An Assessment of the Production Capacity of Incorporating Two Synthetic Carbon Dioxide Fixation Pathways within the Escherichia Coli Model

by

Barry Chen

Department of Chemical Engineering & Applied Chemistry, University of Toronto

December 2023

Table of contents

1 Abstract

In this project, the E. coli model, particularly the iML1515 model from BiGG Models, is studied with a focus on enhancing its production capability. The approach includes the integration of two synthetic CO_2 fixation pathways: the synthetic reductive glycine (rGlyP) pathway and the synthetic acetyl-CoA (SACA) pathway. The goal of the project is to equip the modified model with the ability of yielding two supplementary products, such as fatty acids and lactate, while simultaneously maintaining the growth of the organism. This work not only explores the feasibility of production within the context of the modified metabolic model but also assesses the potential for achieving carbon neutrality through biological methodologies.

2 Introduction

2.1 E. coli and bioproduction

In the context of biopharmaceutical production, the expression of numerous recombinant proteins is employed for the treatment of various diseases or serves as the medium for testing new and novel drugs. Among the various choices of hosts, E. coli stands out as one of the most popular hosts for the expression of diverse recombinant proteins, thanks to its rapid growth, high product yield, cost-effectiveness, and an easily scalable process [1]. Besides its applications in the biopharmaceutical field, E. coli holds great potential in various other domains, such as food engineering, biofuel development, and other biotechnology applications, addressing life-saving challenges, advancing biotechnology, and aligning with global market trends.

2.2 E. coli and carbon footprint

Carbon net-zero has been garnering increasing attention in recent years. In the petrochemical industry, there is a trend of replacing traditional energy and fossil-based products with bulk and fine chemicals produced through bioprocesses to decrease carbon emissions and mitigate negative environmental impacts. Intensive studies have been conducted on E. coli, demonstrating its ability to produce various products through a systematic and integrated approach, such as gene deletion [2]. As biotechnology continues to advance, it is anticipated to promote a greener and more sustainable biochemical industry, producing chemicals with renewable feedstock, high selectivity, and diversity.

3 Methods

3.1 Synthetic reductive glycine pathway

Based on the descriptions in the journal articles, the reactions of fixing CO_2 to form lactate, the product from the integration of the rGlyP pathway, are summarized below. Firstly, CO_2 is converted to formic acid through an electrochemical reaction catalyzed by CdS-CNT (Reaction 3.1.1); then, it undergoes several core reactions in the rGlyP pathway to form pyruvate (Reactions 3.1.2 – 3.1.7), catalyzed by the corresponding enzymes. Lastly, pyruvate is converted to lactate as the final product (Reaction 3.1.8) [3][4][5][7][8]. Native reactions in the iML1515, such as aldehyde dehydrogenase (Reaction 3.1.9) and succinate-semialdehyde dehydrogenase (Reaction 3.1.10), are chosen as the energy source (proton source).

$$CO_2 + 2\,H^+ + 2e^- \xrightarrow{CdS-CNT} CH_2O_2 \tag{3.1.1}$$

$$CH_2O_2 + THF + ATP \xleftrightarrow{MeFtfL} ADP + Pi + 10\text{-formyl-THF} \tag{3.1.2}$$

$$10\text{-formyl-THF} \xleftrightarrow{MeFch} H_2O + 5,10\text{-methenyl-THF} \tag{3.1.3}$$

$$5,10\text{-methenyl-THF} + 2\,NADPH \xleftrightarrow{MeMdtA} 2\,NADP^+ + 5,10\text{-methylene-THF} \tag{3.1.4}$$

$$5,10\text{-methylene-THF} + CO_2 + NH_3 + 2\,NADH \xleftrightarrow{EcGcvT/P/H} 2\,NAD^+ + THF + \text{glycine} \tag{3.1.5}$$

$$\text{glycine} + 5,10\text{-methylene-THF} + H_2O \xleftrightarrow{EcGlyA} THF + \text{serine} \tag{3.1.6}$$

$$\text{serine} \xleftrightarrow{\text{MeSdaA}} \text{NH}_3 + \text{pyruvate} \quad (3.1.7)$$

$$\text{pyruvate} + 2\ \text{NADH} \xleftrightarrow{\text{MeLdhA}} 2\ \text{NAD}^+ + \text{lactate} \quad (3.1.8)$$

$$\text{C}_4\text{H}_8\text{O} + \text{NAD}^+ + \text{H}_2\text{O} \longrightarrow \text{C}_4\text{H}_7\text{O}_2 + 2\ \text{H}^+ + \text{NADH} \quad (3.1.9)$$

$$\text{C}_4\text{H}_5\text{O}_3 + \text{NADP} + \text{H}_2\text{O} \longrightarrow \text{C}_4\text{H}_4\text{O}_4 + 2\ \text{H}^+ + \text{NADPH} \quad (3.1.10)$$

3.2 Synthetic acetyl-CoA pathway

According to the findings of the journal articles, the reactions of fixing CO_2 via the SACA pathway while generating dodecanoate, a medium-chain fatty acid, are listed below. A series of reactions starts by converting CO2 to methanol through a CO2 hydrogenation reaction catalyzed by ZnO (Reaction 3.2.1). Methanol then undergoes the essential reactions in the SACA pathway to form acetyl-CoA (Reactions 3.2.2 – 3.2.5) [6][9]. Finally, acetyl-CoA formed goes through the native fatty acid synthase to produce the final product. No energy source is added since no proton gets consumed during the following reaction processes.

$$\text{CO}_2 + 3\ \text{H}_2 \xrightarrow{\text{ZnO}} \text{methanol} + \text{H}_2\text{O} \quad (3.2.1)$$

$$2\ \text{methanol} + \text{O}_2 \xleftrightarrow{\text{AOX}} 2\ \text{formaldehyde} \quad (3.2.2)$$

$$2\ \text{formaldehyde} \xleftrightarrow{\text{GALS}} \text{glycolaldehyde} \quad (3.2.3)$$

$$\text{glycolaldehyde} + \text{Pi} \xleftrightarrow{\text{ACPS}} \text{acetyl-phosphate} + \text{H}_2\text{O} \quad (3.2.4)$$

$$\text{acetyl-phosphate} + \text{CoA} \xleftrightarrow{\text{PTA}} \text{acetyl-CoA} + \text{Pi} \quad (3.2.5)$$

3.3 Model design

The work was conducted in Google Colab using core Python packages, including cobra, straindesign, and scipy. The base model used is iML1515 from BiGG. Three modified models were investigated, namely, the base model integrated with the rGlyP pathway, the base model integrated with the SACA pathway, and the base model integrated with both pathways. For each tested model, metabolites and reactions were added as necessary, and the maximum biomass growth rate and product generation in the absence of glucose were computed. Phenotypic phase planes were plotted with and without constraints, followed by knock-out analysis. If a knockout strategy was identified, subsequent analyses, such as production prediction using MOMA, were performed.

4 Results

4.1 Integrating the rGlyP pathway

Two constraints, used in both the generation of the feasible design space (green area) and the knock-out analysis, are formulated as follows:

$$\text{biomass} - 0.01 \bullet \text{lactate} \geq 0 \quad (4.1.1)$$

$$\text{lactate} + 0.05 \bullet (\text{CO}_2\ \text{exchange}) \leq 0 \quad (4.1.2)$$

Figure 1 shows that applying these constraints results in potential growth-coupled solutions. Figure 2 illustrates that the maximum biomass growth rate is approximately 2.25 mmol/h/gdw, while the maximum lactate synthesis rate is around 38.20 mmol/h/gdw.

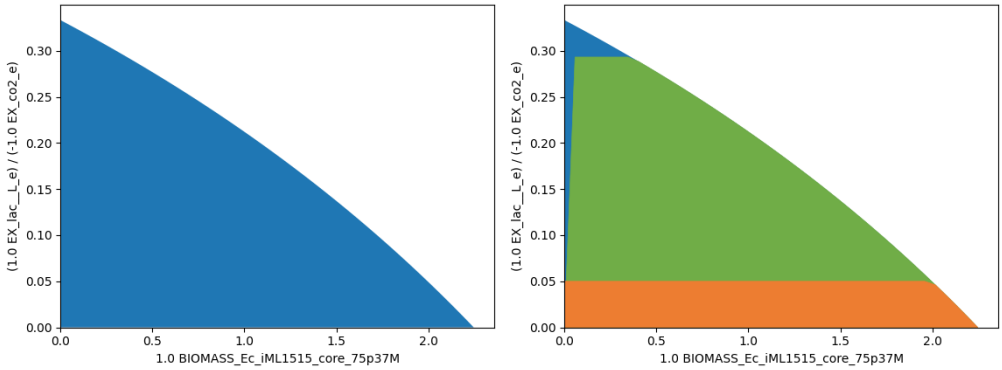


Figure 1: The phenotypic phase plane without constraints (left) and with constraints (right) for the model integrated with the rGlyP pathway.

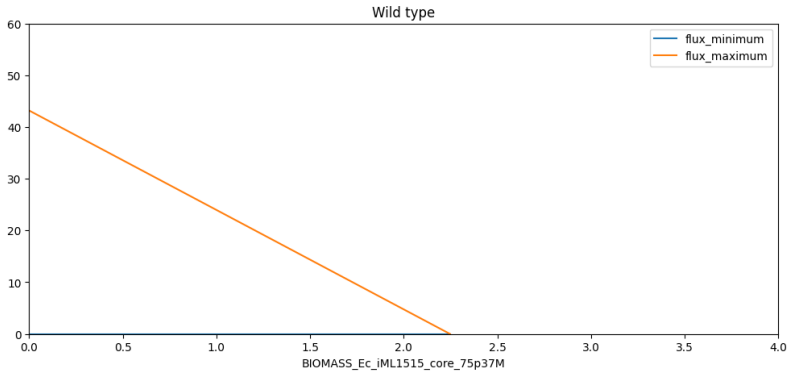


Figure 2: The production envelope (biomass vs. lactate) for the model integrated with the rGlyP pathway.

4.2 Integrating the SACA pathway

In this model, the two constraints take the following form:

$$\text{biomass} - 0.01 \bullet \text{dodecanoate} \geq 0 \qquad (4.2.1)$$

$$\text{dodecanoate} - 0.01 \bullet (\text{methanol generation}) \leq 0 \qquad (4.2.2)$$

Figure 3 demonstrates implementing these constraints could lead to growth-coupled production. Figure 4 illustrates that the maximum flux for biomass generation and the maximum flux for dodecanoate production are approximately 4.81 mmol/h/gdw and 18.72 mmol/h/gdw, respectively.

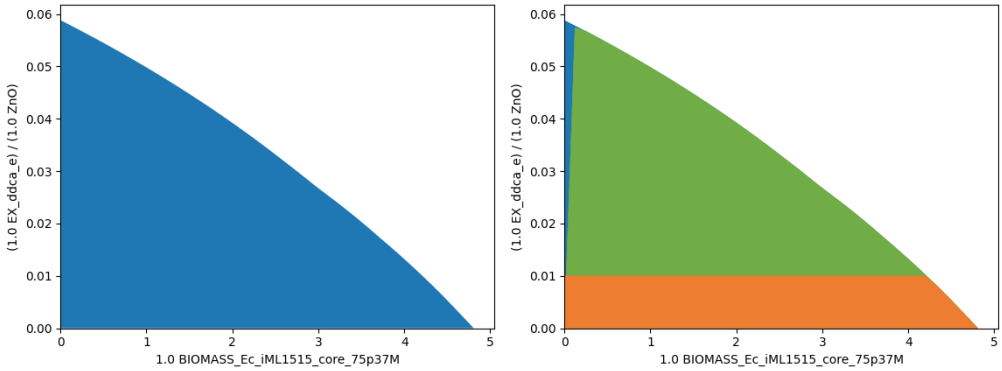


Figure 3: The phenotypic phase plane without constraints (left) and with constraints (right) for the model integrated with the SACA pathway.

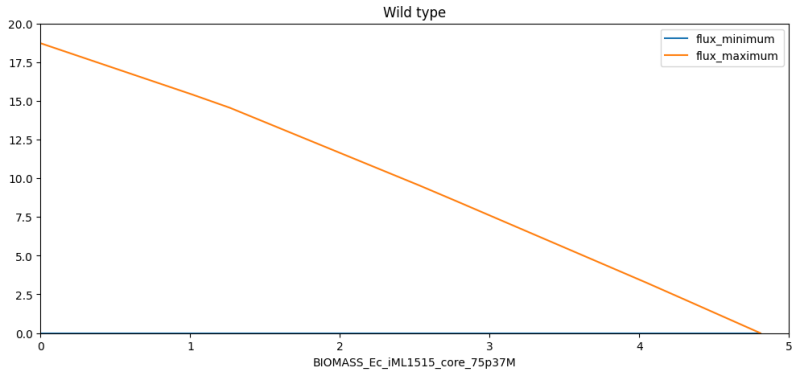


Figure 4: The production envelope (biomass vs. dodecanoate) for the model integrated with the SACA pathway.

4.3 Integrating both pathways

In the last model, the same conclusion can be drawn from Figure A.1. Figure A.2 illustrates that the maximum rates for biomass generation, lactate production, and dodecanoate synthesis are approximately 6.60 mmol/h/gdw, 169.39 mmol/h/gdw, and 27.60 mmol/h/gdw, respectively. Refer to the appendix for figures for this section.

5 Discussion

According to the results, growth-coupled production of the two selected products is possible with the corresponding constraints. However, the desired state was not achieved, and several factors contribute to this outcome.

The most evident factor is the hyperparameters in the strain design, such as the user-defined constraints used for model suppression and protection. While these constraints worked for the phenotypic phase plane plots, they may not be suitable for the strain design, which involves a more in-depth analysis of the tested models. Another hyperparameter that could impact the results is the time limit. In my testing cases, it was set to 1000 seconds. However, based on the results from a journal article where the time limit was set to 1 week, it may imply that the time limit for my testing models was not long enough for the strain design optimization algorithm to find a feasible solution [10].

Another plausible factor is the potential need for knock-in rather than knock-out in the experimental design. It is likely that the model may be missing crucial reactions essential for promoting growth-coupled production. Other factors to consider includes the selection of the base model, the chosen products, the requirement for more robust objective functions that account for product yield, substrate-specific productivity, and coupling strength, as well as the necessity for utilizing more recently developed tools such as OptKnock, OptGene, and OptCouple [10][11].

Reference

[1] M. N. Baeshen et al., “Production of biopharmaceuticals in E. coli: Current scenario and future perspectives,” Journal of Microbiology and Biotechnology, vol. 25, no. 7, pp. 953–962, 2015. doi:10.4014/jmb.1412.12079

[2] X. Chen et al., “Metabolic Engineering of Escherichia coli: A sustainable industrial platform for bio-based chemical production,” Biotechnology Advances, vol. 31, no. 8, pp. 1200–1223, 2013. doi:10.1016/j.biotechadv.2013.02.009

[3] N. J. Claassens *et al.*, “Engineering the reductive glycine pathway: A promising synthetic metabolism approach for C1-assimilation,” *One-Carbon Feedstocks for Sustainable Bioproduction*, pp. 299–350, 2022. doi:10.1007/10_2021_181

[4] B. Dronsella *et al.*, *Engineered synthetic one-carbon fixation exceeds yield of the calvin cycle*, 2022. doi:10.1101/2022.10.19.512895

[5] S. A. Al-Tamreh *et al.*, “Electroreduction of carbon dioxide into formate: A comprehensive review,” *ChemElectroChem*, vol. 8, no. 17, pp. 3207–3220, 2021. doi:10.1002/celc.202100438

[6] X. Lu *et al.*, “Constructing a synthetic pathway for acetyl-coenzyme A from one-carbon through enzyme design,” *Nature Communications*, vol. 10, no. 1, 2019. doi:10.1038/s41467-019-09095-z

[7] S. Kim *et al.*, “Growth of E. coli on formate and methanol via the reductive glycine pathway,” *Nature Chemical Biology*, vol. 16, no. 5, pp. 538–545, 2020. doi:10.1038/s41589-020-0473-5

[8] S. Kim *et al.*, “Optimizing E. coli as a formatotrophic platform for bioproduction via the reductive glycine pathway,” *Frontiers in Bioengineering and Biotechnology*, vol. 11, 2023. doi:10.3389/fbioe.2023.1091899

[9] S. Kattel, P. J. Ramírez, J. G. Chen, J. A. Rodriguez, and P. Liu, “Active sites for CO2 hydrogenation to methanol on Cu/ZnO catalysts,” *Science*, vol. 355, no. 6331, pp. 1296–1299, 2017. doi:10.1126/science.aal3573

[10] A. M. Feist et al., “Model-driven evaluation of the production potential for growth-coupled products of escherichia coli,” Metabolic Engineering, vol. 12, no. 3, pp. 173–186, 2010. doi:10.1016/j.ymben.2009.10.003

[11] K. Jensen, V. Broeken, A. S. Hansen, N. Sonnenschein, and M. J. Herrgård, “OptCouple: Joint simulation of Gene Knockouts, insertions and medium modifications for prediction of growth-coupled strain designs,” Metabolic Engineering Communications, vol. 8, 2019. doi:10.1016/j.mec.2019.e00087

Appendix

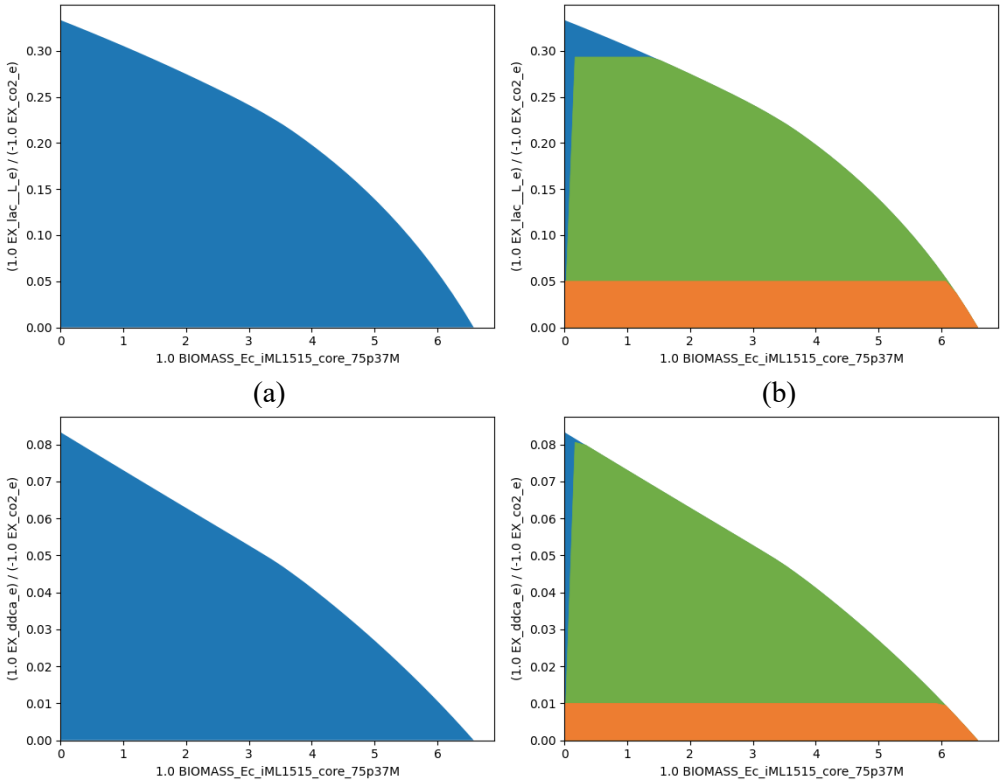
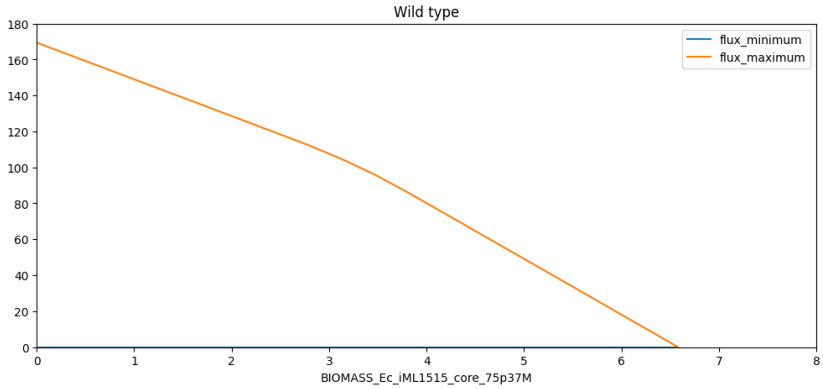
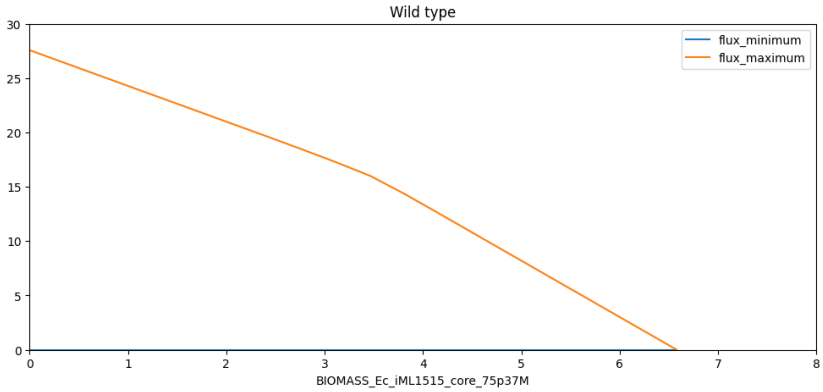


Figure A.1: The phenotypic phase plane without constraints (a, c) and with constraints (b, d) for the model integrated with two pathways. Constraints 4.1.1 and 4.1.2 were applied in (b). Constraints 4.2.1 and 4.2.2 were used in (d).



(a)



(b)

Figure A.2: The production envelopes for the model integrated with two pathways. (a) biomass vs. lactate. (b) biomass vs. dodecanoate.