

Model-guided *in silico* Overexpression of *adhE* Gene Predicts Increased Ethanol Production in *Escherichia coli* from Xylose

Bashir Sajo Mienda*, Mohd Shahir Shamsir

Bioinformatics Research Group (BIRG), Department of Biosciences & Health Sciences, Faculty of Biosciences and Medical Engineering, University Teknologi Malaysia, Malaysia

Copyright © 2015 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract *Escherichia coli* systems metabolic engineering has progressed significantly in guiding future metabolic engineering targets for the production of bioethanol using different carbon sources. However, the use of xylose as substrate coupled with overexpression of *E. coli* native *adhE* using parsimonious enzyme usage (pFBA) algorithm with the OptFlux interface remained largely underexplored. Here, we show for the first time that *in silico* overexpression of the *adhE* and under expression of *pflA* lead to 2 fold increase in ethanol production from xylose using the *E. coli* GEM. The results indicate that 2 NADH molecules have been generated by under expression of *pflA* and *ldhA*. Furthermore, the triple overexpression of the native *adhE/b1241* using xylose as the substrate might have increased the consumption of NADH generated in the cell that lead to 2 fold increase in ethanol production with a growth rate that was 90.8% of the wild-type model. On the bases of these findings, we hypothesize that *E. coli* native *adhE* preferred xylose as substrate when overexpressed to achieve cellular redox balance by oxidizing NADH generated in increasing ethanol production. This study informs other studies that model-guided biological insight could be applied in identifying metabolic engineering targets, paving way for a comprehensive biological inquiry on the role of the *E. coli* native *adhE* overexpression in enhancing ethanol production using xylose as a solitary carbon source.

Keywords Alcohol Dehydrogenase, *E. coli* Genome Scale Model, Ethanol Production, OptFlux software, *In silico* Overexpression Prediction, Xylose, Model-guided Approach

1. Introduction

Bioethanol is currently considered as a single larger biotechnological commodity, which is currently used as a

biofuel [1]. Different carbon sources or substrates such as glucose and sucrose have been used for large scale production in *Escherichia coli* [1]. It was reported elsewhere that 9 billion gallons of ethanol were produced in 2008, and all emanates from corn based production [2] which may likely result to food shortage. Therefore, excellent source of sugars could be from lignocellulosic biomass for conversion into fuel ethanol. Lignocellulosic biomass is the most abundant in nature, which constitutes mainly cellulose and hemicellulose [3] that when decomposed releases two most dominant sugars such as glucose and xylose [3]. It is deemed imperative to design the biocatalysts that can be able to utilize the sugar components in lignocellulose and convert them to ethanol with high level of productivity in a defined mineral salt medium [2]. *E. coli* strains have been reported to utilize all the sugar components of lignocelluloses but ethanol is only a minor fermentation end products, in combination with mixed acid accumulation as the major fermentation product [1, 2].

Recent advances have been reported in native metabolic pathway engineering of *E. coli* using glucose or xylose without the foreign gene for enhanced ethanol production [4]. Redesigning of *E. coli* for ethanol production was largely reported to be dependent on the use of the traditional three approaches: Constructions of a metabolic pathway for ethanol production as the major fermentation end product, elimination of NADH consuming pathway genes/reactions and knocking out of genes responsible for side product formation [2]. One of the successful approaches of these strategies was the use of *Zymomonas mobilis* homoethanol pathways that was introduced into *E. coli*, which in turn enabled redox-balanced production of ethanol at high yield and productivity [5]. Another example of metabolic engineering intervention is inactivation of methylglyoxal synthetase leading to co-metabolism of different five sugar mixture (mannose, glucose, arabinose, galactose and xylose) to ethanol in *E. coli* [2, 6]. After a series of adaptive laboratory evolution, they could be able to achieve a

concurrent metabolism of a complex combination of the aforementioned sugars present in lignocellulosic biomass, where high yield and productivity is achieved using a mineral salt medium [2, 6]. This indicates the significance of substrate selectivity and /or combination thereof in enhancing ethanol production in metabolically engineered *E. coli* strains. A number of researches have been conducted considering different carbon sources to increase ethanol production in wild-type and metabolically engineered *E. coli* [1], but the use of xylose as the only solitary carbon source has not yet been fully elucidated. We used xylose as a carbon source to predict *in silico* over expression of the *E. coli* native alcohol dehydrogenase (*adhE*) to increase the production of ethanol in *E. coli* genome scale model.

In silico metabolic engineering of *E. coli* has been used to predict a series of metabolic engineering targets for increase ethanol production from glucose and gluconate [7], while xylose and glycerol [8] have been applied to predict gene knockouts using *E. coli* genome scale metabolic model [8]. The previous approach used algorithm called Minimization of Metabolic Adjustment (MOMA) for a metabolic gene knockout. While the current study used a variant of flux balance analysis called parsimonious enzyme usage (pFBA), which is specifically used for metabolic pathway genes over/under expression in *E. coli* genome-scale metabolic model to predict *in silico* cellular behavior after the introduction of genetic modification [9] with the OptFlux software platform [10].

Here, we report the implementation of an accurate genome scale metabolic model of *E. coli* [11] and pFBA algorithm [9] to broadly overexpress alcohol dehydrogenase (*adhE/b1241*) and down regulation of *ldhA* and *pflA* pathway genes that predicted increased ethanol production that is twofold higher than the wild-type model using xylose as the solitary carbon source. To the best of our knowledge, this is the first report describing the *in silico* overexpression of native *adhE/b1241* and down regulation of *ldhA/b1380* and *pflA/b0902* for increase ethanol production in *E. coli* using xylose as the substrate. These findings would pave way for a comprehensive model-guided experimental inquiry and/or novel biological discovery as to the role of the native *adhE* in oxidizing excess NADH generated in relation to ethanol production in *E. coli* from xylose substrate.

2. Materials and Method

2.1. Constraint based Modelling with OptFlux Software Platform

The OptFlux strain design methodology [10] using the parsimonious enzyme usage flux balance analysis (pFBA) was implemented within the OtFlux interface to over expressed *adhE/b1241* in *E. coli* genome-scale metabolic

model. pFBA is a bi-level linear programming optimization algorithm [9] that uses the genome scale metabolic model of *E. coli* K12 [11] for *in silico* gene and/or enzyme down or up regulations that is consistent with transcriptomic and proteomics data as described in their original documentation [9]. We implemented the constraint-based modeling for overexpression of *adhE/b1241* in the *E. coli* model *iAF1260* with Java programming within the framework of OptFlux as described previously [10].

The selected solitary carbon source was xylose unless otherwise stated, and the uptake rate of the xylose carbon source was constrained to a maximum of 20 mmol g DW⁻¹ h⁻¹. The oxygen uptake rate was considered to be 0.0 mmol g DW⁻¹ h⁻¹ as the simulation condition was anaerobic for fermentative ethanol production. These values were chosen based on slightly close experimental observation of anaerobic growth of *E. coli* [12-14]. All simulations of mutant models for *in silico* overexpression of *adhE/b1241* were constrained to a range of expression value(s) from 1-3 using the pFBA algorithm as described elsewhere [9] In addition, the pFBA was equally implemented to under expressed *pflA* and *ldhA* pathway genes by imposing a zero flux constraints in the mutant models constructed in this study. While the growth of the wild-type model was performed using the normal FBA for maximization of biomass.

3. Results and Discussions

E. coli carries out mixed acid fermentation reactions in the production of ethanol among other end products. Ethanol is catalyzed principally by two main distinctive step reaction using alcohol dehydrogenase (*adhE*). Conversion of acetyl-coA through acetaldehyde into ethanol with the regeneration of two molecules of NAD⁺ is achieved by this enzyme (*adhE*) [1] (see Fig 1). In practice, two molecules of NADH are consumed to produce 1 molecule of ethanol in *E. coli* under anaerobic condition (see fig 1). While on the other hand, reduction of pyruvate catalyzed by soluble lactate dehydrogenase (*ldhA*) consumes 1 molecule of NADH that is meant for the production D-lactate [1] (see fig 1).

Furthermore, incomplete oxidation of sugars occurs when anaerobic (anoxic) condition is established, which in turn lead to down regulation of citric acid cycle to channel the carbon flux towards the production of ethanol in *E. coli*, and as a result acetate is formed as the main product [1]. Pyruvate dehydrogenase is shut down and it is being replaced by pyruvate formate lyase (*pflA*), which preserves an additional NADH that could be used during consumption, and subsequently catalysis that could lead to the formation of formate and acetyl-CoA [1].

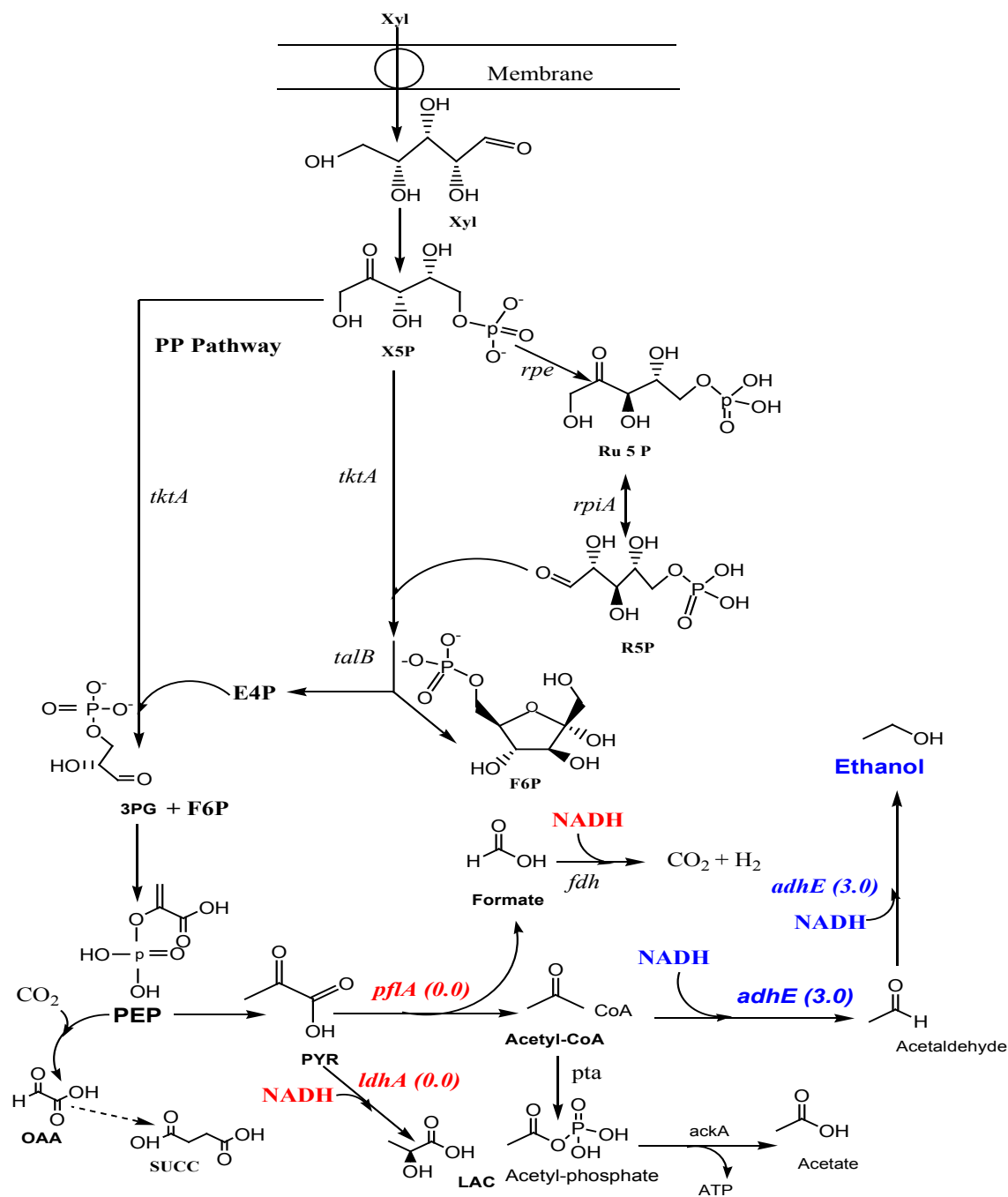


Figure 1. Central metabolic pathway leading to ethanol production in *E. coli* from xylose. Shown are the pathways related to the synthesis of ethanol (partially adopted from ref [3]). The under express/down regulated genes are shown in red (*pflA* and *ldhA*, which carries zero fluxes). The NADH in red are preserved to be used for ethanol production by down regulating the activities of *pflA* and *ldhA*. The overexpressed genes are shown in blue (*adhE* which carries 3.0 fluxes). The NADH in blue are normally produced during the biosynthesis of ethanol using *adhE* as a catalyst. Broken arrows in the pathway indicate multiple steps. The genes encoding the designated enzymes that catalyzes the pathway reaction is indicated as follows: *adhE*, alcohol dehydrogenase; *pflA*, pyruvate formate lyase A; *pta*, phosphate acetyltransferase; *rpe*, ribulose-5-phosphate 3-epimerase; *rpiA*, ribulose-5-phosphate isomerase A; *talB*, transaldolase B; *tktA*, transketolase A; *fdh*, formate dehydrogenase; *ackA*, acetate kinase A. Abbreviations: Xyl, xylose; OAA, oxaloacetate; LAC, lactate; SUCC, succinate; PEP, phosphoenolpyruvate; PYR, pyruvate; Ru5P, ribulose 5-phosphate; R5P, ribose 5-phosphate; E4P, erythrose 4-phosphate; X5P, Xylose 5-phosphate; F6P, Fructose 6-phosphate; 3PG, 3-phosphoglyceraldehyde;

In this study, we used the knowledge based reported above and reason that overexpression of *adhE* and down regulation of the *ldhA* and *pflA* might increase the pool of NADH in the cell that could direct more carbon flux towards ethanol production in *E. coli*. The role of *pflA* in the formation of formate from pyruvate and subsequent conversion of the formate to CO₂ and H₂ by consumption of NADH could be minimized by down regulating and/or inactivation of the *pflA* (see table 1 and Fig 1). On the other hand, the formation of Lactate, as an NADH oxidizing pathway has been minimized by the inactivation and/or down regulation of *ldhA/b1380*. This leads to an additional molecule of NADH that could be used during ethanol production in *E. coli*.

Table 1. *E. coli* design properties for enhanced ethanol flux from xylose substrate

E. coli model strains	Over expressed genes	Under expressed genes	Growth Rate (h ⁻¹)	% growth rate	Ethanol Flux (mmol g DW ⁻¹ h ⁻¹)	% Ethanol flux	Acetate Flux (mmol g DW ⁻¹ h ⁻¹)	Succinate Flux (mmol g DW ⁻¹ h ⁻¹)
WT (Feist model)	-		0.53601481	100	20.38003	100	20.72615	0.17889
BMR 101	adhE/b1241 (1)	pflA (0.0) ldhA (0.0)	0.50891133	94.9	31.60845	155	9.91272	0.16984
BMR102	adhE/b1241 (2)	pflA (0.0) ldhA (0.0)	0.50891133	94.9	31.60845	155	9.91272	0.16984
BMR 103	adhE/b1241 (3)	pflA (0.0) ldhA (0.0)	0.48680531	90.8	40.7645	200	0.28283	0.16246

E. coli systems metabolic engineering has evolved to use genome-scale metabolic reconstruction (GEM) for the prediction of accurate metabolic engineering targets that could guide future experimental studies. These models (GEM) have been developed, tested and validated against previously established experimental data [11, 15-22], enabling a model-guided biological inquiry and/or providing novel biological insight. Model-guided approach has been particularly useful in experimental and novel biological discovery. This approach led to the discovery of two new reactions carried out by the classical enzymes phosphofructokinase and aldolase in *E. coli* [23, 24]. In the current study, we applied the *E. coli* GEM using an algorithm called parsimonious enzyme usage (pFBA) under the OptFlux software platform to *in silico* over/under expressed a number of genes to increase the pool of intracellular NADH for the increase production of ethanol in *E. coli*.

The results obtained in strain BMR101 after a single over expression of the *adhE* and under expression/down regulation of the *pflA* and *ldhA* lead to 155% increase in ethanol production from the wide-type control model (see table 1). The mutant model (BMR 101) maintained a growth rate that is 94.9% of the wild-type. On the other hand, the same ethanol flux and growth rate was achieved with mutant model BMR102 after a double overexpression of *adhE* and under expression of the *pflA* and the *ldhA* (see table). The similarities in the flux for ethanol and growth rates of these two mutant models reported in this study could be attributed to the reactome of the *E. coli* GEM. This is because it was established that the accounting exercise of the GEM is predicated on the accurate reconstructions of its reactome [23, 25], although GEM was established to be capable of performing accurate, true-positive and true-negative predictions that generally exceeds 80-90% of the total models previously examined [23].

The results obtained in mutant model BMR103 for the *in silico* triple over expression of *adhE/b1241* and under expression of the *pflA* and *ldhA* showed 200% (2 fold) increase in ethanol production in *E. coli* from xylose and 90.8% growth rate relative to the wild-type model (see table 1). We reason that under expression of *pflA* might have generated 1 molecule of NADH, by reducing the formation of formate that usually served as a precursor for formate dehydrogenase (*fdh*) to oxidize NADH to produce CO₂ and H₂ under anaerobic condition in *E. coli* (see Fig 1). Although

pyruvate formate lyase B (*pflB*) was reported to be activated and produced acetate and probably formate under anaerobic condition using gluconate as substrate in *E. coli* [26]. The activity of this enzyme was previously hypothesized to be substrate dependent under anaerobic condition [7]. It was already reported that deletion of *pflB* can block the conversion of pyruvate to acetate and to ethanol with an impaired growth profile [27], as such, it is considered not a suitable target for engineering when using xylose as a substrate under anaerobic condition in *E. coli*. On the other hand, down regulation of *ldhA* might have equally generated 1 molecule of NADH that would have been used for lactate production (see Fig 1). In addition, the overexpression of *adhE/b1241* using the pFBA algorithm reached its maximum limits at 3 beyond which the *E. coli* mutant model could not grow under anaerobic conditions (data not shown). This indicates that the model can only achieve cellular redox balance at expression level of three (3). While at the expression level higher than three (3), the cell generates excess NADH that cannot be recycled due to lack of an electron sink [28]. On the bases of these observations, we hypothesize that the consumption of the additional 2 molecules of NADH generated in the cell can only be achieved via xylose metabolism with triple overexpression of native *adhE* in *E. coli*. This increase in ethanol flux has not been observed when glucose or glycerol is used as carbon sources (data not shown). The triple over expression of *adhE* might have utilized the 2 NADH generated by down regulating *ldhA* and *pflA* to increase an additional molecule of ethanol that could not be obtained using other solitary carbon sources such as glucose and glycerol. On the bases of these findings, we can hypothesize that the *E. coli* native *adhE/b1241* preferred xylose as substrate when over expressed to oxidize NADH generated in increasing ethanol production. These findings could serve as a starting point for a comprehensive model-guided experimental and/or novel biological inquiry about the use of xylose substrate in relation to native *adhE* overexpression for increased NADH utilization and ethanol production in *E. coli*.

REFERENCES

- [1] Förster, A.H. and J. Gescher, *Metabolic engineering of*

- Escherichia coli* for production of mixed-acid fermentation end products. *Frontiers in Bioengineering and Biotechnology*, 2014. **2**(16): p. 1-12.
- [2] Jarboe, L.R., et al., *Metabolic Engineering for Production of Biorenewable Fuels and Chemicals: Contributions of Synthetic Biology*. *Journal of Biomedicine and Biotechnology*, 2010. **2010**: p. 1-18.
- [3] Chiang, C.-J., et al., *Systematic Approach To Engineer Escherichia coli Pathways for Co-utilization of a Glucose–Xylose Mixture*. *Journal of Agricultural and Food Chemistry*, 2013. **61**(31): p. 7583-7590.
- [4] Kim, Y., L.O. Ingram, and K.T. Shanmugam, *Construction of an Escherichia coli K-12 mutant for homoethanologenic fermentation of glucose or xylose without foreign genes*. *Appl Environ Microbiol*, 2007. **73**(6): p. 1766-71.
- [5] L. O. INGRAM, et al., *Genetic Engineering of Ethanol Production in Escherichia coli*. *Appl. Environ. Microbiol*, 1987. **53**(10): p. 2420-2425.
- [6] Yomano, L.P., et al., *Deletion of methylglyoxal synthase gene (mgsA) increased sugar co-metabolism in ethanol-producing Escherichia coli*. *Biotechnology Letters*, 2009. **31**(9): p. 1389-1398.
- [7] Mienda, B.S., et al., *IN SILICO METABOLIC ENGINEERING INTERVENTIONS OF ESCHERICHIA COLI FOR ENHANCED ETHANOL PRODUCTION, BASED ON GENE KNOCKOUT SIMULATION*. *IIOABJ*, 2014. **5**(2): p. 16-23.
- [8] Mienda, B.S. and M.S. Shamsir, *In silico Gene knockout metabolic interventions in Escherichia coli for Enhanced Ethanol production on Glycerol*. *RES J PHARM BIOL CHEM SCI*, 2014. **5**(4): p. 964-974.
- [9] Lewis, N.E., et al., *Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models*. *Mol Syst Biol*, 2010. **6**: p. 390.
- [10] Rocha, I., et al., *OptFlux: an open-source software platform for in silico metabolic engineering*. *BMC Syst Biol*, 2010. **4**: p. 45.
- [11] Feist, A.M., et al., *A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information*. *Mol Syst Biol*, 2007. **3**: p. 121.
- [12] Varma, A., B.W. Boesch, and B.O. Palsson, *STOICHIOMETRIC INTERPRETATION OF ESCHERICHIA-COLI GLUCOSE CATABOLISM UNDER VARIOUS OXYGENATION RATES*. *Applied and Environmental Microbiology*, 1993. **59**(8): p. 2465-2473.
- [13] Edwards, J.S., R.U. Ibarra, and B.O. Palsson, *In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data*. *Nat Biotechnol*, 2001. **19**(2): p. 125-30.
- [14] Fischer, E., N. Zamboni, and U. Sauer, *High-throughput metabolic flux analysis based on gas chromatography–mass spectrometry derived ¹³C constraints*. *Analytical Biochemistry*, 2004. **325**(2): p. 308-316.
- [15] Schellenberger, J., et al., *Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0*. *Nat Protoc*, 2011. **6**(9): p. 1290-307.
- [16] Becker, S.A., et al., *Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox*. *Nat Protoc*, 2007. **2**(3): p. 727-38.
- [17] King, Z.A. and A.M. Feist, *Optimizing Cofactor Specificity of Oxidoreductase Enzymes for the Generation of Microbial Production Strains—OptSwap*. *Industrial Biotechnology*, 2013. **9**(4): p. 236-246.
- [18] Feist, A.M., et al., *Modeling methanogenesis with a genome-scale metabolic reconstruction of Methanosarcina barkeri*. *Mol Syst Biol*, 2006. **2**: p. 2006 0004.
- [19] Feist, A.M., et al., *Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli*. *Metab Eng*, 2010. **12**(3): p. 173-86.
- [20] Lee, J., et al., *Genome-scale reconstruction and in silico analysis of the Clostridium acetobutylicum ATCC 824 metabolic network*. *Appl Microbiol Biotechnol*, 2008. **80**(5): p. 849-62.
- [21] Orth, J.D., et al., *A comprehensive genome-scale reconstruction of Escherichia coli metabolism--2011*. *Mol Syst Biol*, 2011. **7**: p. 535.
- [22] McCloskey, D., B.O. Palsson, and A.M. Feist, *Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli*. *Mol Syst Biol*, 2013. **9**: p. 661.
- [23] Monk, J. and B.O. Palsson, *Genetics. Predicting microbial growth*. *Science*, 2014. **344**(6191): p. 1448-9.
- [24] Yamamoto, N., et al., *Update on the Keio collection of Escherichia coli single-gene deletion mutants*. *Molecular Systems Biology*, 2009. **5**.
- [25] Monk, J., J. Nogales, and B.O. Palsson, *Optimizing genome-scale network reconstructions*. *Nature Biotechnology*, 2014. **32**(5): p. 447-452.
- [26] Hildebrand, A., et al., *Engineering Escherichia coli for improved ethanol production from gluconate*. *J Biotechnol*, 2013. **168**(1): p. 101-6.
- [27] Hasona, A., et al., *Pyruvate formate lyase and acetate kinase are essential for anaerobic growth of Escherichia coli on xylose*. *J Bacteriol*, 2004. **186**(22): p. 7593-600.
- [28] Shen, C.R., et al., *Driving Forces Enable High-Titer Anaerobic 1-Butanol Synthesis in Escherichia coli*. *Applied and Environmental Microbiology*, 2011. **77**(9): p. 2905-2915.