ISSN (Online) : 0974-5645
ISSN (Print) : 0974-6846
Indian Journal of Science and Technology *Vol 8(2), 172–177, January 2015* DOI : 10.17485/ijst/2015/v8i1/56501

In silico Evaluation of the Effect of pfl Gene Knockout on the Production of D-lactate by *Escherichia coli* Genome Scale Model using the OptFlux Software Platform

Bashir Sajo Mienda*, **Mohd Shahir Shamsir and Faezah Mohd Salleh**

Bioinformatics Research Group (BIRG), Biosciences & Health Sciences Department, Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, Skudai 81310 Johor Bahru, Malaysia; bsmienda@gmial.com

Abstract

The increase availability of genome scale metabolic models of *Escherichia coli* and computational successes is revolutionizing the field of metabolic engineering and synthetic microbiology. *E. coli* has been experimentally established to produce D-lactate under micro-aerobic conditions when pyruvate formate lyase (PFL) genes are knocked out. However, investigation on the in silico prediction and for evaluation of the effect of PFL genes knockout on the production of D-lactate using E. coli genome scale metabolic model with regulatory on/off minimization (ROOM) under the OptFlux software platform remained under explored. Here, we demonstrate that metabolic engineering strategies using the OptFlux software platform by gene knockout simulation of pflA/b0902, pflB/b0903, pflC/b3952 and pflD/b3951 have been predicted to increase D-lactate production in E. coli and hence maintaining a growth rate that is 96% of the wild-type model. The deletion of the PFL genes have been established to increase D-lactate production in *E. coli.* The results obtained in this study is in agreement with the previously established experimental studies. These findings suggests that the OptFlux software platform using ROOM as the simulation algorithm, can prospectively and effectively predict future metabolic engineering targets for increased D-lactate production in *E. coli* and/or other microbial chemical syntheses.

Keywords: D-lactate, *Escherichia coli* Model, Gene Knockout Simulation, Metabolic Engineering, OptFlux Software

1. Introduction

Lactic acid (lactate) and some its derivatives have many application in the food, pharmaceutical and polymer industries. *Escherichia coli* as a well-engineered chassis host used to have a significant metabolic shift between aerobic and anaerobic conditions as a result *E. coli* produces a mixture of organic acids such as formate, acetate, D-Lactate, succinate and ethanol under oxygen limited conditions without other electron acceptors[1]. D-lactate was known to be produced by *E. coli* under micro aerobic conditions when pyruvate formate lyase genes are knocked out[1]. D-lactate industrial production is beset with a number of challenges, which entails improved *E. coli* strain with increase production potential. Metabolic engineering as a discipline has been and will continue to provide engineered *E. coli strains* with great potential for industrial production of D-lactate. Although several laboratories have reported alternative biocatalysts[1, 3] many of which constitutes engineered *E. coli* strains that produce D- or L-lactate[1,4-6.] The identification of suitable engineering targets using *in silico* biotechnology approaches would offer an easy solution to guide future systems metabolic engineering for improved D-lactate production in *E. coli.*

The availability of genome scale models, particularly of *E. coli*[7, 8] and the use of *in silico* metabolic engineering software platform to predict and/or identify metabolic gene knockouts targets have received remarkable attention

*Author for correspondence

in recent years[9–11]. One of the notable examples are the use of the OptFlux software to predict metabolic engineering interventions using *E. coli* genome scale as described previously[11, 12]. OptKnock, is another software that was proposed to predict metabolic engineering for lactic acid production in *E. coli*[9, 13]. In addition, OptGene, was reported among other alternative methods that is capable of *in silico* gene knockout in genome scale metabolic models[14]. A peculiar drawback of the use of OptKnock and OptGene in metabolic engineering interventions is that they only use metabolic information, determining sets of reactions to be eliminated from the metabolic models, instead of sets of genes to knockout, which is the real purpose[9, 15]. Therefore, in order to have an appropriate mutant in the experimental setting, one needs to determine which sets of genes can lead to the elimination of a given set of reactions[9, 15]. This is because the rule – 1 gene: 1 enzyme: 1 reaction – was not universal[9]. However, there are many exceptions, such as isoenzymes, protein complexes, or enzymes that catalyze several reactions[9, 15].

On the other hand, *E. coli* systems metabolic engineering could be synergistically combine with computational tools and synthetic biology to offer promising solutions for industrial chemical productions. An example is seen in a computational tool/software interface for *in silico* metabolic engineering interventions and constraint based modelling called OptFlux[12]. The OptFlux software platform uses genome scale metabolic models of *E. coli* to predict the phenotype simulation of both the wild-type and the mutant strain using the method of flux balance analysis (FBA)[9–12]. The software has a peculiar feature of plug-in architecture, where an algorithm such as Regulatory on/off minimization of metabolic flux changes (ROOM) can be used to introduce genetic perturbations in the *E. coli* genome scale model, there by paving ways for model-guide experimental inquiry and/or understanding novel biological insight on the behavior of the mutant model after gene knockout[9, 12, 16, 17].

Computational tool such as Opt Knock was reported to be a reference algorithm for studying a number of metabolic gene knockout using *E. coli* genome scale model using bi-level optimization approach for the D-lactate production, but it does not allow non-linear objective function and need a considerable to compute a solution[13]. We previously reiterated how computational breakthroughs can help to redesign microbial chassis host for robust bioethanol production[18]. *In silico* metabolic engineering interventions for increased ethanol production from glucose and gluconate using the OptFlux software platform was reported[11]. In a similar study, enhanced ethanol production by model-guided in silico metabolic engineering in *E. coli* was predicted using glycerol and xylose as the main solitary carbon sources[10]. Furthermore, D-lactate production from glycerol was predicted using *E. coli* genome-scale model with the OptFlux software platform[9]. A similar software called Meta Flux Net was used to investigate in silico metabolic engineering targets by comparative genome approach by Lee and co-workers[19, 20] to increase succinate production in *E. coli*[19]. Nevertheless, we used the OptFlux software platform, with a more advanced *in silico* metabolic engineering capabilities and peculiar plug-in architecture in current work. To the best of our knowledge, we report for the first time the implementation of this software to predict gene knockout in *E. coli* iJ01366[8]. We used ROOM as the algorithm for simulation to predict whole cell's post perturbation behavior after gene knockout (*pflA/b0902, pflB/b0903, pflC/b3952 and pflD/b3951*) in relation to D-lactate production.

2. Materials and Methods

2.1 Model

The Metabolic reconstruction of *Escherichia coli* iJ01366[8] was used as a model for all the wild-type and mutant strains described in this study. The model was previously tested and validated against experimental data, and was shown to be capable of predicting accurate growth rates, metabolite excretion rates, and a growth phenotypes on a number of substrates and genetic conditions[7, 21, 22]. The substrates used in this study is glucose unless otherwise stated.

2.2 Flux Balance Analysis

OptFlux software, as an open source platform www.optflux.org[12] and a reference computational tool for metabolic engineering was used for the Flux Balance Analysis (FBA). Regulatory on/off minimization of metabolic flux changes after genetic perturbations (ROOM)[16] was used as a simulation method for gene knockouts, and it was implemented using the Java programming within the framework of the OptFlux as described elsewhere[9, 12]. All simulation of mutant strains and wild-type models were performed using the OptFluxv3.06.

The chosen solitary carbon source is glucose, and the uptake rate of the carbon source was constrained to a maximum of 20 mmolgDW^{-1}h$^{-1.}$ The oxygen uptake rate was considered to be 5 mmolgDW^{-1}h^{-1} as the simulation condition was micro-aerobic for fermentative production of D-lactate. These values were chosen based on slightly close experimental observation of micro-aerobic and anaerobic growth of *E. coli*[23–25].

2.3 Gene Knockout under the OptFlux Software Platform

Gene knockout simulation was conducted under the OptFlux software platform using ROOM[16] as simulation method. Flux Balance Analysis (FBA) was used for simulation of the wild-type model, which predicts metabolic flux distributions at steady state by using linear programming, while ROOM uses Mixed Integer Linear Programing (MILP) to find flux distribution that predicts the same constraint as FBA while minimizing the number of significant flux changes[16]. The algorithm accounted only for flux changes (0.001 flux prediction) that is considered significant[16]. This is because inherent noise some time exists in biological systems and by using small flux changes, reduced running time is achieved as described in their original documentation[16]. The wild-type model obtained from the Biomodels database[26], constructed by Orth[8] was designated as WT (Orth Model) and the mutant models/strains with pyruvate formate lyase single gene knockouts were designated as *pflA*$^{-}$, *pflB*$^{-}$, *pflC*$^{-}$ and *pflD*$^{-}$ strains respectively. The *in silico* gene knockouts simulation were run to completion using ROOM, as previously described in their original documentation[16].

3. Results and Discussion

Metabolism of *E. coli* experienced significant changes between aerobic and anaerobic conditions. Under oxygen limited conditions *E. coli* produces a mixture of organic acids such as formate, acetate, D-lactate, succinate and ethanol without other electron acceptors[1] pyruvate is mainly assimilated via Pyruvate Formate Lyase (PFL) and form formic acid and acetyl-CoA (AcCoA)[1]. It was previously established that knocking out of PFL genes (*pflA, pflB, pflC* and *pflD*) triggereda metabolic turn over towards the production of D-lactate under micro-aerobic conditions[1].

The *in silico* results of D-lactate production from glucose as substrate were indicated in Table 1 and Figure 1. It has been shown that the wild-type model (WT Orth) showed no D-lactate production when oxygen limited (micro-aerobic) condition was used for the simulation, indicating clearly that under micro-aerobic conditions, *E. coli* does not naturally produce D-lactate. On one hand the deletion of PFL genes in the *E. coli* central metabolism indicate a positive D-lactate production under micro-aerobic condition with a growth rate that is 96% of the wild-type model (see Table 1). While on the other hand, no D-lactate was produced when the wild-type model was simulated under the same condition. These findings are in agreement with previously established experimental studies reported elsewhere[1], where *pflA, pflB, pflC* and *pflD* were knocked out under micro-aerobic conditions to achieve a desired D-lactate production in *E. coli.*

It was previously reported that the conversion of one molecule of pyruvate to lactate requires one molecule of NADH and ethanol formation on the other hand needs two molecules of NADH. Alteration in the production of each metabolite, *E. coli* can modulates its metabolism fermentatively to grow on a number of substrates[1]. Alternative pathways could be activated to direct the carbon fluxes through the role of some key metabolites such as energy and electron donors as well as the metabolites

Table 1. *E. coli* strain design properties on glucose under the OptFlux software platform

E. coli Knock out genes/ strains	Growth rate (h^{-1})	% Biomass	D-lactate (mmolgDW^{-1}h^{-1})	Acetate (mmolgDW^{-1}h^{-1})	% Acetate	Ethanol (mmolgDW^{-1}h^{-1})	Formate (mmolgDW^{-1}h^{-1})
WT (orth Model	0.77075859	100	0.00	19.35334	100	9.15405	31.4814
pflA/b0902	0.746628	96.86	0.001	18.7887	97.1	8.87983	30.5699
pflB/b0903	0.746628	96.86	0.001	18.813	97.2	8.93473	30.6245
pflC/b3952	0.746628	96.86	0.001	18.813	97.2	8.93473	30.6245
pflD/b3951	0.746628	96.86	0.001	18.813	97.2	8.93473	30.6245

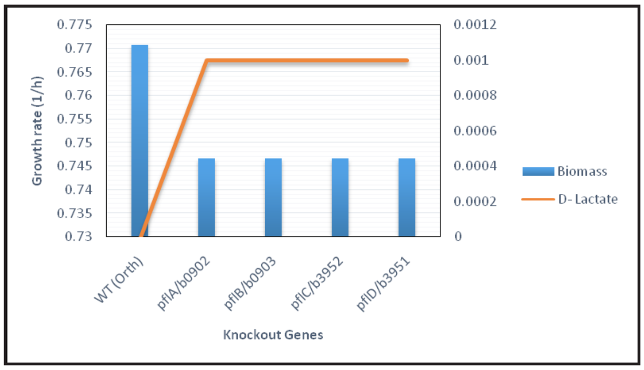

Figure 1. The Growth rates and D-lactate productivity of the wild-type and mutant E. coli models constructed in this study.

at the important branch points such as pyruvate and AcCoA[27] (see Figure 2).

In addition, it was previously established and known that mutation in the specific metabolic pathways such as those reported in this study, significantly affect the overall fermentation characteristics. Pyruvate is assimilated through PFL under micro-aerobic conditions[28]. Therefore knocking out the genes that codes for PFL, lead to a significant metabolic interventions that results in increase or decrease production of a particular compound such as D-lactate for instance. It is well known that that in *E. coli*, PFL has several genes classes with corresponding subunits, primarily designated as *pflA*, *pflB*, *pflC* and *pflD* as mentioned earlier it was also previously reported that PFL is a homodimeric protein with two subunits, PFL activating enzyme I and formate acetyl transferase I, are encoded by *pflA* and *pflB* genes respectively[1]. These two genes were believed to have constituted the *E. coli* PFL operon together with the probable formate transporter gene *foc*A and anaerobically regulated promoters[29]. While the other two genes (*pflC* and *pflD*) probably code for PFL activating enzyme II and formate acetyltransferase II respectively[1]. On the bases of these findings *pflC* and *pflD* are not involved in the *E. coli* PFL operon as reported elsewhere[1].

Maximum uptake rates for glucose were set to be 20 mmolgDW^{-1}h^{-1} and the corresponding Oxygen uptake rate was 5 mmolgDW^{-1}h^{-1} for micro-aerobic simulation.

Furthermore, the production of acetate and formate were reduced to nearly 97% of the wild-type model in all the mutant models examined (see Table 1). This might be attributed to the deletion of pfl genes that are directly involved in acetate and formic acid formation (see Figure 2).

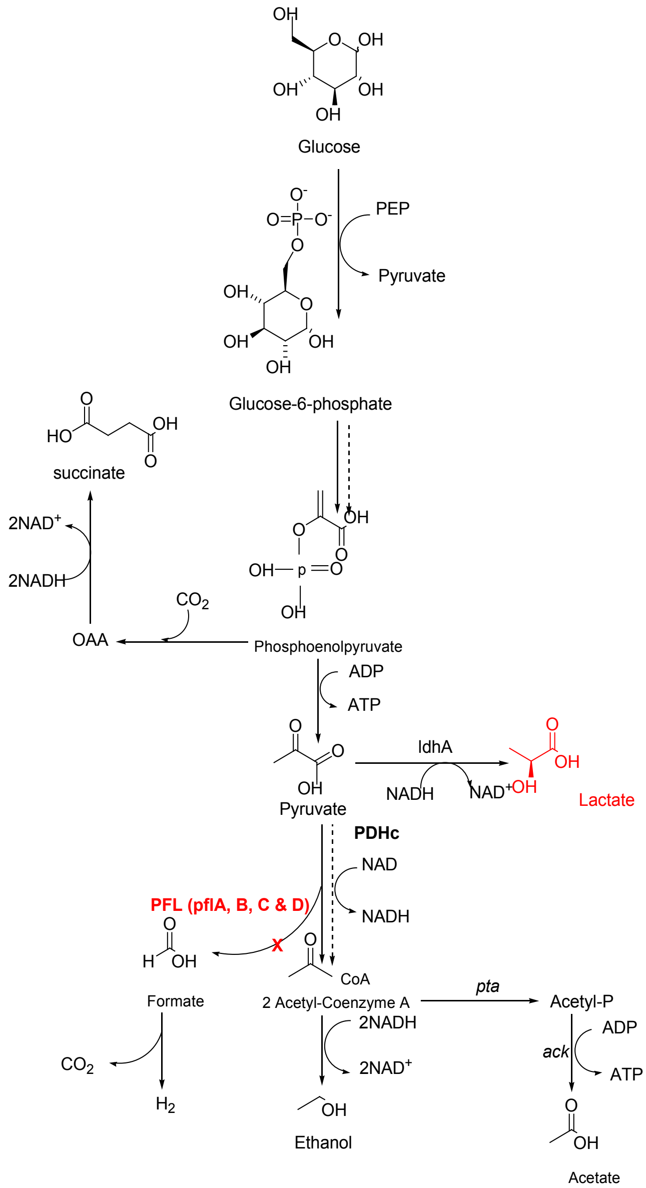

Figure 2. Pathways involved in micro-aerobic utilization of glucose by E. coli (WT Orth Model) to produce optically D-lactate and its constructed mutant strains (partially adopted from ref:[1, 6, 9, 10]. The pathways along with the deleted competing gene(s) are shown. The red colour enzymes represents the pathways that were inactivated via gene knockout. The sign "X" indicates gene knockouts. The knockout genes encode for Pyruvate formate lyase (pflA, pflB, pflC and pflD).

Although acetate, formate and ethanol were also produced, indicating that other alternative pathways such as the use of Pyruvate dehydrogenase complex (PDHc) may have been activated (see Table 1 and Figure 2). This is because it was previously reported that *E. coli* PDHc usually become activated when there is limited oxygen concentration or aerobic condition is established[28] (see Figure 2).

4. Conclusion

In conclusion, this study informed other studies by demonstrating that the genome scale metabolic reconstruction of *E. coli* iJ01366[8] in combination with the OptFlux software platform[12] using ROOM[16] can prospectively and effectively predict metabolic engineering targets by *in silico* deletion of *pflA/b0902*, *pflB/b0903*, *pflC/b3952* and *pflD/b3951* for increased optically pure D-lactate production in *E. coli*. This would guide future experimental studies by accurately predicting the target for engineering using the OptFlux software platform. In addition, the behavior of the *E. coli* perturbed cells could be well understood, creating a path to model-guided novel biological insight for the production of value added chemicals via microbial cell factories.

5. Acknowledgement

We would like to thank Pablo Carbonell, Sara Careia and Paulo Maia, for providing insight on the use of the OptFlux software platform.

6. References

1. Zhu J, Shimizu K. The effect of pfl gene knockout on the metabolism for optically pure D-lactate production by Escherichia coli. Appl Microbiol Biotechnol. 2004; 64(3):367–75.
2. Okano K, Tanaka T, Ogino C, Fukuda H, Kondo A. Biotechnological production of enantiomeric pure lactic acid from renewable resources: recent achievements, perspectives, and limits. Appl Microbiol Biotechnol. 2010; 85(3):413–23.
3. Mazumdar S, Clomburg JM, Gonzalez R. Escherichia coli strains engineered for homofermentative production of D-lactic acid from glycerol. Appl Environ Microbiol. 2010; 76(13):4327–36.
4. ZhouS, Shanmugam KT, Ingram LO. Functional Replacement of the Escherichia coliD-(-)-Lactate Dehydrogenase Gene (ldhA) with the L-(+)-Lactate Dehydrogenase Gene (ldhL) from Pediococcus acidilactici. Applied and Environmental Microbiology. 2003; 69(4):2237–44.
5. Zhou S, Causey TB, Hasona A, Shanmugam KT, Ingram LO: Production of Optically Pure D-Lactic Acid in Mineral Salts Medium by Metabolically Engineered Escherichia coli W3110. Applied and Environmental Microbiology. 2003; 69(1):399–407.
6. Zhu Y, Eiteman MA, DeWitt K, Altman E. Homolactate fermentation by metabolically engineered Escherichia coli strains. Appl Environ Microbiol. 2007; 73(2):456–64.
7. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BO. A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Molecular Systems Biology. 2007; 3:121.
8. Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, Palsson BO. A comprehensive genome-scale reconstruction of Escherichia coli metabolism-2011. Molecular Systems Biology. 2011; 7:535.
9. Mienda BS, Shamsir MS, Salleh FM. In silico metabolic engineering prediction of Escherichia coli genome model for production of D-lactic acid from glycerol using the OptFlux software platform. International Journal of Computational Bioinformatics and In Silico Modeling. 2014; 3(4):460–5.
10. Mienda BS, Shamsir MS. In silico Gene knockout metabolic interventions in Escherichia coli for Enhanced Ethanol production on Glycerol. Res J Pharm Biol Chem Sci. 2014; 5(4):964–74.
11. Mienda BS, Shamsir MS, Shehu I, Deba AA, Galadima IA. In Silico Metabolic Engineering Interventions of Escherichia Coli for Enhanced Ethanol Production, Based on Gene Knockout Simulation. IIOABJ. 2014; 5(2):16–23.
12. Rocha I, Maia P, Evangelista P, Vilaca P, Soares S, Pinto JP, Nielsen J, Patil KR, Ferreira EC, Rocha M: OptFlux: an open-source software platform for in silico metabolic engineering. BMC Systems Biology. 2010; 4:45.
13. Fong SS, Burgard AP, Herring CD, Knight EM, Blattner FR, Maranas CD, Palsson BO. In silico design and adaptive evolution of Escherichia coli for production of lactic acid. Biotechnology and Bioengineering. 2005; 91(5):643–8.
14. Rocha I, Maia P, Rocha M, Ferreira EC. OptGene - a framework for in silico metabolic engineering. Book of Abstracts of the 10th International Chemical and Biological Engineering Conference - CHEMPOR. 2008.
15. Vilaca P, Rocha I, Rocha M. A computational tool for the simulation and optimization of microbial strains accounting integrated metabolic/regulatory information. Bio Systems. 2011; 103(3):435–41.
16. Shlomi T, Berkman O, Ruppin E. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. Proceedings of the National Academy of Sciences of the United States of America 2005. 102(21):7695–700.
17. Monk J, Palsson BO. Genetics. Predicting microbial growth. *Science* 2014, 344(6191):1448-1449.

18. Mienda BS, Shamsir MS. Thermotolerant micro-organisms in Consolidated Bioprocessing for ethanol production: A review. Res Biotechnol. 2013; 4(4):1–6.
19. Lee SJ, Lee DY, 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. Appl Environ Microbiol 2005. 71(12):7880–7.
20. Lee DY, Yun H, Park S, Lee SY. MetaFluxNet: the management of metabolic reaction information and quantitative metabolic flux analysis. Bioinformatics. 2003; 19(16):2144–6.
21. Feist AM, Zielinski DC, Orth JD, Schellenberger J, Herrgard MJ, Palsson BO. Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli. Metabolic Engineering. 2010; 12(3):173–186.
22. McCloskey D, Palsson BO, Feist AM. Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli. Molecular Systems Biology. 2013; 9:661.
23. 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.
24. Edwards JS, Ibarra RU, Palsson BO. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. Nature Biotechnology. 2001; 19(2):125–30.
25. Fischer E, Zamboni N, Sauer U: High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived 13C constraints. Analytical Biochemistry. 2004; 325(2):308–16.
26. Le Novere N, Bornstein B, Broicher A, Courtot M, Donizelli M, Dharuri H, Li L, Sauro H, Schilstra M, Shapiro B *et al.* BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic Acids Research. 2006; 34(Database issue):D689–91.
27. Zu Berstenhorst SM, Hohmann HP, Stahmann KP. Vitamins and Vitamin-like Compounds: Microbial Production. Encyclopedia of Microbiology. 3rd ed. Oxford: Academic Press; 2009. p. 549–61.
28. Alexeeva SBDK, Sawers G, Klaas J. Hellingwerf, Amjtd M. Effects of Limited Aeration and of the ArcAB System on intermediary metabolism in E. coli. Journal of Bacteriology. 2000; 182(17):4935–40.
29. Ndag S. Promoter 7 of the Escherichia coli pfl Operon Is a Major Determinant in the anaerobic Regulation of Expression by ArcA. J Bacteriol. 1995; 177(18):5338–41.