Differential equation with embedded FBA

1 **function** `step`(`model1...modelM`, `m1...mN`, `x1...xM`, `s1...sN`);

Input : bacteria models `modelj`, exchange metabolites `mi` in environment, bacteria densities `xj`, metabolite densities `si`

Output : slope of bacteria and metabolite densities \dot{x}_j, \dot{s}_i

```

2 for j := 1 to M do
3   for i := 1 to M do
4     modelj :=
4       update_intake_bounds(modelj, sj, mi)
5   end
6 end
7 for j := 1 to M do
8   μj, vj := FBA(modelj, wgrowth)
9 end
10 μ := μ - μmort
11 ẋ := diag(μ) x
12 for j := 1 to M do
13   for i := 1 to N do
14     ḡ[mi] := ḡ[mi] + vj[mi] xj
15   end
16 end
17 return ẋ, ḡ
```

In a first step the upper bounds of the intake fluxes are updated for each bacteria j and exchange metabolite i . The function $update_intake_bounds(model_j, s_j, m_i)$ calculates the upper bounds using the formula 4 if the metabolite m_i is contained in $model_j$ as a exchange metabolite and updates this value in the model.

In a next step the GEMs are optimized for growth using FBA, the results are used as growth rate μ_j and actual input and output fluxes v_j of bacteria j in this time step.

The mortality is considered by subtracting the constants μ from the growth rates μ .

At last step the slopes \dot{x} and \dot{s} are calculated according to 2 and 3 and returned to the ODE solver.

D. Simulation Setup

The goal of the simulation is to validate the basic functionality of the simulator using a simplified setup of a realistic future simulation scenario. As defined in our project goals, this simulation scenario is the dynamic flux balance analysis (DFBA) of a co-culture of *Saccharomyces cerevisiae* and *Lactobacillus plantarum*.

As genome-scale models a model of *Lactobacillus plantarum* published by Teusink et al. [26]. A decision about a yeast model is not made yet.

The simulation will consider two input metabolites: oxygen and glucose. Table II and table III contain all values needed to define the initial metabolite conditions and kinetics.

Table II: Model constants used in the simulation setup

Constant	<i>S. cerevisiae</i>	<i>L. plantarum</i>
Maximum glucose uptake rate (mmol/g/h)	18.5	18.5
Maximum oxygen uptake rate (mmol/g/h)	2.5	2.5
Glucose uptake saturation constant (g/l)	0.5	0.5
Oxygen uptake saturation constant (mmol/l)	0.005	0.005
Mortality (?)	?	?
Glucose uptake inhibition by ethanol constant (g/l)	10 [12]	-
Glucose uptake inhibition by lactic acid constant (mol/mol ?)	TBD	-

Table III: Simulation parameters used in the simulation setup

Parameter	value	reference
Initial glucose density (mmol/l)	272.9 ... 1230.755	equation 7, table IV
Initial oxygen density (mmol/l)	0.5039	equation 8, table IV
Initial density of <i>S. cerevisiae</i> (mmol/l)	?	-
Initial density of <i>L. plantarum</i> (mmol/l)	?	-

To verify the basic functionality of the simulator the resulting bacteria densities and metabolite densities of ethanol, d- and l-lactate, oxygen and glucose will be compared to existing data.

III. RESULTS

IV. CONCLUSIONS

APPENDIX A

The following approximation is used to convert °C (“degree plato”) to a density measure (g/l) [7].

$$d_{total} = 4.13 \frac{g}{l} \frac{1}{^{\circ}P} p + 997 \frac{g}{l} \quad (5)$$

As the simulation framework expects metabolite densities relative to the total volume of the solution (mmol of metabolite per liter solution, mmol/l) the total density d_{total} must to be converted to a density s_{glc} . It is assumed that $V_{total} = V_{glc} + V_W$.

$$\begin{aligned} d_{total} &= \frac{m_{total}}{V_{total}} \\ d_{total} &= \frac{m_{glc} + m_w}{V_{total}} \\ d_{total} &= \frac{m_{glc} + d_w V_w}{V_{total}} \\ d_{total} &= \frac{m_{glc} + d_w (V_{total} - V_{glc})}{V_{total}} \\ d_{total} &= \frac{m_{glc} + d_w \left(V_{total} - \frac{m_{glc}}{d_{glc}} \right)}{V_{total}} \\ s_{glc} &= \frac{m_{glc}}{V_{total}} = \frac{d_{total} - d_w}{1 - \frac{d_w}{d_{glc}}} \end{aligned} \quad (6)$$

Combining equation 5 and 6, including all constants and converting it to mmol/l leads to:

$$s_{glc} = \left(63.857 \frac{1}{^{\circ}P} p - 46.385 \right) \frac{mmol}{l} \quad (7)$$

Table IV: Constants used in this document

Constant	symbol	value	reference
Oxygen saturation of water at 20°C (mg/l)	-	9.077	[1]
Molar mass of water (g/mol)	-	18.015	[5]
Molar mass of glucose (g/mol)	-	180.156	[5]
Density of water (g/l)	d_w	1.00	[5]
Density of glucose (g/l)	d_{glc}	1.56	[5]
Typical glucose/water solution density to brew beer (°P)	-	5...20	-

To calculate the initial oxygen density in the solution it is assumed that the solution is at 20 °C and fully saturated with oxygen:

$$s_{init,ox} = 9.077 \frac{mg}{l} = 9.077 \cdot 10^{-3} \frac{g}{l} = \frac{9.077 \cdot 10^{-3} \frac{g}{l}}{18.015 \cdot 10^{-3} \frac{g}{mmol}} = 0,5039 \frac{mmol}{l} \quad (8)$$

APPENDIX B

Table V: Rating of considered DFBA methods

Method	comp. effort	impl. complexity	flexibility
dynamic optimization approach (DOA)	high	medium-high	?
static optimization approach (SOA)	low	low	low
direct approach (DA)	medium	low	medium
reformulation to a differential-algebraic equation system	low-medium	high	?

ACKNOWLEDGMENT

This work was supported by the Technology Development Fund Spring 2018, Háskóli Íslands.

REFERENCES

- [1] "Chemical features of water," 1987. [Online]. Available: <http://www.fao.org/docrep/field/003/AC183E/AC183E04.htm>
- [2] "Vorläufiges biergesetz, bundesgesetzblatt 1993 teil i seite 1400," 1993. [Online]. Available: <http://archiv.jura.uni-saarland.de/BGBI/TEIL1/1993/19931400.1.HTML>
- [3] "Weingesetz nach änderung durch artikel 9 des gesetzes vom 27. juni 2017 (bgbl. i s. 1966)," 1994. [Online]. Available: https://www.gesetze-im-internet.de/weing_1994/BJNR146710994.html
- [4] "Demeter processing standards," 2017. [Online]. Available: http://www.demeter.net/sites/default/files/di_processing_stds_demeter_biodynamic_17_-_e.pdf
- [5] "Pubchem: Open chemistry database," 2018. [Online]. Available: <https://pubchem.ncbi.nlm.nih.gov/>
- [6] R. P. Brexó and A. S. Sant'Ana, "Impact and significance of microbial contamination during fermentation for bioethanol production," *Renewable and Sustainable Energy Reviews*, vol. 73, pp. 423–434, 2017.
- [7] Z. Bubnk, P. Kadlec, D. Urban, and M. Bruhns, *Sugar technologists manual: chemical and physical data for sugar manufacturers and users*. Bartens, 1995.
- [8] T. J. Hanly and M. A. Henson, "Dynamic flux balance modeling of microbial co-cultures for efficient batch fermentation of glucose and xylose mixtures," *Biotechnology and Bioengineering*, vol. 108, pp. 376–385, 2011.
- [9] L. Heirendt, S. Arreckx, T. Pfau, S. N. Mendoza, A. Richelle, A. Heinken, H. S. Haraldsdottir, S. M. Keating, V. Vlasov, J. Wachowiak *et al.*, "Creation and analysis of biochemical constraint-based models: the cobra toolbox v3. 0," 2017.
- [10] C. S. Henry, M. DeJongh, A. A. Best, P. M. Frybarger, B. Linsay, and R. L. Stevens, "High-throughput generation, optimization and analysis of genome-scale metabolic models," *Nature biotechnology*, vol. 28, no. 9, p. 977, 2010.
- [11] M. A. Henson and T. J. Hanly, "Dynamic flux balance analysis for synthetic microbial communities," *IET Systems Biology*, vol. 8, pp. 214–229, 2014.
- [12] J. L. Hjersted, M. A. Henson, and R. Mahadevan, "Genome-scale analysis of *Saccharomyces cerevisiae* metabolism and ethanol production in fed-batch culture," vol. 97, pp. 1190–1204, 2007.
- [13] K. Höffner, S. M. Harwood, and P. I. Barton, "A reliable simulator for dynamic flux balance analysis," *Biotechnology and Bioengineering*, vol. 110, no. 3, pp. 792–802, 2013.
- [14] K. A. Johnson and R. S. Goody, "The original michaelis constant: translation of the 1913 michaelis–menten paper," *Biochemistry*, vol. 50, no. 39, pp. 8264–8269, 2011.
- [15] P. D. Karp, S. M. Paley, M. Krummenacker, M. Latendresse, J. M. Dale, T. J. Lee, P. Kaipa, F. Gilham, A. Spaulding, L. Popescu *et al.*, "Pathway tools version 13.0: integrated software for pathway/genome informatics and systems biology," *Briefings in bioinformatics*, vol. 11, no. 1, pp. 40–79, 2009.
- [16] R. Luo, H. Wei, L. Ye, K. Wang, F. Chen, L. Luo, L. Liu, Y. Li, M. J. C. Crabbe, L. Jin, Y. Li, and Y. Zhong, "Photosynthetic metabolism of *C3* plants shows highly cooperative regulation under changing environments: A systems biological analysis," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 847–852, 2009.
- [17] R.-Y. Luo, S. Liao, G.-Y. Tao, Y.-Y. Li, S. Zeng, Y.-X. Li, and Q. Luo, "Dynamic analysis of optimality in myocardial energy metabolism under normal and ischemic conditions," *Molecular Systems Biology*, vol. 2, 2006.
- [18] R. Mahadevan, J. S. Edwards, and F. J. Doyle, "Dynamic Flux Balance Analysis of Diauxic Growth in *Escherichia coli*," *Biophysical Journal*, vol. 83, pp. 1331–1340, 2002.
- [19] N. V. NARENDRANATH, S. H. HYNES, K. C. THOMAS, and W. M. INGLEDEW, "Effects of lactobacilli on yeast-catalyzed ethanol fermentations," *APPL. ENVIRON. MICROBIOL.*, vol. 63, no. 11, p. 6, 1997.
- [20] R. A. Notebaart, F. H. Van Enckevort, C. Francke, R. J. Siezen, and B. Teusink, "Accelerating the reconstruction of genome-scale metabolic networks," *BMC bioinformatics*, vol. 7, no. 1, p. 296, 2006.
- [21] J. D. Orth, I. Thiele, and B. Palsson, "What is flux balance analysis?" *Nature Biotechnology*, vol. 28, no. 3, pp. 245–248, 2010.
- [22] B. Palsson, *Systems biology*. Cambridge university press, 2015.
- [23] F. Pizarro, C. Varela, C. Martabit, C. Bruno, J. R. Pérez-Correa, and E. Agosin, "Coupling kinetic expressions and metabolic networks for predicting wine fermentations," *Biotechnology and Bioengineering*, vol. 98, pp. 986–998, 2007.
- [24] J. Sainz, F. Pizarro, J. R. Pérez-Correa, and E. Agosin, "Modeling of yeast metabolism and process dynamics in batch fermentation: Modeling yeast metabolism," *Biotechnology and Bioengineering*, vol. 81, pp. 818–828, 2003.
- [25] F. Santos, J. Boele, and B. Teusink, "A practical guide to genome-scale metabolic models and their analysis." Elsevier, 2011, vol. 500, pp. 509–532.
- [26] B. Teusink, A. Wiersma, D. Molenaar, C. Francke, W. M. de Vos, R. J. Siezen, and E. J. Smid, "Analysis of growth of *Lactobacillus plantarum* wcf1 on a complex medium using a genome-scale metabolic model," *Journal of Biological Chemistry*, vol. 281, no. 52, pp. 40 041–40 048, 2006.
- [27] K. Zhuang, E. Ma, D. R. Lovley, and R. Mahadevan, "The design of long-term effective uranium bioremediation strategy using a community metabolic model," *Biotechnology and Bioengineering*, vol. 109, pp. 2475–2483, 2012.
- [28] K. Zhuang, M. Izallalen, P. Mouser, H. Richter, C. Risso, R. Mahadevan, and D. R. Lovley, "Genome-scale dynamic modeling of the competition between *Rhodospirillum rubrum* and *Geobacter* in anoxic subsurface environments," *The ISME Journal*, vol. 5, pp. 305–316, 2011.
- [29] A. R. Zomorodi and D. Segrè, "Synthetic ecology of microbes: Mathematical models and applications," *Journal of Molecular Biology*, vol. 428, pp. 837–861, 2016.