# Genome-scale metabolic design for improved target production of *Escherichia coli* based on flux balance analysis

Submitted to
Graduate School of Information Science and Technology, Osaka University

January 2018

Kento Tokuyama

# List of publications

## 1. Journal paper

[1-1] Kento Tokuyama, Satoshi Ohno, Katsunori Yoshikawa, Takashi Hirasawa, Shotaro Tanaka, Chikara Furusawa and Hiroshi Shimizu. “Increased 3-hydroxypropionic acid production from glycerol, by modification of central metabolism in *Escherichia coli.” Microb Cell Fact,* 13:64, 2014. (Chapter 3)

## 2. International conferences

[2-1] Kento Tokuyama, Yoshihiro Toya, Takaaki Horinouchi, Chikara Furusawa, Hiroshi Shimizu, " Expanding the space of actual metabolic network for growth-coupled succinate production in *Escherichia coli* by adaptive laboratory evolution”, *The 23th Symposium of Young Asian Biological Engineers’ Community*, Poster number: PC-15, Xian, China, October 2017, Poster (Chapter 2 & 4)

[2-2] Kento Tokuyama, Yoshihiro Toya, Takaaki Horinouchi, Chikara Furusawa, Hiroshi Shimizu, "Adaptive laboratory evolution of *in silico* designed succinate producing *Escherichia coli*”, *International Union of Microbiological Societies Congress 2017*, Poster number: PO 047, Singapore, July 2017, Poster (Chapter 4)

[2-3] Kento Tokuyama, Satoshi Ohno, Katsunori Yoshikawa, Takashi Hirasawa, Shotaro Tanaka, Chikara Furusawa, Hiroshi Shimizu. “Increased 3-hydroxypropionic acid production from glycerol based on the genome-scale metabolic simulation in *Escherichia coli*.”, *Metabolic Engineering X,* Poster number: 201, Vancouver, Canada, June 2014, Poster (Chapter 3)

## Abstract

Strain improvement plays a central role in development of microbial process for production of valuable chemical. Recent development of computational approach enable to predict metabolic state, and strain improvements have been performed based on the prediction. Applying flux balance analysis (FBA) on genome-scale metabolic model, which contains all known metabolic reactions and genes, metabolic network for enhancing growth-coupled target production can be designed with considering mass-balance. Although target productivities of various microbial processes has been improved by strain improvement based on the metabolic design, the target productivities were always lower than of the expected values. One possible reason of the decreased productivity is that FBA only considers mass-balance equation on metabolic network, while some of reactions in actual cells exists as rate-limiting reactions by metabolic constraints by substrate level, enzyme level and enzyme kinetics, and as a result, the engineered strain could not achieve the predicted state by FBA. This thesis focused on understanding the mechanism of rate-limiting reaction based on the metabolic design by FBA, and strain improvement by overcoming the rate-limiting reactions in productions of 3-hydroxypropionic acid and succinic acid, which are precursors of bio-plastics.

This doctor thesis consists of five chapters:

Chapter 1: Introduction and general objective of this study are described.

Chapter 2: In order to evaluate the effect of limitation of flux of metabolic reaction on the target production, rapid screening method of possible rate-limiting reaction was developed by using genome-scale metabolic model. The method assumes the metabolic constraints as an upper-bound constraint on each enzymatic flux on FBA calculation and simulate which reaction is thought to be possible rate-limiting reactions for the target production. The experimental results of strain improvement for 1,4-butanediol production reported by Yim et al. were consistency with the predicted rate-limiting reactions. This proposed method also screened the possible rate-limiting reactions for various microbial process of growth-coupled target production.

Chapter 3: Strain improvement for 3-hydroxypropionic acid production by the knockout mutant of *E. coli* was achieved by additional genetic manipulation for overcoming the predicted rate-limiting reactions. Experimental evaluation of the

knockout mutant suggested that overflow of 1,3-propanediol (1,3PDO) from 3-hydroxypropionaldehyde led to decrease the flux into 3-hydroxypropionaldehyde dehydrogenase, which was predicted as one of the rate-limiting reactions. Additional deletion of *yqhD* encoding NADPH-dependent aldehyde reductase successfully improved 3HP production by rewiring carbon flux from 1,3PDO overflow to 3HP synthesis.

Chapter 4: Molecular mechanism causing the rate-limiting reaction on the growth-coupled succinate production was revealed through strain improvement by adaptive laboratory evolution (ALE). The proposed method developed in the chapter 2 displayed that the knockout mutant of *adhE-pykAF-gldA-pflB* had 9 possible rate-limiting reactions. High succinate producing strains were isolated from ALE experiment of the knockout mutant by selecting faster growing strain, since the strain has to produce succinate in order to grow faster. Mutation analysis of high succinic acid producing strains obtained from ALE revealed that phosphoenolpyruvate carboxylase (Ppc), which was predicted as one of possible rate-limiting reaction, was a dominant rate-limiting reaction. Reverse engineering and functional analysis of the identified mutations revealed that the mutants of Ppc desensitized allosteric inhibition by aspartate, and improved enzymatic activity of Ppc toward increasing succinate production.

Chapter 5: General conclusion and future perspective of this research are described.

Table of contents

# Chapter 1: Introduction

## 1-1 Highlights

- Microbial production is a promising way to compensate the increasing demand for commercial use due to depletion of oil resource.
- Strain improvement for enhancing the target productivity is essential for development of the microbial process.
- Flux balance analysis (FBA) is a powerful method to design the genome-scale metabolic network for enhancing the growth-coupled target production with considering mass-balance.
- Although target productivities of various fermentation processes has been improved based on the metabolic design, the target productivities were always lower than of the expected value by FBA.
- The general objective of this study was to unveil what causes the miss-prediction by FBA and achieve further strain improvements.

## 1-2 Microbial production of chemicals

Microbial fermentation for chemical production is a promising way to compensate the increasing demand of commercially used chemicals. Strain improvement for increasing target productivity is an essential for development of the microbial process. Recombinant DNA techniques have been applied to improve target production by modification of metabolic pathway such as gene deletion. It is important for further strain

development to build a metabolic engineering strategy for optimizing whole metabolic network toward increasing target production, because metabolism in actual cells constructs huge and complex network structure via intermediate metabolites and cofactors.

### 1-2-1 Importance of biotechnological production process

Biosynthesis of target chemicals using microorganism's metabolism has been attracted for compensating the increasing demand for commercially valuable chemicals due to depletion of oil resource. Metabolism is a general term meaning a set of chemical reactions essential for life-sustaining within the cells of organisms. The microbial process of chemical production converts abundantly available substrates such as renewable biomass resources and waste glycerol to the valuable chemicals such as bio-plastic, bio-fuel and pharmaceutical. Therefore, the process is expected to contribute to the sustainable development of highly civilized human society.

Actually, the microbial process has contributed to develop our life in various industrial fields. The oldest record about the microbial process is the conversion of sugar to alcohol by yeast in B.C (1). In the early 1900's, Alexander Fleming firstly discovered antibiotic effect of penicillin and this led to begin the golden era of microbial production of antibiotics. The microbial production of monosodium glutamate, which is known as an umami tastant, has been developed by using *Corynebacterium glutamicum* before World War II in Japanese food companies of Ajinomoto Corporation Inc. and Kyowa Hakko Kogyo Corporation Ltd. (2). Recently, the microbial

production processes have been commercially developed for 1,3-propanediol production in DuPont Tate and Lyle Bioproducts Company (USA), succinate in Myriant Corporation (USA) and 1,4-butanediol in Genomatica Inc. (USA). In the 2000's, the US Department of Energy addressed the microbial process of valuable chemical production and listed the potential target chemicals produced from biomass (3)(4).

### 1-2-2 History of strain improvement

Industrial strain improvement plays a central role in the development of the microbial process. Before the 1980s, the strain improvement has been achieved by random mutagenesis/selection/screening (5). Mutagenesis is conducted by using physical mutagen like ultra violet or chemical mutagen like nitrosomethyl guanidine. Physiological characteristics such as target productivity, growth ability and stability of surviving clones, which can grow in designated selective pressure, are evaluated in the step of screening. To repeat the cycle of random mutagenesis/selection/screening is a beneficial way to improve target productivity. Since the biggest advantage of this approach is capable to apply for any organism independent a priori knowledges about them, strain improvements have been achieved for various target products including alcohol (6), anti-biotics (7), amino acids (8) and nucleic acids (9). However, there are also several disadvantages in the approach. Firstly, sufficient time is needed to develop the ideal strain. Second, identification of beneficial mutations for improved target productivity is difficult because they have a large number of mutations, some

of which does not affected on the productivity. This complexity for understanding the mechanism of improved target production hampers to apply them for the other target chemical productions.

After 1970, development of site-directed mutagenesis approach, which is known as recombinant DNA technology, enable us to introduce artificial mutations into designated site *in vitro* and express them as plasmid DNAs *in vivo* (10). Sanger and Coulson developed that a novel method for determining sequence in DNA using dideoxynucleotide in 1977 (11). These technologies contributed to identify the beneficial mutations on strain improvement. In the 1980s, development of chromosomal transformation technology made it possible to construct knockout mutant of designated chromosomal genes replaced by antibiotic marker (12). In the 1990s, these techniques have accelerated metabolic pathway modification for improving the properties of microorganisms and their target productivities. The term *metabolic engineering* is defined as the directed improvement of product formation or cellular properties through the modification of specific biochemical reactions or introduction of new ones with the use of recombinant DNA technology (13).

### 1-2-3 Metabolic engineering

Metabolic engineering is a powerful approach to improve target productivity by recombinant DNA technology. Optimizing metabolic flux distributions for the target production is important in order to apply metabolic engineering for strain improvement. The *flux* is the term meaning

reaction rate (or velocity) in this thesis, and a fundamental determinant of cell physiology and the most critical parameter of a metabolic pathway (14). In general metabolic engineering strategy, biosynthesis pathways of target metabolite are activated, while the other pathways which do not related on target production are inactivated (15)(16)(17)(18)(19)(20)(21)(22)(23)(24).

Gene deletion (or gene knockout) is mostly used technique to block unnecessary pathways and rewire metabolic flux to target synthesis. There are several ways to disrupt target gene in chromosomal DNA. The mostly used approach is replacement of the target gene by antibiotic resistance gene such as *kan* encoding kanamycin resistance gene derived from *Streptomyces kanamyceticus* (25)(26). Recent development of genome editing such as CRISPR has improved efficiency of gene deletion without antibiotic resistance gene in various organisms including *E. coli* (27), yeast (28), and cyanobacteria (29). Target gene is often determined based on the culture result of a parental strain and priori knowledges of metabolism. For example, when the parental strain produces a large amount of byproduct, the gene involved in the synthetic pathway of the byproduct is a candidate to be disrupted. Since cofactors also play an important role in the biosynthesis pathway, gene deletion strategy have to consider not only carbon flow, but also supply and regeneration of cofactors for strain improvement (30). However, it is difficult to make metabolic engineering strategy by only relying on our experiences and intuitions, because there are a large number of reactions constructing complex metabolic network via intermediate metabolites and cofactors in the cells. Therefore, rational engineering of

metabolism with considering whole metabolic network or wide view of metabolic pathways is essential for the practice of metabolic engineering for strain improvement.

### 1-3 Genome-scale metabolic design

Genome-scale metabolic reconstruction has been developed based on whole genome sequence information and gene-protein-reaction associations. Flux balance analysis (FBA) is a widely used technique to calculate a flux distribution using the genome-scale metabolic model based on assumptions of metabolic steady state and maximizing cell growth. Gene knockout targets for growth-coupled target production can be screened by using FBA. The term *genome-scale metabolic design* is meaning rational design of whole metabolic network for enhancing target production in this thesis. Although various fermentation processes has been improved by gene deletion based on the prediction, the target productivity is always lower than of the predicted value. One possible reason causing the inconsistency was that FBA only considers mass-balance equation on the metabolic network, while metabolism in actual cells is complexity regulated by low enzymatic activity (31), enzymatic regulation (32), expression regulation (33), and thermodynamic feasibility (34). If such regulation system cause rate-limiting reactions on *de novo* biosynthesis of target metabolite, the cells could not achieve the predicted metabolic state. It is essential for genome-scale metabolic design to identify rate-limiting reactions and build a metabolic engineering strategy.

### 1-3-1 Reconstruction of genome-scale metabolic model

To simulate the physiological characteristics of microorganism and the behavior of metabolic network, constraint-based metabolic modeling has been developed. Metabolism is sequential chemical reactions catalyzed by enzymes, and the products produced by a metabolic reaction are rapidly catalyzed by the other metabolic reaction. In the modeling approach, a pseudo-steady state assumption is generally applicable for them, because concentrations of internal metabolite do not dynamically changed during stable culture conditions, such as exponential cell growth phase. Therefore, the sum of fluxes in and out of each internal metabolite is balanced. In first studies, the constrain-based modeling was applied to determine the internal flux through local metabolic network such as tricarboxylic acid cycle (35)(36). Recent development of whole genome-sequencing techniques and accumulation of knowledges of metabolism enable us to extend reconstruction of the metabolic network from local reactions to genome-scale metabolic network. Genome-scale metabolic models are constructed from genome sequence annotation and information about biochemical reactions from literatures and databases such as KEGG (37)(URL: http://www.genome.jp/kegg/) and BRENDA (38)(URL: https://www.brenda-enzymes.org/). The genome-scale metabolic models usually contain gene-protein-reaction associations, stoichiometry of the biochemical reactions including transport reaction and chemical reaction, thermodynamic information (i.e. reversibility) and biomass reaction constructed from a molecular weight of components of cell of 1 g $mmol^{-1}$.

The first genome-scale metabolic model was built for *Haemophilus influenza* in 1999 (39). Subsequently, the genome-scale metabolic models were developed on well-characterized model organisms such as *Escherichia coli* (40) and *Saccharomyces cerevisiae* (41). During past ten years, the number of the genome-scale metabolic reconstructions has rapidly grown and these have been widely applied in studies related in metabolic engineering, biological discovery, phenotypic behavior, network analysis and evolutionary biology (42). The genome-scale metabolic models are now available for more than 100 species including bacteria, eukaryotes and archaea (43)

### 1-3-2 Solution space of metabolic network

Steady state assumption of metabolic network constraints space of feasible fluxes of each reaction in genome-scale metabolic model. The term *solution space* is used in this thesis in reference to the feasible space of fluxes. The number of dimensions of solution space is equal to the number of reaction included in the genome-scale metabolic model. Here, an example of metabolic solution space is described using simple metabolic network model.

Figure 1-1 shows a simple example of metabolic network, which contains 6 intracellular reactions, 3 transport reactions and biomass reaction. There are 5 intracellular metabolites (A-E), 4 extracellular metabolites ($A_{ext}$, $D_{ext}$, $E_{ext}$ and Biomass). The stoichiometric equation of the network are also shown in Figure 1-1.

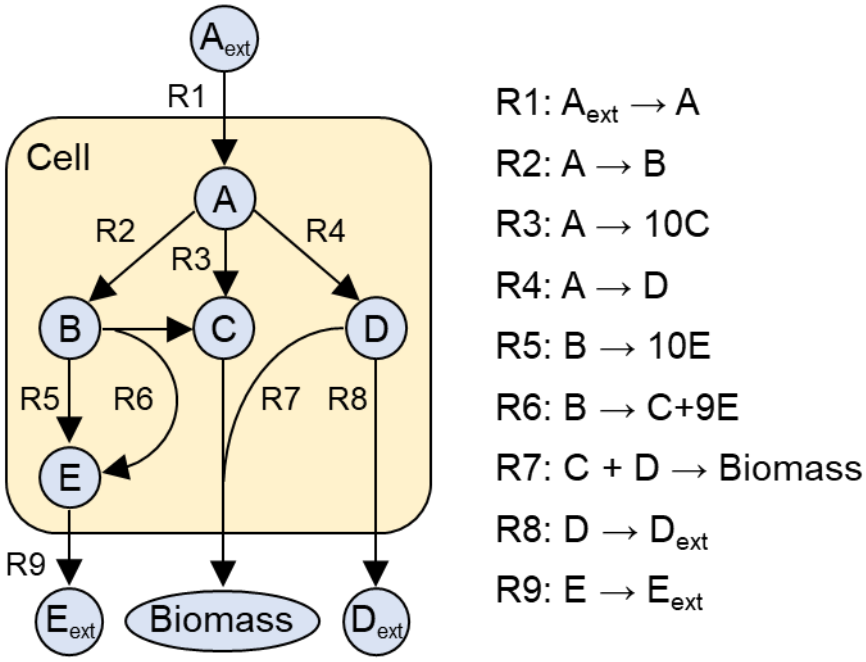



**Figure 1-1 A simple example of metabolic network.** Arrows indicate metabolic reaction. A ~ E indicate intracellular metabolites. $A_{ext}$, $D_{ext}$ and $E_{ext}$ indicate extracellular metabolites. R1~R9 indicate each metabolic reactions, whose stoichiometric equations are described in next to the metabolic network.

In the metabolic network shown in the Figure 1-1, a dynamic mass balance is described as following equations (1.1a-e). $v_j$ represents metabolic fluxes of the $j^{th}$ reaction (i.e. R1-R10) and the unit of $v_j$ is mmol gCDW$^{-1}$ hr$^{-1}$.

$$\frac{dA}{dt} = v_1 - v_2 - v_3 - v_4 \quad \text{(Equation 1.1a)}$$

$$\frac{dB}{dt} = v_2 - v_5 - v_6 \quad \text{(Equation 1.1b)}$$

$$\frac{dC}{dt} = 10v_3 + v_6 - v_7 \quad \text{(Equation 1.1c)}$$

$$\frac{dD}{dt} = v_4 - v_7 - v_8 \quad \text{(Equation 1.1d)}$$

$$\frac{dE}{dt} = 10v_5 + 9v_6 - v_9 \quad \text{(Equation 1.1e)}$$

Assuming the metabolic steady state on the metabolic network, constraint equations on each metabolic fluxes are obtained as follows (equation 1.2a-e), because intracellular metabolite's concentration are not dynamically changed.

$$v_1 - v_2 - v_3 - v_4 = 0 \quad \text{(Equation 1.2a)}$$

$$v_2 - v_5 - v_6 = 0 \quad \text{(Equation 1.2b)}$$

$$10v_3 + v_6 - v_7 = 0 \quad \text{(Equation 1.2c)}$$

$$v_4 - v_7 - v_8 = 0 \quad \text{(Equation 1.2d)}$$

$$10v_5 + 9v_6 - v_9 = 0 \quad \text{(Equation 1.2e)}$$

These constraint equations can be represented using a matrix notation,

$$\begin{bmatrix} \mathbf{1} & \mathbf{-1} & \mathbf{-1} & \mathbf{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{10} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{10} & \mathbf{9} & \mathbf{0} & \mathbf{0} & \mathbf{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{v_1} \\ \boldsymbol{v_2} \\ \boldsymbol{v_3} \\ \boldsymbol{v_4} \\ \boldsymbol{v_5} \\ \boldsymbol{v_6} \\ \boldsymbol{v_7} \\ \boldsymbol{v_8} \\ \boldsymbol{v_9} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad \text{(Equation 1.3)}$$

$$\Leftrightarrow \sum_{j \in \boldsymbol{R}} \boldsymbol{S}_{i,j} \cdot v_j = 0 \quad (\forall_i \in \boldsymbol{M} \text{ and } \forall_j \in \boldsymbol{R}) \qquad \text{(Equation 1.4)}$$

where $\boldsymbol{S}$ is called stoichiometric matrix indicating stoichiometric coefficients of all metabolic reactions. $\boldsymbol{M}$ is a set of metabolites in the system or cell. $\boldsymbol{R}$ is a set of reactions in the system or cell.

In general, stoichiometric metabolic reconstruction does not have any constraint for fluxes of intracellular reactions, and only consider uptake flux in order to calculate the solution space satisfying the assumption of metabolic steady state. Furthermore, the values of fluxes are assumed continuous explanatory variables for calculating feasible solution space. Here, the following constraints are assumed to describe the solution space of the metabolic network in Figure 1-1.

$$v_1 = 1 \text{ mmol gCDW}^{-1} \text{ hr}^{-1} \qquad \text{(Equation 1.5)}$$

$$0 \leq v_j \leq \infty \quad (j = 2 \sim 9) \qquad \text{(Equation 1.6)}$$

Since the number of dimensions of solution space is equal to the number

of reaction included in the metabolic network, it is difficult to visualize the overall solution space. In this thesis, solution space is described by focusing molar yield of target metabolite and biomass reaction per consumed substrate. Here, the solution space of the metabolic network is described with regarding metabolite $E_{ext}$ as target metabolite (Figure 1-2).

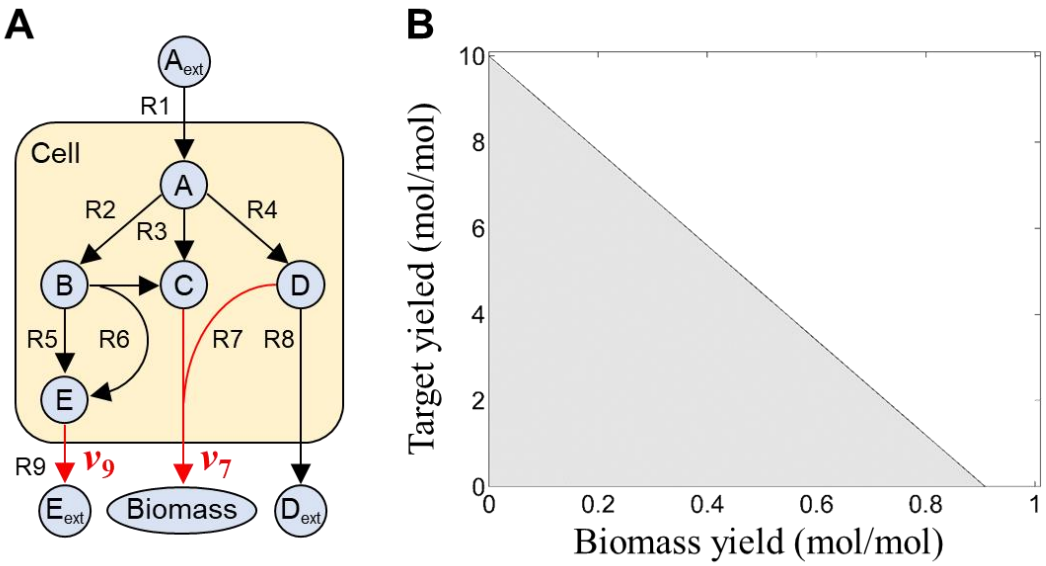



**Figure 1-2 A simple example of metabolic network (A) and solution space (B).** Red arrows in pane A indicate metabolic reactions describing their fluxes in the x-axis and y-axis in pane B.

The y-axis and x-axis in Figure 1-2B are equivalent of the molar yield of the target metabolite $E_{ext}$ and Biomass, respectively.

The structure of solution space is depend on the structure of metabolic network. The solution space means the relationship of target production and cell growth with considering mass-balance equation. In wildtype

microorganisms, the relationship is usually trade-off except for in case of natural byproducts such as acetate in *E. coli*.

The shape of solution space is depend on network structure. For example, considering disruption of reaction 3 (R3), the additional constraint is imposed as following equation 1.7),

$$v_3 = 0 \qquad \text{(Equation 1.7)}$$

and the solution space is changed as shown in Figure 1-3. In this case, the cells have to produce target metabolite of $E_{ext}$ in order to grow, since intracellular metabolite C required for biomass reaction is only supplied by reaction 6, which is coupled with production of target metabolite E. This type of the solution space is thought to be favorable for growth-coupled production process. The term *growth-coupled target production* is used in this thesis in reference to the target production process coupled with cell growth. Since the shape of solution space is depend on the network structure, rational metabolic design by disruption and addition of metabolic reactions is important for the development of the growth-coupled target production.

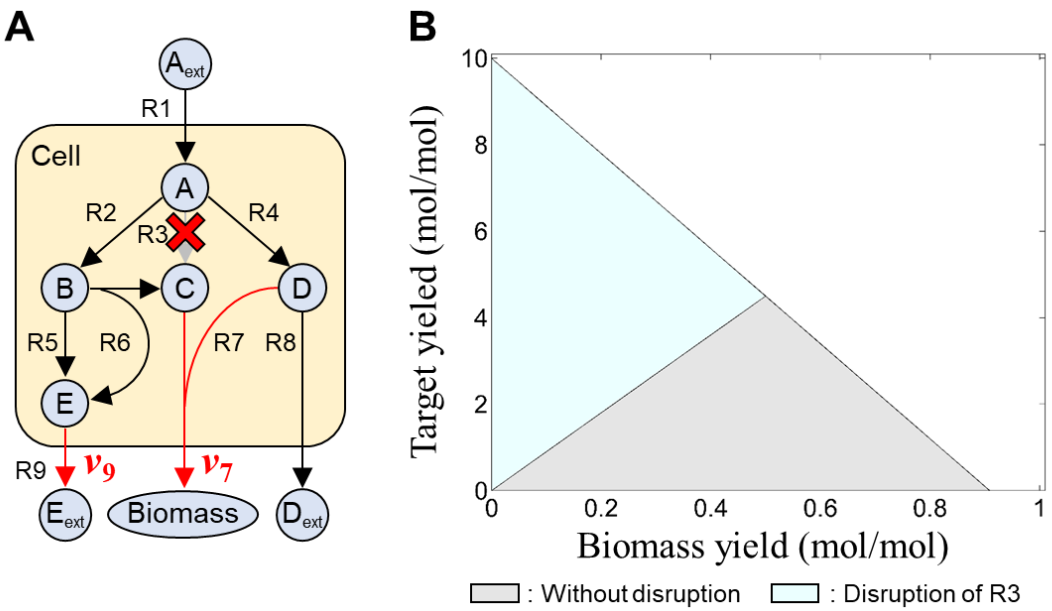



**Figure 1-3 A simple example of metabolic network with disruption of reaction 3 (A) and its solution space (B).** (B) Gray and blue area indicate the feasible solution spaces of the metabolic network without disruption of R3 and with the disruption, respectively.

### 1-3-3 Flux balance analysis

Flux balance analysis (FBA) is a widely used approach for studying biochemical networks, in particular the genome-scale metabolic network reconstructions (44)(45)(46)(47). FBA uses the steady state assumption and an objective function to model a desired phenotype as an optimization problem (Figure 1-4). Since steady state assumption of metabolic network reduces the system to a set of linear equations as shown in the chapter 1-3-2, a flux distribution is calculated using linear programming as following equations:

(Equation 1.8)

$$\text{Maximize} \quad \boldsymbol{c}^T \cdot \boldsymbol{v}$$

$$\text{Subject to} \quad \sum_{j \in \boldsymbol{R}} S_{i,j} \cdot v_j = 0 \quad (\forall_i \in \boldsymbol{M} \text{ and } \forall_j \in \boldsymbol{R})$$

$$v_{j,min} \leq v_j \leq v_{j,max}$$

Where

$\boldsymbol{R}$ is a set of reactions in the system or cell.

$\boldsymbol{M}$ is a set of metabolites in the system or cell.

$\boldsymbol{c}$ is a vector that represents coefficients of an objective function to be maximized.

$\boldsymbol{v}$ is a vector of metabolic fluxes in the system or cell.

$v_j$ represents metabolic fluxes of the $j^{\text{th}}$ reaction.

$v_{j,min}$ represents lower bound of metabolic fluxes of the $j^{\text{th}}$ reaction.

$v_{j,max}$ represents upper bound of metabolic fluxes of the $j^{\text{th}}$ reaction.

$S_{i,j}$ represents the stoichiometric coefficient indicating the amount of the

$i^{th}$ metabolite produced per unit of flux of the $j^{th}$ reaction

For example, the flux distribution of the simple metabolic network at the optimal growth state is shown in Figure 1-5.

In genome-scale metabolic model, the biomass reaction is generated based on molecular weight of metabolic precursors of cell of 1g $mmol^{-1}$. These precursors include nucleotides, amino acids, energy metabolites and components of cell membrane that a cell requires for replication and maintenance. When the coefficient is set to 1 for biomass reaction, $v_{biomass\ reaction}$ indicates maximum specific cell growth rate under the designated conditions. In general, the maximum cell growth rate and flux distributions at the optimal growth state are highly consistency with the experimental results in various wild-type microorganisms (48)(49).

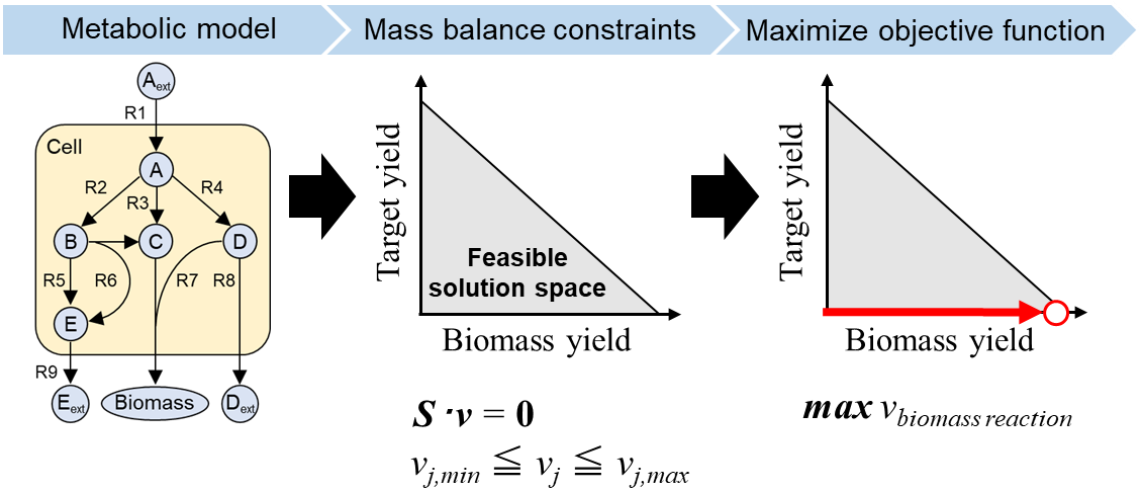



**Figure 1-4 Principal of flux balance analysis.**

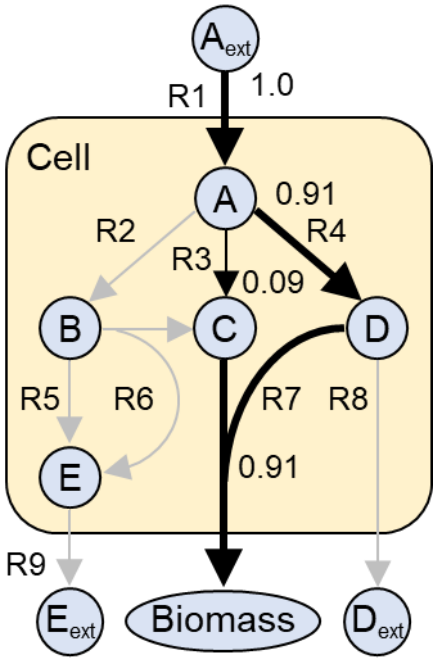



**Figure 1-5 Flux distribution of a simple metabolic network calculated by FBA.** Black arrows indicate activate metabolic reactions. Gray arrows indicate inactivate metabolic reactions (i.e. the fluxes is equal to zero.).

### 1-3-4 Importance of metabolic design for growth-coupled target production

Strain improvement for optimizing the production of valuable chemicals is an overarching challenge in biotechnology. However, wildtype microorganisms almost show low productivities for industrial applications. Since target production consume carbon and energy sources, the relationship of cell growth and target production is usually trade-off. This means that increasing target production leads to decrease cell growth rate. The trade-off relationship is expected to be unfavorable for the microbial process, because the final volume of target metabolite (i.e. production titer) is depend on not only production rate, but also cell volume and cell growth rate. Furthermore, during fermentation process of such strain, population of the cells having high growth rate and low productivity likely overgrow the overall population. This leads unstable production of the target metabolite and decrease the productivity with increasing the population of high-growing cells (50)(51).

Growth-coupled target production is a promising approach to develop the microbial process. Since the strain designed for the growth-coupled target production produces target metabolite in order to grow, the process is expected to contribute stable production and increase target productivity. The other advantage of the process is that high target producing strains can be easily selected by selecting faster growing cells through adaptive laboratory evolution (52)(53), which is described in the chapter 4. Since the relationship is depend on the metabolic network structure as shown in the chapter 1-3-2, genetic manipulation of metabolic genes can change the network structure and the relationship. Rational design of metabolic

network is an important for development of the growth-coupled target production.

### 1-3-5 Metabolic engineering strategy based on FBA

Effective gene knockout for the growth-coupled target production can be screened by evaluating the predicted target productivity at the optimal growth state. The knockout mutant must produce a target metabolite to produce the biomass components with satisfying the mass balance constraints. For example, the metabolic engineering strategy for the growth-coupled production of target metabolite $E_{ext}$ in the simple metabolic network is displayed in Figure 1-6. Disruption of reaction 3 leads to increase target production yield at optimal growth state (Figure 1-6B), because precursor metabolite C for biomass reaction is only supplied via target metabolite biosynthesis in the metabolic network disrupting reaction 3. On the other hand, disruption of reaction 8 do not change the production yield (Figure 1-6C). These results indicate that the disruption of reaction 3 blocks unnecessary pathways for target production and rewire metabolic flux to target synthesis coupled with cell growth.

Recent development of calculation algorithm enable us to evaluate the effect of combinatorial gene knockouts on the growth-coupled target production (54)(55)(56). Using such simulations, strain improvements have been achieved for various target production processes (Table 1-1). For example, Fong et al. constructed high lactate producing *E. coli* by double knockout of *pta* and *adhE* predicted by Optknock algorithm (54), which is a

most used the calculation algorithm, while wildtype *E. coli* do not secrete lactate (52).

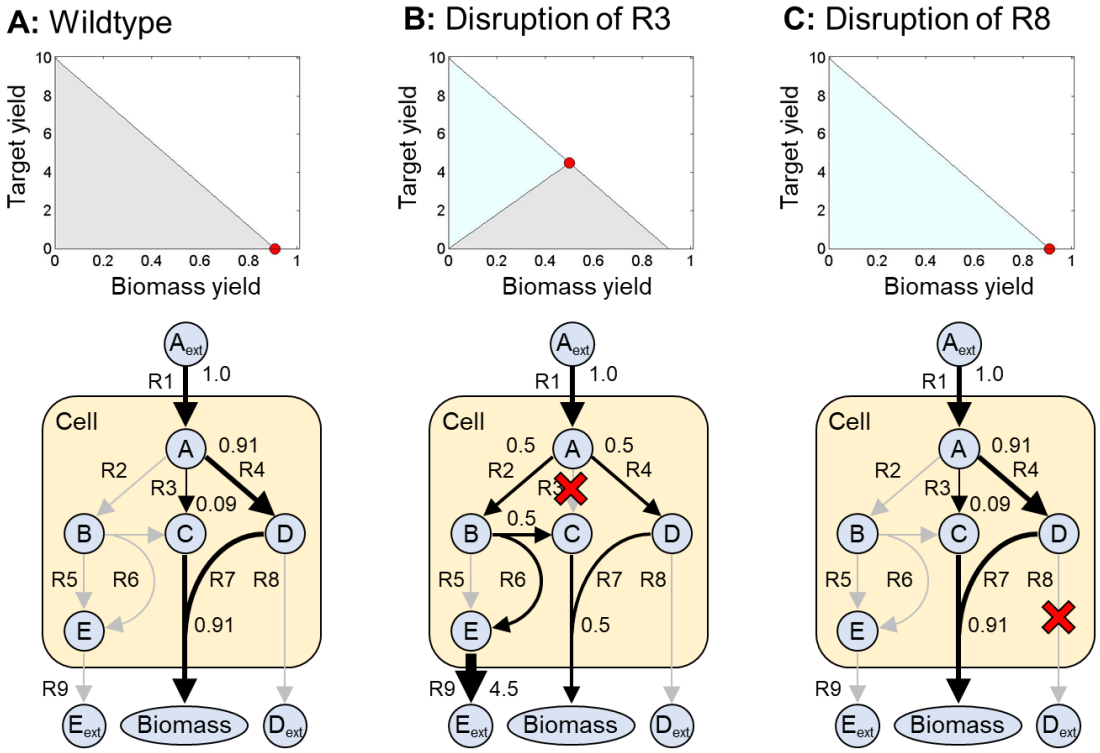



**Figure 1-6 Comparison of solution spaces and flux distributions between different metabolic networks.** Upper panes indicate the solution spaces, and lower panes indicate flux distributions calculated by FBA. Red points in the solution space indicate optimal solution calculated by FBA.

**Table 1-1 Experimental evaluation of gene knockout simulation for strain improvements for target production.**

| Target process | Host<br>Knockout gene | Carbon molar yield of target product (C-mol/C-mol)[*1] | | | Ref. |
|---|---|---|---|---|---|
| | | Pre-knockout[*2] | Knockout mutant | Predicted value | |
| Glucose -> Succinate | *E. coli*<br>*ptsG, pykAF* | 0 | 0.23 | 0.46 | (57) |
| Glucose -> Lactate | *E. coli*<br>*pta, adhE* | 0 | 0.61 | 0.85 | (52) |
| Glucose -> Lactate | *E. coli*<br>*pta. adhE* + **ALE**[*3] | 0 | 0.86 | 0.85 | (52) |
| Glucose -> 1,4-butanediol | *E. coli*<br>*adhE, ldh, pflB, mdh* | 0.02 | 0.10[*4] | 0.49 | (58) |
| Glucose -> 2,3-butanediol | *S. cerevisiae*<br>*ADH1, ADH3, ADH5* | 0 | 0.15 | 0.42 | (59) |
| Glycerol<br>-> 3-hydroxypropionic acid | *E. coli*<br>*tpiA, zwf* | 0.05 | 0.20 | 0.71 | This study |
| Glycerol<br>-> 3-hydroxypropionic acid | *E. coli*<br>*tpiA, zwf* + ***yqhD*** **knockout** | 0.05 | 0.34 | 0.71 | This study |
| Glycerol -> Succinate | *E. coli*<br>*adhE, pykAF, gldA, pflB* | 0.03 | 0.08 | 0.44 | This study |
| Glycerol -> Succinate | *E. coli*<br>*adhE, pykAF, gldA, pflB*<br>**+ ALE**[*3] | 0.03 | 0.45 | 0.44 | This study |

*1 Carbon molar yields were calculated from the carbon-mol of the product per cabon-mol of the consumed substrate.

*2 Pre-knockout indicates wild-type microorganisms or host strains expressing genes involved in synthetic pathway of target metabolite.

*3 "+ALE" indicates the evolved strain obtained from adaptive laboratory evolution of the knockout mutant

*4 The value was refer to the culture result of ECKh-401 strain (*E. coli ΔadhE ΔldhA ΔpflB Δmdh ΔarcA lpdA::lpdA*$^{D354K}$ expressing heterologous genes involved in 1,4BDO synthetic pathway) in the previous study (58).

### 1-3-6 Problem of the metabolic engineering strategy based on FBA

FBA is a useful to predict to build metabolic engineering strategy for development of growth-coupled target production. Although the engineered strains based on FBA have actually improved the target productivity, these knockout mutants always showed lower cell growth ability and target production yield as compared to FBA predictions as shown in Table 1-1. The possible reasons why the knockout mutant decreased the productivity of growth-coupled target production are summarized as following.

1) Environmental stress
   - 1-A) Growth inhibition by oxidative stress and culture environment such as pH, temperature and osmotic stress
   - 1-B) Regulation of specific metabolic reaction by chemicals (Ex. feedback inhibition of amino acid biosynthesis)

2) Limitation of metabolic reaction
   - 2-A) Shortage of enzyme level
   - 2-B) Shortage of substrate level
   - 2-C) Kinetic constraints

3) Difference of composition of biomass reaction

(1) Environmental stress is caused by physical factors such as pH, temperature and osmotic pressure, and chemicals in culture medium including substrate and product. Environmental stress can effect on both of overall physiological characteristic of microorganism and specific metabolic

reaction. For example, oxidative stress, which is the metabolic condition arisen from imbalance between toxic reactive oxygen species and antioxidant systems, leads redox imbalance and increases energy requirement for cell growth or damages macromolecules such as DNA (60). Chemicals existing in culture broth also damage macro molecules and inhibit specific metabolic reaction (61)(62)(63). In such case, construction of tolerable strain against the environmental stress is important for increasing target production (62)(64).

(2) Metabolic reaction in actual cells is regulated by kinetic constraints including enzyme level, substrate level and enzymatic activity, while FBA simplify the complicated metabolic system by considering only mass-balance equation of the stoichiometry. Product inhibition discussed as the environmental stress (1-B) is also considered to be caused by kinetic constraint such as feedback inhibition of amino acid biosynthesis. Enzymatic activity is limited by not only its kinetic parameters, but also interaction with other molecules such as inhibitor. Since there are several factors regulating metabolism such as limitation of enzymatic activity (31), enzymatic regulation (32), expression regulation (33), and thermodynamic feasibility (34), and causing rate-limiting reactions, the flux space of actual metabolic network must be diminished from the solution space in FBA. Therefore, these simplification of metabolic system potentially lead to overestimate the productivity by FBA and that the knockout mutant always display the lower productivity than of the predicted value. For development of genome-scale metabolic design for strain improvement, it is absolutely

essential to know which reaction is rate-limiting step, what cause the rate-limiting and how do we overcome the metabolic limitation.

For example, assuming the reaction 2 as a sole rate-limiting reaction in the simple metabolic network with disruption of the reaction 3, and the maximum flux through the reaction 2 is limited lower than 0.4 mmol $gCDW^{-1}$ $hr^{-1}$, an additional constraint is added to the linear programming problem as following.

$$0 \leq v_2 \leq 0.4 \text{ mmol } gCDW^{-1} hr^{-1} \quad \text{(Equation 1.9)}$$

The solution space and the flux distribution with considering the rate-limiting reaction are calculated as shown in Figure 1-7. The biomass yield and the target production yield at the optimal growth state is decreased by considering the rate-limiting reaction, because the flux through the reaction 3 determine the precursor supply for biomass reaction and target synthesis. For overcoming the kinetic constraints causing rate-limiting reaction, further strain improvement is required.

(3) The biomass reaction of genome-scale metabolic model is constructed by measuring composition of biomass components of wildtype microorganism. The biomass compositions might be changed based on their genetic backgrounds. In previous study, Long CP et al. reported that measured molecular weight of wildtype and 22 types of single gene knockout mutant of *E. coli* were 37.3 ± 0.4 C-mmol $gCDW^{-1}$ (65). Therefore, although the difference of composition of biomass reaction is a potential cause leading the inconsistency between FBA prediction and culture result, the effect is considered to be small.

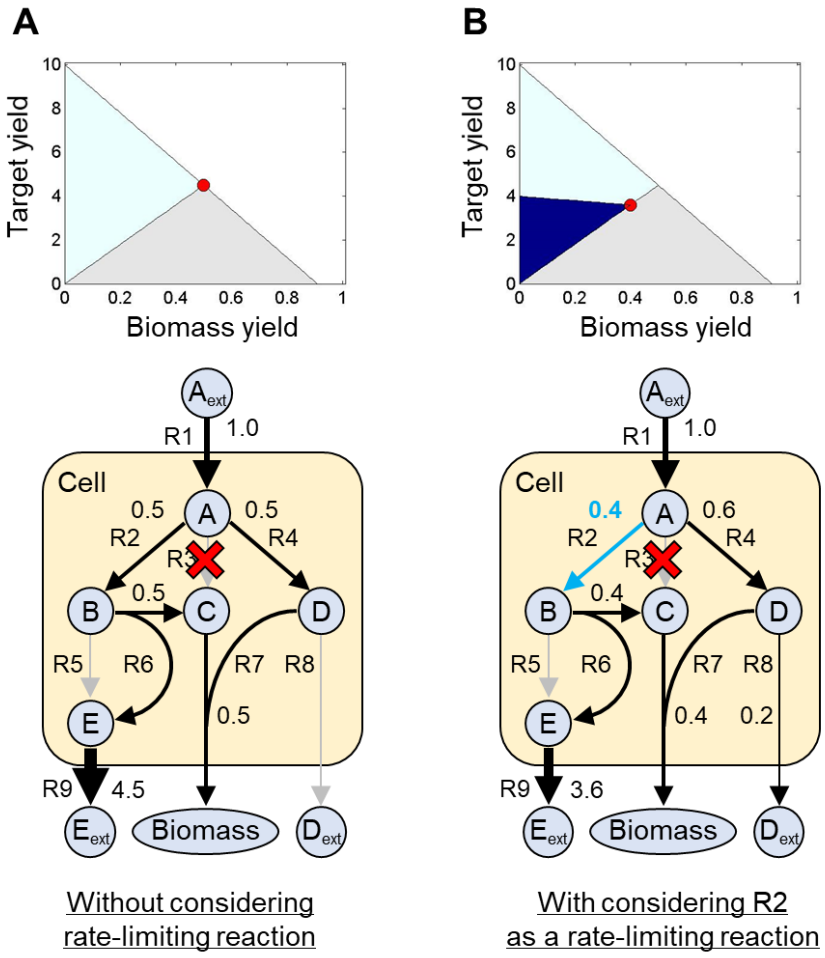



**Figure 1-7 Comparison of solution spaces with considering reaction 2 as a rate-limiting reaction.** (A) Solution space of the simple metabolic network with the disruption of reaction 3, and the flux distribution calculated by FBA. (B) Solution space of the simple metabolic network with the disruption of reaction 3 and considering reaction 2 indicated by blue arrow as a rate-limiting reaction, and the flux distribution calculated by FBA. The solution space and flux distribution was calculated by setting the upper-bound of reaction 3 to 0.4.

## 1-4 General objective

The general objective of this study was to unveil what causes the inconsistencies on growth-coupled target production between FBA prediction and experimental results. For this purpose, novel method to screen possible rate-limiting reaction for the growth-coupled target production based on FBA is developed (chapter 2), and reveal the molecular mechanism causing the rate-limiting steps in the microbial production processes of 3-hydroxypropionic (chapter 3) acid and succinate (chapter 4) through strain improvements.

## 1-5 Outline of the thesis

The thesis consists of 5 chapters, and a schematic outline of the thesis is shown in Figure 1-3.

**Chapter 1** describes the general introduction of this thesis. The background about the microbial process and genome-scale metabolic design are summarized in this chapter. The objectives and schematic of this thesis are also described.

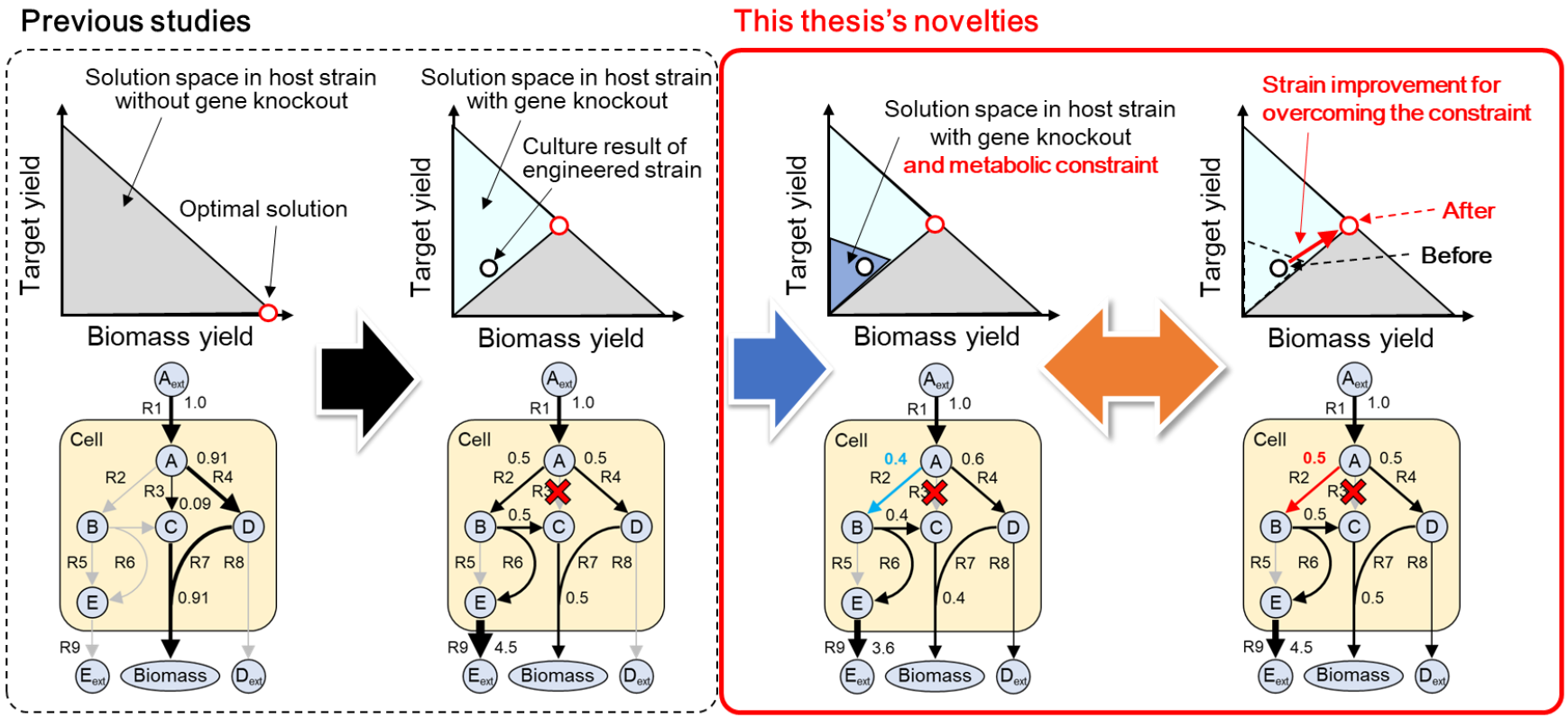
**Chapter 2** describes the development of novel screening method of key enzyme for metabolic engineering by using genome-scale metabolic model. The method assumes the complex metabolic regulating system as an upper-bound constraint on each enzymatic flux on FBA calculation and simulate which reaction is thought to be possible rate-limiting reactions for

the target production. The experimental results of strain improvement for 1,4-butanediol production reported by Yim et al. were consistency with the predicted rate-limiting reactions. And this proposed method also successfully screened the possible rate-limiting reactions for various microbial process of growth-coupled target production. To understanding the mechanism causing the rate-limiting reactions is essential to build metabolic engineering strategies, since the proposed approach cannot deal the molecular mechanism.

**Chapter 3** describes the strain improvement for 3-hydroxypropionic acid production based on genome-scale metabolic design. Gene knockout simulation based on conventional FBA identified the double knockout of *tpiA*-*zwf* improve the growth-coupled 3HP production in *E. coli*. Actually, the production yield of 3HP was improved to 0.20 C-mol/C-mol, which was 4.4-fold higher than before the gene deletions. Next, the metabolic engineering strategy for further improvement of 3HP production in the knockout mutant was build based on experimental evaluation of the strain. Overflow of 1,3-propanediol was thought to led to decrease the flux through 3-hydroxypropionaldehyde dehydrogenase, which was predicted as one of the possible rate-limiting reactions for the 3HP production. This overflow was caused by *yqhD* expression induced by activation of methylglyoxal pathway in the knockout mutant. Additional disruption of *yqhD* successfully increased the production yield of 3HP from 0.20 C-mol/C-mol to 0.34 C-mol/C-mol.

**Chapter 4** describes the strain improvement for succinic acid production based on genome-scale metabolic design. Gene knockout simulation based on conventional FBA identified the five knockout of *adhE-pykAF-gldA-pflB* improve the growth-coupled succinic acid production in *E. coli*. Although the production yield of 3HP was improved to 0.08 C-mol/C-mol, the productivity was lower than of the predicted value of 0.45 C-mol/C-mol at the optimal growth state. 9 enzymatic reactions including phosphoenolpyruvate carboxylase (Ppc) were identified as candidates of rate-limiting reactions that can diminish flux space of the actual metabolic network and hamper the optimization of metabolic state for succinate production. For improving succinate production and understanding the molecular mechanism causing the rate-limiting reaction, adaptive laboratory evolution (ALE) was applied for the knockout mutant, since its metabolic network was designed for growth-coupled succinate production. All of the evolved strains obtained from ALE successfully increased cell growth rate and succinate production yield. Furthermore, the all evolved strains had novel mutations in *ppc*. These mutations expanded the flux space of actual metabolic network for succinate production by desensitizing an inhibition by aspartate on Ppc.

**Chapter 5** describes general conclusion and future perspective of this research




**Problem:** The engineered strain show the lower productivity than of predicted values by FBA.
**Objective:** Demonstrating the hypothesis that some of metabolic constraints diminished flux space and decrease the target production.

Chapter2: Prediction of rate-limiting reaction by considering metabolic constraint as upper bound constraint on FBA.
Chapter3&4: Understanding molecular mechanism causing the rate-limiting reactions by experimental strain improvement and experimental evaluation.

Figure 1-8 The schematic and novelties of this thesis.

# Chapter 2: Prediction of possible rate-limiting reactions for target production based on genome-scale metabolic model

## 2-1 Highlights

- Simple screening method of rate-limiting reactions for growth-coupling target production was developed.
- The proposed method screens the rate-limiting reactions with none of experimental data and only imposing an upper-bound constraint of each enzymatic reaction's flux on conventional FBA calculation.
- Predicted results for 1,4-butanediol production were consistency with the experimental results of stain improvement performed in the previous study.
- The proposed method was applied to screen possible rate-limiting reactions in various fermentation processes.

## 2-2 Introduction

### 2-2-1 Genome-scale metabolic design for growth-coupled target production

One of the goals of metabolic engineering is to optimize a metabolic system for overproduction of a target compound based on the understanding of the system. Recent developments in genome-scale metabolic reconstructions have enabled the simulation of the physiological

characteristics of organisms (66)(67)(68). Flux balance analysis (FBA) is a method to estimate a metabolic flux distribution using a genome-scale metabolic model (44)(46)(45)(47), and it has been widely used to predict the effect of gene knockouts in enhancing the target production (54)(69)(56)(70)(55)(71).

Gene knockout mutant for growth-associated target production can be screened by comparing production yield at its optimal growth state as predicted by FBA (54)(56)(55). The knockout mutant must produce a target metabolite to produce the biomass components with satisfying the mass balance constraints. Several studies successfully achieved strain improvement for production of various target metabolites by using such simulations (57)(58)(59). However, these knockout mutants often showed lower cell growth ability and target production yield as compared to FBA predictions. These inconsistencies should be arising from that FBA simplify the complicated metabolic system by considering only mass-balance equation of the stoichiometry. Since there are several factors regulating metabolism such as enzymatic regulation (32), expression regulation, and thermodynamic feasibility (34) and causing rate-limiting steps, the flux space of actual metabolic network must be diminished from the solution space in FBA. These lead to overestimate the productivity by FBA and that the knockout mutants always display the lower productivity than of the predicted value.

### 2-2-2 Identification of rate-limiting steps to metabolic fluxes in biotechnological process

Overcoming the metabolic limitation is important for strain improvement. Rate-limiting reactions in biotechnological process are caused by low enzymatic activity (31), enzymatic regulation (32), expression regulation (33), and thermodynamic feasibility (34). Screening the rate-limiting step for target production has been achieved by experimental approaches (72)(73)(74)(75). For example, Lu et al. identified rate-limiting step in Coenzyme Q10 biosynthesis by overexpression of individual enzymes involved in quinone modification pathway at different levels (72). Pitera et al. identified rate-limiting step in isoprenoid production by exogenously supplementing a precursor in the culture medium (73). Although such experimental approaches can directly evaluate the metabolic bottleneck *in vivo*, it is difficult to expand the target pathways for evaluation due to experimental cost and technical limitations. Computational approaches using a metabolic model considering kinetics have the potential to overcome the disadvantage.

Kinetic modeling of metabolism quantitatively describes the dynamics of metabolic system. Reconstruction of kinetic models has been achieved by integration of priori knowledges on the network structure, kinetics of their enzymes and kinetic parameters (e.g., rate constants such as $K_m$ or $K_{cat}$ and enzyme concentrations) measured *in vitro* and assumed by using experimental data (76)(77)(78)(79). Sensitivity analysis using kinetic models have the potential to screen the metabolic bottleneck from the entire

metabolic system. For example, Rizk et al. achieved to rank which metabolic reactions mostly limited on aromatic production (78). Andreozzi et al. predicted the gene targets for upregulation or downregulation for enhancing 1,4-butanediol production using the kinetic models (80). The kinetic modeling approaches enable us to evaluate which enzymatic reactions are most sensitive for the target production, and enable to build metabolic engineering strategies, such as up-regulation, down-regulation, and knockout. On the other hand, there are several disadvantages in such conventional modeling approaches for identification of rate-limiting steps.

Firstly, they require considerable information regarding the detailed reaction mechanism including regulation and accurate kinetic parameters to construct the model. Second, they have to validate the model accuracy with a large amount of experimental data of culture results and omics data such as fluxomics, metabolomics, transcriptomics and proteomics. And last, although quantitative reconstruction using kinetic modeling can reproduce the metabolic behavior in the range of used training data sets, estimation of the metabolic state overproducing target metabolite is potentially difficult since the training data sets do not contain the abnormal metabolic state. In these studies, further strain improvement based on the predicted rate-limiting reactions were not experimentally validated (78)(79). Furthermore, genetic engineering tools could not quantitatively change enzymatic properties as expected, and trial-and-error procedures of genetic manipulation are needed to obtain the ideal one in the end. Therefore, simple screening method without any experimental data such as conventional FBA approach is

expected to be useful to screen the first candidate of metabolic reaction for strain improvement.

### 2-2-3 Objective of this chapter

In this chapter, I firstly developed *in silico* screening method for possible rate-limiting reactions for growth-coupled target production by using only existing genome-scale metabolic model without any experimental data. The proposed method simplify the metabolic limitation arisen from the complicated interaction of several factors as an upper-bound constraint of each enzymatic reaction's flux on FBA calculation (Figure 2-1). When imposing upper-bound constraint on a reaction decreased target productivity at the optimal growth state, the reaction is thought to be possible a rate-limiting step. When imposing upper-bound constraint on a reaction do not decrease target productivity at the optimal state, the reaction is thought to be potentially non rate-limiting step. The calculation of the maximum productivity at the optimal state was performed by using a linear programming problem as same as FBA calculation.

For validating the predictions of this method, previously reported the metabolic engineering results for 1,4-butanediol production in *E. coli* was used in this chapter. The proposed approach was also applied to predict the possible rate-limiting reactions for various biotechnological processes such as lactate or succinate production from glucose, and 3-hydroxypropionic acid (3HP) or succinate production from glycerol. Furthermore, molecular mechanism causing the rate-limiting steps and metabolic engineering

strategies were discussed with priori knowledges on metabolic system.

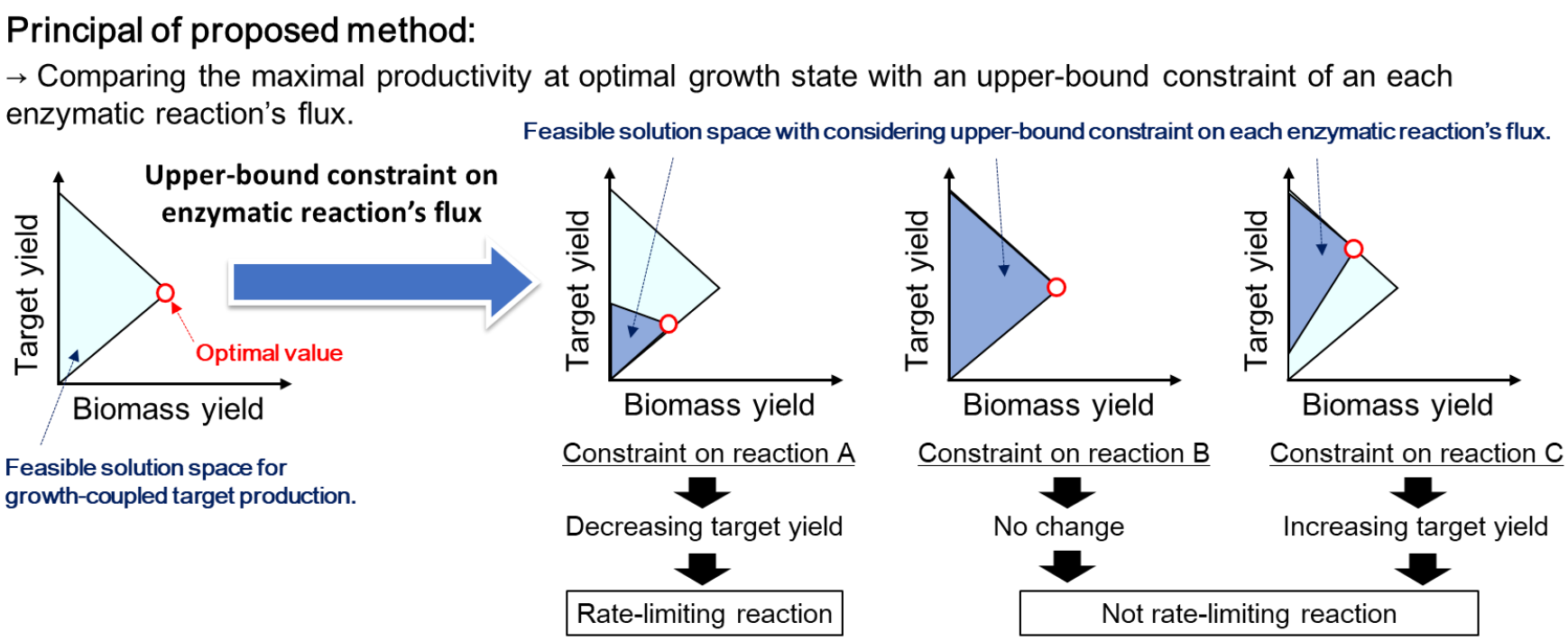



Figure 2-1 The principle of *in silico* prediction of possible rate-limiting reactions using genome-scale metabolic model.

## 2-3 Methods

### 2-3-1 Metabolic model

Genome scale metabolic model of *Escherichia coli* iAF1260 (81) was used in this study. This model contains 1,260 ORFs, 2,077 metabolic and transport reactions, and 1,038 unique metabolites. To simulate heterologous metabolite production of 1,4-BDO or 3HP, synthetic pathway, transport reaction and exchange reactions were added to the genome-scale metabolic model, respectively (Figure 2-2 and 2-3, Table 2-1 and 2-2). In case of simulation of 3HP production, the fluxes of exchange reaction, which serves to uptake compounds to the cell or secrete compounds from the cell, of dihydroxyacetone and D-glyceraldehyde were set to zero due to avoid undetermined production fluxes. Substrate uptake rate (SUR) and oxygen uptake rate (OUR) were respectively set to the values shown in Table 2-3 considering each fermentation process. Other external metabolites such as $CO_2$ and $NH_3$ were allowed to transport freely through the cell membrane. For gene knockout simulations, minimum and maximum fluxes of the corresponding reactions were set to zero.

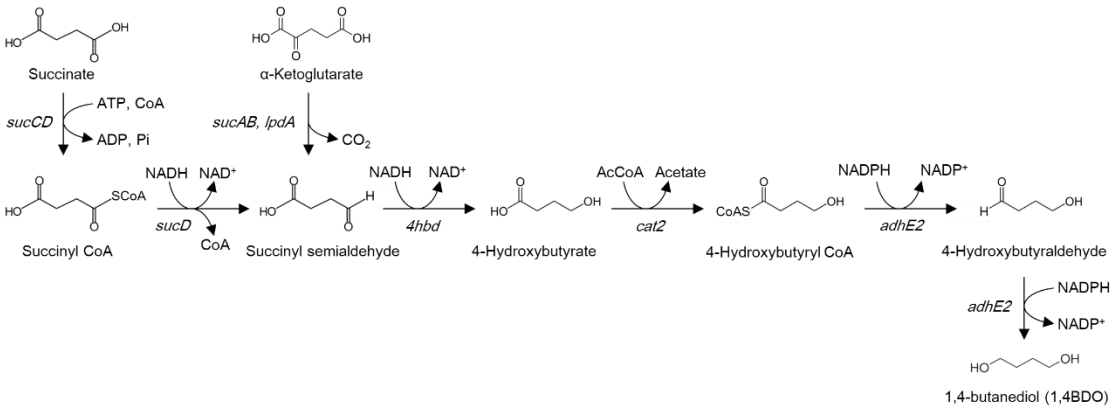



**Figure 2-2 1,4-Butanediol synthetic pathway.** The genes encoding relevant enzymes are given in *italics*.

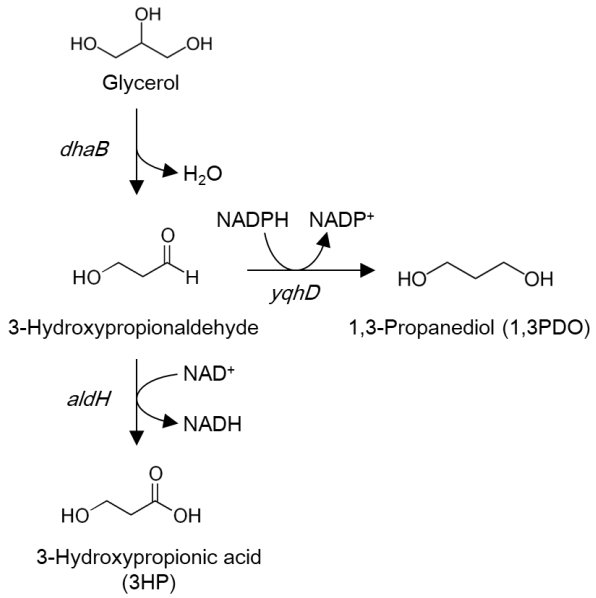



**Figure 2-3 3HP synthetic pathway from glycerol.** The genes encoding relevant enzymes are given in *italics*.

Table 2-1 List of added reactions for 1,4-butanediol production.

| Reaction name | Formular | Gene |
|---|---|---|
| 2-oxoglutarate decarboxylase | akg[c] + h[c] -> co2[c] + sucsal[c] | *M. bovis sucA* |
| CoA-dependent succinate semialdehyde dehydrogenase | h[c] + nadh[c] + succoa[c] -> coa[c] + nad[c] + sucsal[c] | *P. gingivalis sucD* |
| 4-hydroxybutyrate dehydrogenase | h[c] + nadh[c] + sucsal[c] <=> nad[c] + 4hdxbutn[c] | *P. gingivalis 4hbd* |
| 4-hydroxybutyryl-CoA transferase | accoa[c] + 4hdxbutn[c] -> ac[c] + 4hbutcoa[c] | *P. gingivalis cat2* |
| 4-hydroxybutyryl-CoA reductase | h[c] + nadh[c] + 4hbutcoa[c] <=> coa[c] + nad[c] + 4hdxbld[c] | *C. acetobutylicum adhE2* |
| 1,4-butanediol dehydrogenase | h[c] + nadh[c] + 4hdxbld[c] <=> nad[c] + 14btd[c] | *C. acetobutylicum adhE2* |
| 4-hydroxybutyrate transport via proton symport | h[c] + 4hdxbutn[c] <=> h[e] + 4hdxbutn[e] | – |
| 1,4-butanediol transport via diffusion | 14btd[c] <=> 14btd[e] | – |
| 4-hydroxybutyrate exchange | 4hdxbutn[e] <=> | – |
| 1,4-butanediol exchange | 14btd[e] <=> | – |

Table 2-2 List of added reactions for 3-hydroxyporpionic acid production.

| Reaction name | Formular | Gene |
|---|---|---|
| Glycerol dehydratase | glyc[c] -> 3hpa[c] + h2o[c] | *K. pneumoniae dhaB* |
| 3-hydroxypropionaldehyde dehydrogenase | 3hpa[c] + nad[c] + h2o[c] <=> 3hp[c] + nadh[c] + 2h[c] | *Escherichia coli aldH* |
| 3-hydroxypropionic acid transport via proton symport | 3hp[c] + h[c] <=> 3hp[e] + h[e] | – |
| 3-hydroxypropionic acid exchange | 3hp[e] <=> | – |
| 1,3-propanediol oxidoreductase | 3hpa[c] + nadph[c] + h[c] <=> 13pd[c] + nadp[c] | *Escherichia coli yqhD* |
| 1,3-propanediol transporter via diffusion | 13pd[c] <=> 13pd[e] | – |
| 1,3-propanediol exchange | 13pd[e] <=> | – |

Table 2-3 Substrate uptake rate and oxygen uptake rate

| Target metabolites | Substrate uptake rate (mmol $gCDW^{-1}hr^{-1}$) | Oxygen uptake rate* |
|---|---|---|
| 1,4-butanediol | Glucose, 10 | 0 |
| Lactate | Glucose, 10 | 0 |
| Succinate | Glucose, 10 | 0 |
| 3-hydroxypropionic acid | Glycerol, 15 | 10 |
| Succinate | Glycerol, 15 | 10 |

* Oxygen uptake rates on target productions from glucose were set to zero indicating anaerobic condition in refer to the previous studies (52)(57)(58). Oxygen uptake rates on target productions from glycerol were set to 10 indicating micro-aerobic condition because the knockout model cannot grow anaerobically on glycerol medium.

### 2-3-2 Screening algorithm of key enzyme for metabolic engineering

Conventional FBA only considers mass balance constraints on metabolic network and optimizing objective function such as maximum cell growth rate. To predict the possible rate-limiting reactions for target production in the knockout mutant, FBA was performed with some of constraints as shown below:

$$\text{max} \quad \boldsymbol{c}^T \cdot \boldsymbol{v}$$

$$\text{subject to} \quad \sum_{j \in \boldsymbol{R}} S_{i,j} \cdot v_j = 0 \quad (\forall i \in \boldsymbol{M})$$

$$v_{substrate\_uptake} = \text{SUR}$$

$$v_{oxygen\_uptake} = \text{OUR}$$

$$v_{knockout\_reactions} = 0$$

$$v_j \leqq k \cdot \text{SUR} \qquad \text{(if reaction } j \text{ has kinetic constraints.)}$$

$$v_j \geqq k \cdot \text{SUR} \qquad \text{(if reaction } j \text{ is reversible and the reverse reaction has kinetic constraints.)}$$

where $S_{i,j}$ is the stoichiometric coefficient of the metabolite $i$ in the reaction $j$. The $v_j$ is the metabolic flux of enzymatic reaction $j$. $\boldsymbol{M}$ and $\boldsymbol{R}$ are the set of metabolites and reactions, respectively. $\boldsymbol{c}$ is a vector that represents coefficients of an objective function to be maximized or minimized. For determining the maximal succinate production flux at optimal growth, the coefficients were set to 1 and $1\times10^{-6}$ for biomass equation and target metabolite's exchange reaction, respectively. $v_{glycerol\_uptake}$, and $v_{oxygen\_uptake}$ are SUR and OUR, respectively as shown in Table 2-3. $v_{knockout\_reactions}$ corresponds to the knockout reaction for growth-coupled target production. For screening

of the rate-limiting reactions, constant value $k$ was set to 0.2 for glycerol fermentation or 0.4 for glucose fermentation. When a constraint on an enzymatic reaction resulted in decrease target production rate at the optimal growth state, that reaction was considered as a candidate for rate-limiting reaction (Figure 2-1). The calculation was implemented in Matlab (MathWorks Inc., Natick, MA, USA) with a solver for linear programming, Gurobi (http://www.gurobi.com).

## 2-4 Results and discussion

### 2-4-1 Prediction of possible rate-limiting reactions for 1,4-butanediol production from glucose.

Since *E. coli* does not have 1,4BDO synthetic pathway in natural, Yim et al. demonstrated 1,4BDO production from glucose in *E. coli* by expressing heterologous pathway shown in Fig. 2-2 (58). In the study, growth-associated 1,4BDO production was increased by deletion of *adhE-pflB-ldhA-mdhA* based on the FBA prediction. For identification of key enzymatic reaction on the growth-associated 1,4BDO production, the knockout mutant model was reconstructed from *E. coli* iAF1260 by addition of 1,4BDO synthetic pathway and disruption of the reaction encoded by the knockout genes. In the knockout model, glucose was converted to acetate and 1,4BDO via glycolysis and reductive TCA cycle at optimal growth state (Figure 2-4B). Deletion of *adhE, pflB, mdh* and *ldhA* enhanced 1,4BDO production in order to maintain the redox balance.

In the knockout model, 7 enzymatic reactions were predicted as candidates of the rate-limiting reaction for 1,4BDO production (Table 2-5 and Figure 2-4C-E). The limitation of 1,4BDO synthesis pathway directly reduced carbon flow into 1,4BDO formation. The limitation of pyruvate dehydrogenase decreased ATP regeneration by reducing acetate production. This led to decrease 1,4BDO synthesis and cell growth because 1,4BDO synthesis from glucose via reductive TCA cycle cannot produce ATP. The limitation of Ppc reduced the flux into the reductive TCA cycle, and led to decrease 1,4BDO formation.

In the previous study, Yim et al. optimized the metabolic network in the

knockout mutant for enhancing 1,4BDO production by additional gene manipulations. Native *lpdA* gene was replaced by the mutant gene of $lpdA^{D354K}$ derived from *K. pneumoniae*. Dihydrolipoamide dehydrogenase encoded by *lpdA* is a component of pyruvate dehydrogenase, which was predicted as the rate-limiting reaction in the proposed method, and *E. coli* LpdA activity is tightly regulated under the anaerobic condition owing to NADH sensitivity (82). Because the mutant has higher activity under anaerobic condition compared with its wildtype, the *lpdA* replacement is thought to increase the flux into pyruvate dehydrogenase (83). The replacement of native *lpdA* by $lpdA^{D354K}$ in the knockout mutant increased not only 1,4BDO production, but also cell growth ability under the oxygen limited condition. This indicated that production of 1,4BDO in the knockout mutant was coupled with cell growth as predicted by FBA.

Yim et al. also deleted *arcA* to activate pyruvate dehydrogenase, succinyl-CoA synthase and CoA-dependent succinate semialdehyde dehydrogenase, because *arcA* gene product work as a global regulator and repress the expressions of *aceEF* encoding subunits of pyruvate dehydrogenase, *sucCD* encoding succinyl-CoA synthase and *sucD* encoding CoA-dependent succinate semialdehyde dehydrogenase (84)(85)(86). The reactions catalyzed by these enzymes were also predicted as rate-limiting reactions.

Furthermore, the enzymes of downstream pathway of 1,4BDO synthesis, which were also predicted as rate-limiting reactions, were optimized by screening high active enzymes from *Clostridium acetobutylicum* and related

organisms. These genetic manipulations resulted in overcoming the metabolic limitations on growth-coupled 1,4BDO production in the knockout mutant.

A recent study by Andreozzi et al. applied kinetic modeling approach to identify metabolic engineering targets for the enhancement of 1,4BDO production in the engineered strain derived from the knockout mutant (80). In order to compare the predicted reaction, possible-rate limiting reactions were screened under aerobic condition setting OUR to 2 mmol $gCDW^{-1}hr^{-1}$, because the previous study consider aerobic fermentation of 1,4BDO. The predicted rate-limiting reactions were almost consistency with the result of the proposed approach (Table 2-6). The main difference was that the kinetic modeling approach identified the metabolic engineering targets for down regulating, while the proposed method cannot identify those. The other difference was that some of reactions in oxidative TCA cycle was determined as the metabolic engineering targets in the kinetic modeling approach, while the proposed method could not predict that. This was caused by metabolic flux through oxidative TCA cycle was not responsible for 1,4BDO production at the optimal growth state predicted by FBA, since the knockout mutant's metabolic network was design for 1,4BDO production under anaerobic condition (58).

Increasing flux of Ppc is a rational way for further improvement of 1,4BDO production in the engineered strain, because the previous study by Yim et al. did not try to do genetic manipulation for *ppc*. Furthermore, the study using kinetic model by Andreozzi et al. also reported that enhancing

enzymatic activity of Ppc is most positively effect on the 1,4BDO production yield in the engineer strain (80). The way to enhance the Ppc activity was subsequently discussed in the chapter 4

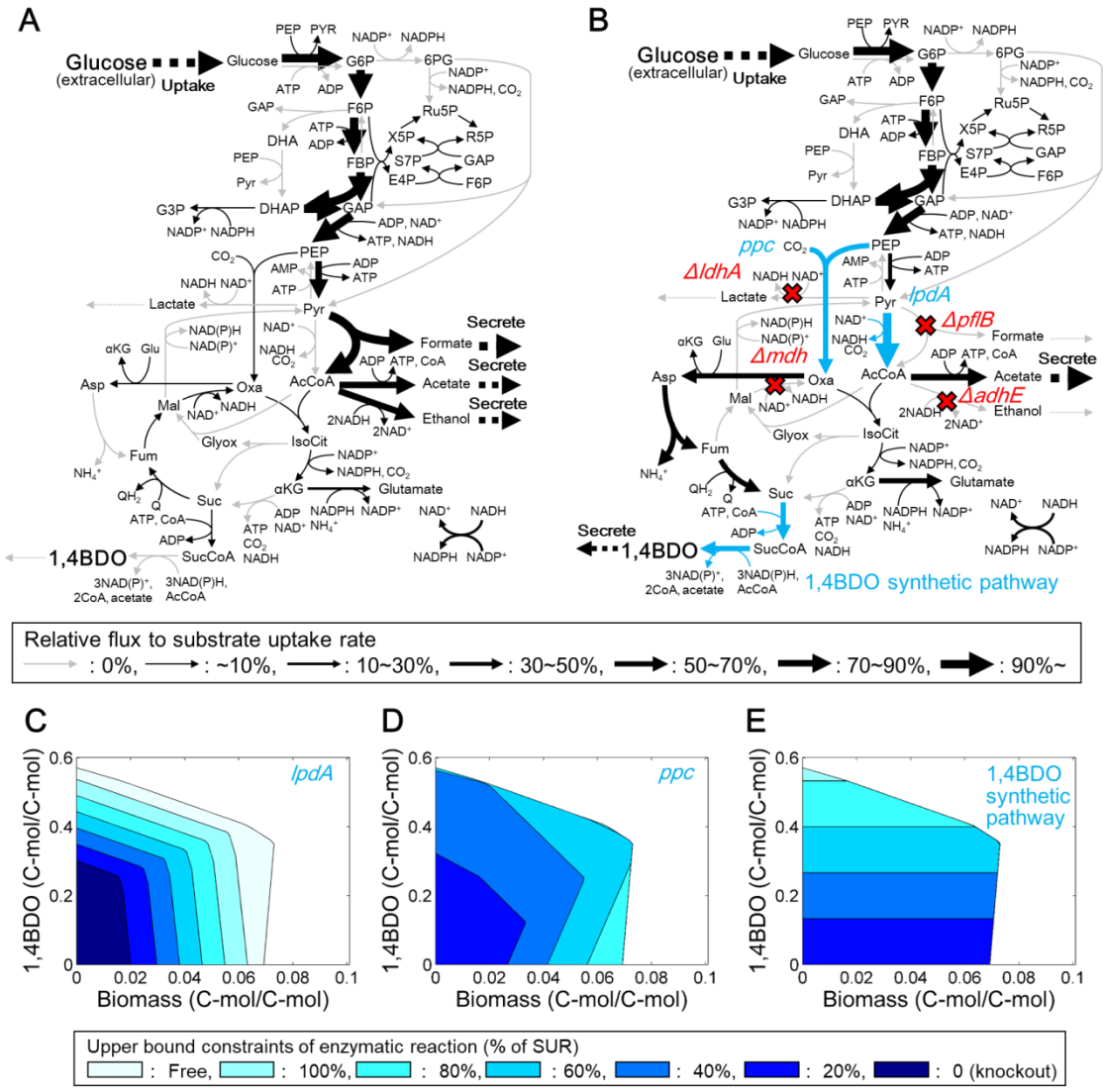



**Figure 2-4 Predicted rate-limiting reactions of growth coupled 1,4BDO production from glucose.** (A, B) Flux distributions of wildtype (A) and *ΔadhE ΔpflB ΔldhA ΔmdhA* (B) were calculated by using iAF1260. Blue arrows indicated predicted rate-limiting reactions. (C-E) Feasible solution space of the knockout mutant with the upper bound constraint of pyruvate

dehydrogenase (C), Ppc (D) and 1,4-butanediol dehydrogenase (E) at 0-100% of SUR and free.

Table 2-5 Predicted rate-limiting reactions for growth-coupling target productions.

| Target | Substrate | Aerobic condition | Knockout genes | Rate-limiting reactions or pathways (Corresponding genes) |
|---|---|---|---|---|
| 1,4BDO | Glucose | Anaerobic | *[1]*adhE, pflB, ldhA, mdh* | Pyruvate dehydrogenase (*lpdA*)<br>Phosphoenolpyruvate carboxylase (*ppc*)<br>1,4-butanediol synthetic pathways (*sucA, sucD, adhE, P. gingivalis 4hbd, P. gingivalis cat2*) |
| Lactate | Glucose | Anaerobic | *[2]*pta, adhE* | Lactate dehydrogenase (*ldhA*) |
| Succinate | Glucose | Anaerobic | *[3]*ptsG, pykAF* | Triosephosphate isomerase (*tpiA*)<br>Phosphoenolpyruvate carboxylase (*ppc*)<br>Reductive TCA cycle (*mdh, fumA*)<br>Acetate synthetic pathway (*pta, ackA*) |
| 3HP | Glycerol | Microaerobic | *[4]*tpiA, zwf* | Glycerol dehydratase (*K. pneumoniae dhaB*)<br>3HPA dehydrogenase (*aldH*) |
| Succinate | Glycerol | Microaerobic | *[4]*adhE, pykAF, gldA, pflB* | Glycerol assimilation pathway (*glpK, gpsA*)<br>Lower glycolysis (*tpiA, gapA, eno, pgk, pgm, lpdA*)<br>Phosphoenolpyruvate carboxylase (*ppc*) |

*1: The knockout genes were predicted in the previous report (58).

*2: The knockout genes were predicted in the previous report (52).

*3: The knockout genes were predicted in the previous report (57).

*4: The knockout genes were predicted in this study

**Table 2-6 A comparison of the proposed method with kinetic modeling approach**

| | Kinetic modeling approach (80) | The proposed method |
|---|---|---|
| Requirement | Reconstruction of kinetic models using stoichiometry, thermodynamics, diverse experimental data (fluxomics, metabolomics, transcriptomics, proteomics and kinetics) | Existing genome-scale metabolic model |
| Type of prediction | Up or down regulation of target genes | Upregulation |
| Prediction method | Metabolic control analysis using ensemble of kinetic models having different kinetic parameters | FBA with an upper-bound constraint of each single enzymatic reaction. |
| Predicted reactions* | | |
| Glycolysis | Upregulation: **ENO**, **PGM**, **Ppc**<br>Downregulation: FBA, PFK, PGI | Upregulation: **ENO**, GAPD, PDH, PGK, **PGM**, **Ppc** |
| TCA cycle | Upregulation: ACONTa, CS, ICDHyr | – |
| 1,4BDO biosynthesis | Upregulation: AKGD, **SUCSALD**, **4HBD**, **4HBCOAT**, **4HBCOAR**, **4HBDH** | Upregulation: **SUCSALD**, **4HBD**, **4HBCOAT**, **4HBCOAR**, **4HBDH** |

*The metabolic reaction names were abbreviations in the genome-scale metabolic model and engineering targets predicted by both method was highlighted in bold.

### 2-4-2 Application of the proposed method for various fermentation process.

The proposed method was applied to identify the possible rate-limiting reactions in the other growth-coupled target productions. The proposed method was successful to screen possible rate-limiting reactions for all case (Figure 2-5). The predicted possible rate-limiting reactions were increased by decreasing the value of *k*, which is constant parameter for the constraint of upper-bound of enzymatic flux. When the *k* was set to 0 (i.e. knockout), unintended reactions which directly involve in *de novo* synthesis of biomass components were predicted. When the *k* was set to 1 (i.e. the upper bound of the flux was set to substrate uptake rate.), no reactions were predicted in some cases. In order to avoid the unintended reactions and evaluate the predicted rate-limiting reaction as same manner, the parameter *k* was set to 0.2 for fermentation on glycerol, or 0.4 for fermentation on glucose by standardizing the intensity of the constraint based on the number of carbon of each substrate.

The calculation time using the genome-scale metabolic model of *E. coli* iAF1260, which contains 1,387 metabolic reactions and 1,038 unique metabolites, was less than 3 minutes by a single Xeon CPU (2.93 GHz). Since the proposed method solves linear programming problems for the metabolic models with constrained upper-bound of each enzymatic reaction's flux, and extracts the possible rate-limiting reactions by comparing the optimal value, the calculation time is dependent on the number of enzymatic reactions in the model. These results indicated that calculation time is less than 0.13 sec per a limitation of each single reaction. The knockout models and the

predicted rate-limiting reactions for the growth-coupled target productions were summarized in Table 2-5. Molecular mechanisms of predicted rate-limiting reactions on target production from glycerol were subsequently discussed in the chapter 3 and 4 thorough experimental strain improvements.

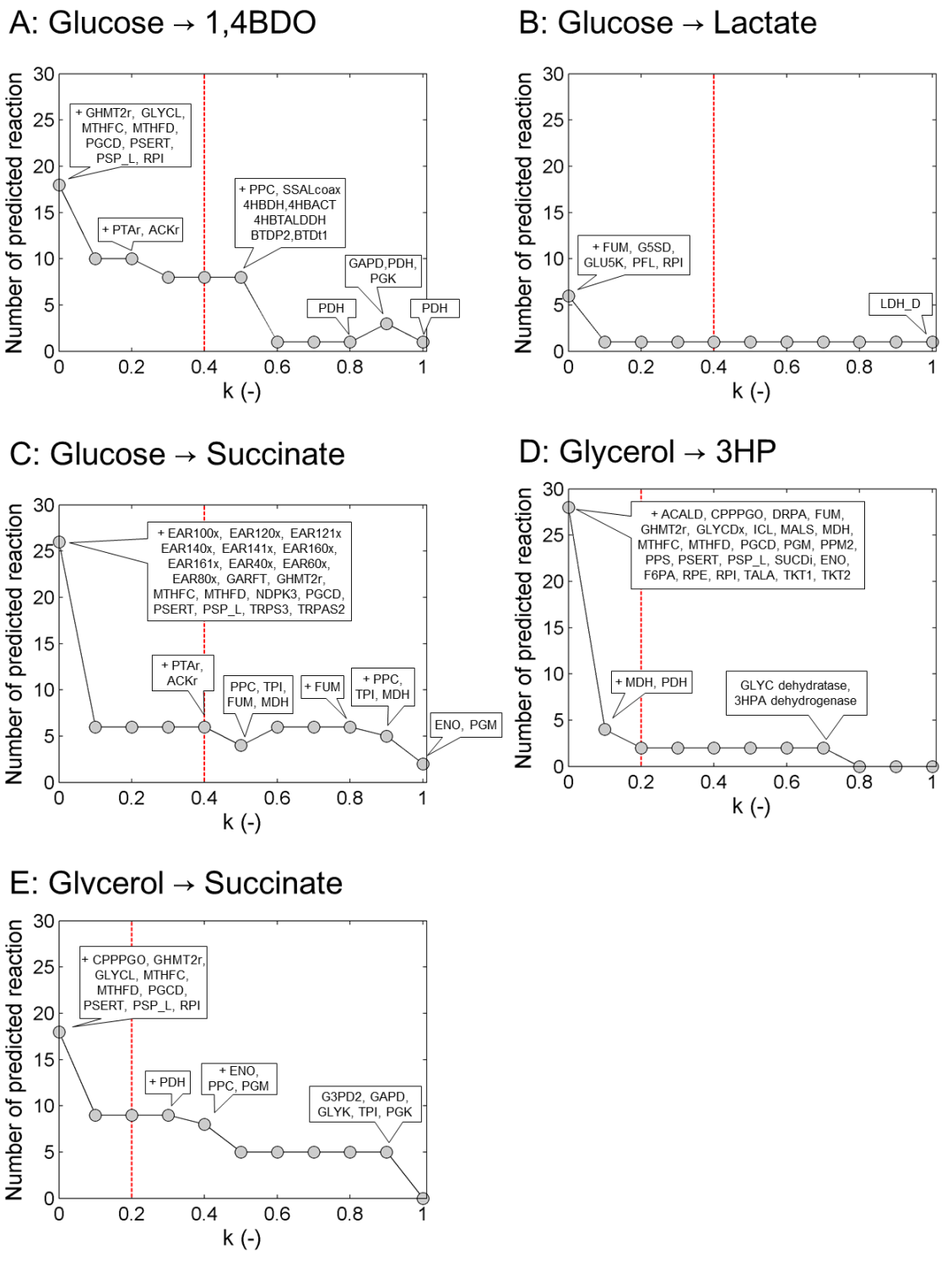



**Figure 2-5 Effect of parameter *k* on the prediction of rate-limiting reactions.**

### Case1: Lactate production from glucose

Lactate is one of anaerobic fermentation products in microorganism. Fong et al. successfully increased lactate production from glucose in *E. coli* by knockout of *pta-adhE* based on FBA prediction (52). Wildtype *E. coli* model produced formate, acetate and ethanol under the anaerobic condition because of re-oxidation of NADH generated in glycolysis (Figure 2-6A). The flux distribution in the knockout model constructed from *E. coli* iAF1260 displayed that the deletion of *pta* and *adhE* increased lactate production in order to generate high amount of ATP in glycolysis and maintain the redox balance (Figure 2-6B). The knockout of *pta* and *adhE* increased the production yield of lactate to 0.61 C-mol/C-mol, which was lower than 0.86 C-mol/C-mol as the predicted yield at the optimal growth state by FBA.

By imposing the upper-bound constraint on the flux of each enzymatic reaction, lactate dehydrogenase was predicted as a sole rate-limiting reaction (Figure 2-6C). Lactate dehydrogenase, which encoded by *ldhA*, converts pyruvate to lactate coupling with NADH oxidation. Fong et al. demonstrated that ALE of the knockout mutant of *pta* and *adhE* successfully increased the growth-coupled lactate production (52). Further experimental analysis of the evolved strains revealed that the ALE experiments optimized the activities in central metabolism for enhancement lactate production (87). Increased phosphoglucose isomerase and phosphofructokinase and decreased glucose 6-phosphate dehydrogenase contributed to improve the flux ratio of glycolytic to oxidative pentose phosphate pathway (Figure 2-7A). The enhanced glycolytic flux improved the supply of NADH and pyruvate as

precursors of lactate dehydrogenase and this led to overcome the kinetic limitation of lactate dehydrogenase, which was predicted as a sole rate-limiting step by the proposed method (Figure 2-7B).

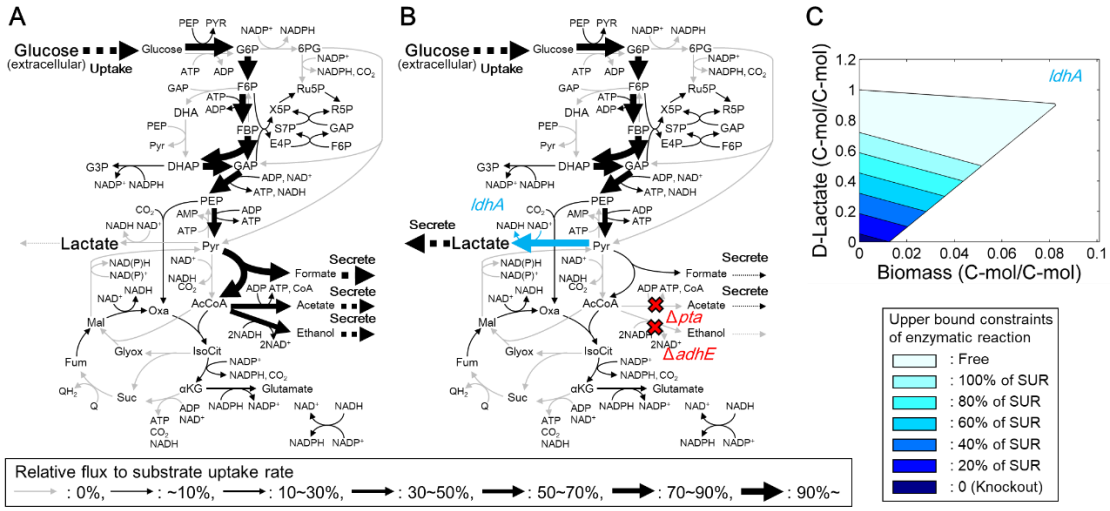



**Figure 2-6 Predicted rate-limiting reactions of growth coupled lactate production from glucose.** (A, B) Flux distributions of wildtype (A) and *Δpta ΔadhE* (B) were calculated by using iAF1260. Blue arrows indicated lactate dehydrogenase predicted as a rate-limiting reaction. (C) Feasible solution space of the knockout mutant with the upper bound constraint of lactate dehydrogenase (C) at 0-100% of SUR and free.

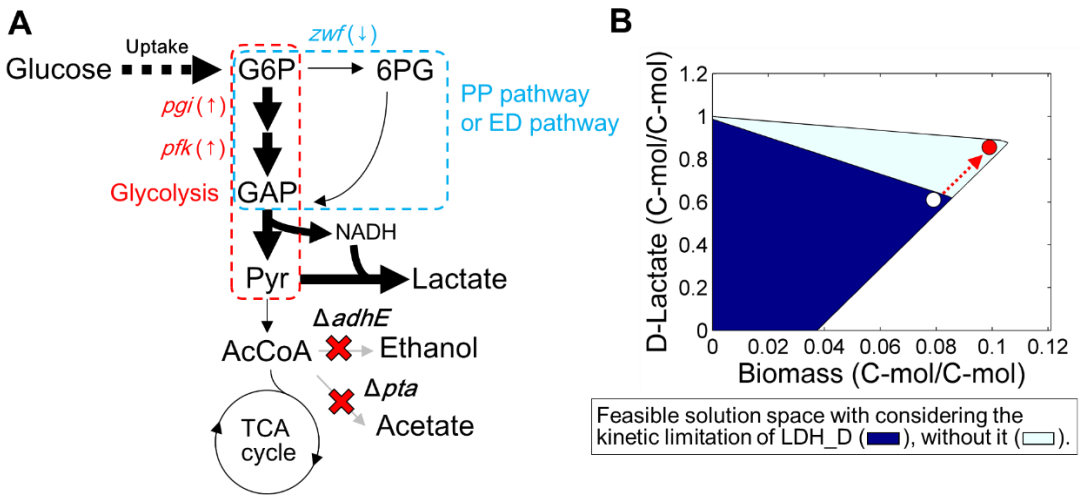



**Figure 2-7 Metabolic optimization for growth-coupled lactate production by adaptive laboratory evolution.** (A) Overview of metabolic adjustment for the lactate production in the knockout mutant of *pta* and *adhE* during adaptive laboratory evolution (87). (B) Direct comparison of feasible solution space and experimental results of the lactate producing strains. White and red circles indicate the experimental results of the unevolved knockout mutant and the average of the all evolved strains obtained in the previous study (52). Metabolic solution space was calculated with the kinetic constant value (*k*) of 1.0 and the measured value of SUR of 16.1 mmol $gCDW^{-1}hr^{-1}$ in iAF1260.

### Case2: Succinate production from glucose

Lee et al. successfully increased succinate production from glucose in *E. coli* by knockout of *ptsG* and *pykAF* based on FBA prediction (57). The flux distribution in the knockout model constructed from *E. coli* iAF1260 displayed that the deletion of *ptsG* and *pykAF* increased succinate production in order to generate high amount of ATP via acetate secretion and maintain the redox balance (Figure 2-8B). The multiple knockout of *ptsG* and *pykAF* increased the production yield of succinate from 0.04 to 0.23 C-mol/C-mol, which was lower than of the predicted yield of 0.46 C-mol/C-mol at the optimal growth state by FBA.

By imposing the upper-bound constraint on the flux of each enzymatic reaction, 6 enzymatic reactions were predicted as rate-limiting reactions for succinate production in the knockout mutant (Figure 2-8C~E and Table 2-5).

The flux limitations of triosephosphate isomerase encoded by *tpiA*, phosphate acetyltransferase encoded by *pta*, and acetate kinase encoded by *ackA* reduced the flux into acetate synthetic pathway and this led to reduce ATP supply and cause redox imbalance. Limiting phosphoenolpyruvate carboxylase encoded by *ppc*, malate dehydrogenase encoded by *mdh*, and fumarase encoded by *fumA, fumB, fumC* directly reduced the flux into succinate synthesis via reductive TCA cycle.

Because the knockout mutant produced large amount of acetate and formate in the flask cultivation, *pta* and *ackA* products should not be rate-limiting reactions. The flux into lower glycolysis such as triosephosphate isomerase is regulated by PEP consuming reaction (88).

Since the knockout mutant does not have pyruvate kinase or phosphotransferase consuming PEP, Ppc is thought to be a possible rate-limiting reaction. Although there is not further experimental strain improvement on the knockout mutant in the study, overexpression of *ppc, mdh* and *fumABC* is a possible way to increase the succinate production.

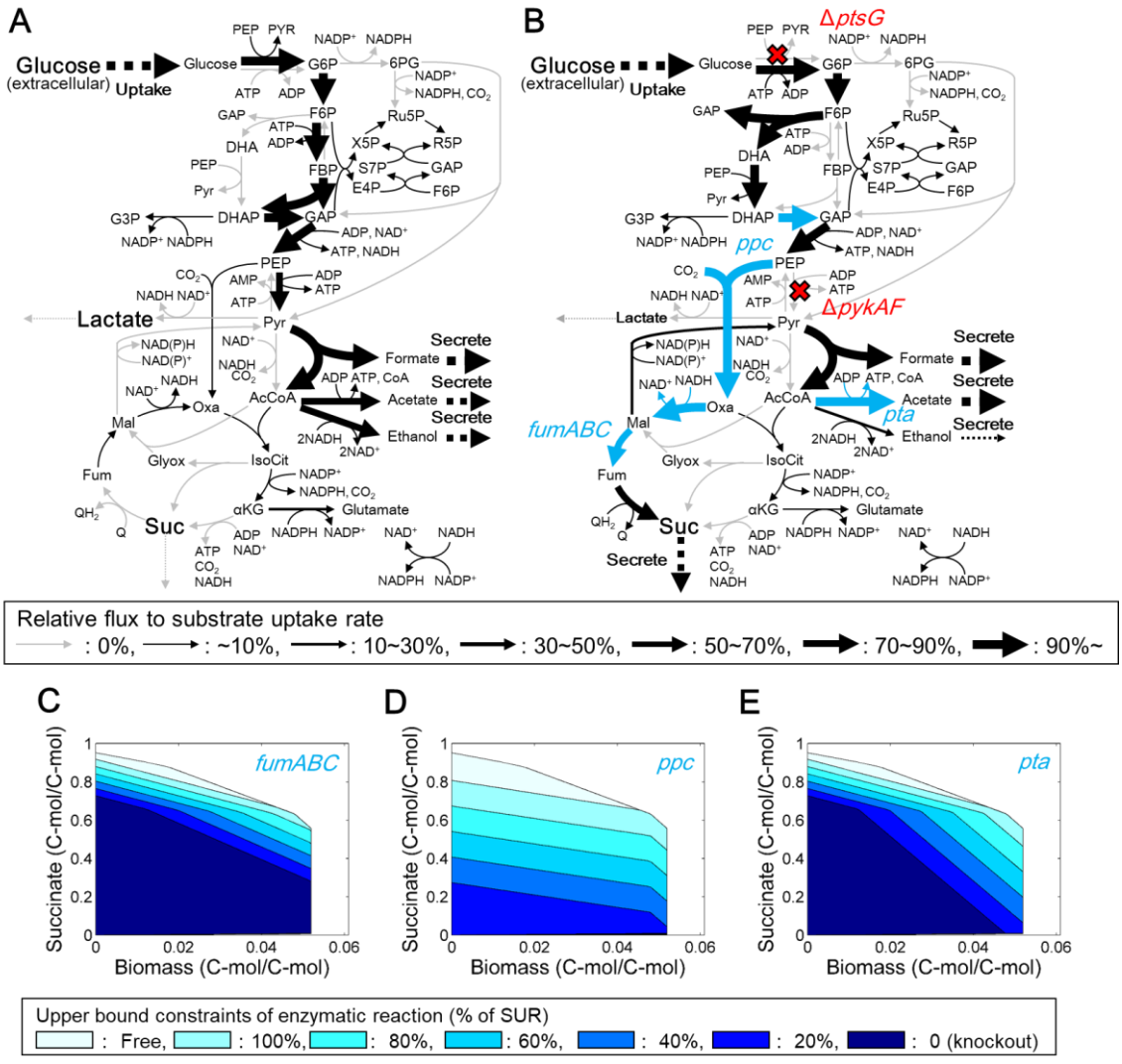



**Figure 2-8 Predicted rate-limiting reactions of growth coupled succinate production from glucose** (A, B) Flux distributions of wildtype (A) and *Δptsg ΔpykAF* (B) were calculated by using iAF1260. Blue arrows indicated

predicted rate-limiting reactions. (C~E) Feasible solution space of the knockout mutant with the upper bound constraint of fumarase (C), phosphoenolpyruvate carboxylase (D) and phosphate acetyltransferase (E) at 0-100% of SUR and free.

### Case3: 3-hydroxypropionic acid production from glycerol

Although wildtype *E. coli* does not have 3HP synthetic pathway, Raj et al. demonstrated 3HP production from glucose in *E. coli* by expressing heterologous pathway shown in Fig. 2-3 (89). In order to simulate the metabolic state of 3HP producing *E. coli*, biosynthetic pathway of 3HP (Figure 2-3 and Table2-2) was added on the genome-scale metabolic model iAF1260. The double knockout models of *tpiA-zwf* was used to identify the rate-limiting reaction for 3HP production. In the all models, 3HP synthesis pathway which involves 2 reactions was predicted as the rate-limiting reactions (Figure 2-9C, Figure 2-9D and Table 2-5).

Glycerol dehydratase is coenzyme $B_{12}$-dependent dehydratase converting glycerol to 3-hydroxypropion aldehyde. Because *E. coli* strains cannot synthesize cobalamin *de novo*, addition of cobalamin to the culture broth is needed to activate the 3HP synthesis (89). Furthermore, expression of *gdrAB* from *K. pneumoniae* is needed to reactivate glycerol-inactivated DhaB and regenerate active enzyme–cyanocobalamin complex for keep activation of glycerol dehydratase in *E. coli* (90)(91)(92). The next reaction of 3-hydroxypropionaldehyde dehydrogenase converts 3-hydroxypropion aldehyde to 3HP. Because accumulation of 3-hydroxypropionaldehyde decrease the cell growth due to its toxicity, the activity of 3-hydroxypropionaldehyde dehydrogenase must be increased by over-expression of the enzyme (89). In addition to overexpression strategy, many reports has tried to increase the activity of 3-hydroxypropionaldehyde dehydrogenase by screening more efficient enzyme from *K. pneumoniae* (93),

*Azospirillum brasilense* (92) and *Cupriavidus necator* (94). Directed evolution approach is also thought to be a possible approach to enhance the activity.

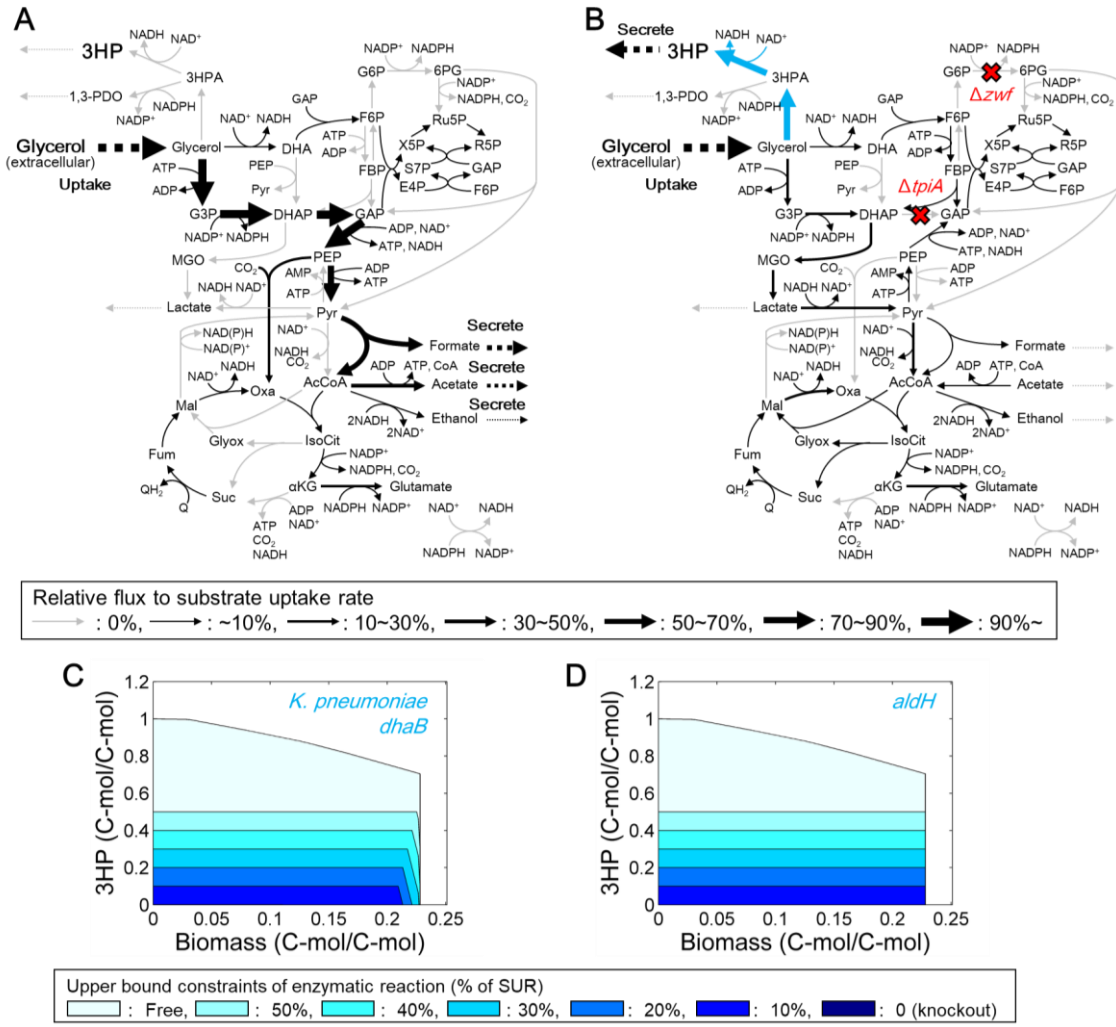



**Figure 2-9 Predicted rate-limiting reactions of growth coupled 3HP production from glycerol.** (A, B) Flux distributions of wildtype (A) and *Δ tpiA Δ zwf* (B) were calculated by using iAF1260. Blue arrows indicated predicted rate-limiting reactions. (C, D) Feasible solution space of the knockout mutant with the upper bound constraint of glycerol dehydratase (C) and

3-hydroxypropionaldehyde dehydrogenase (D) at 0-50% of SUR and free.

### Case4: Succinate production from glycerol

The knockout model of *Δ adhE Δ pykAF Δ gldA Δ pflB* was used to identify the rate-limiting reactions for the growth-coupled succinate production. By limiting the upper bound flux of each reaction to 20% of SUR, 9 enzymatic reactions were screened as candidates of the rate-limiting reaction (Figure 2-10 and Table 2-5). These candidates were classified into three categories according to their subsystems of glycerol assimilation (e.g. glycerol kinase, Figure 2-10C), glycolysis (e.g. pyruvate dehydrogenase, Figure 2-10D), and anaplerotic reaction (phosphoenolpryvate carboxylase, Figure 2-10E).

Category 1: glycerol assimilation

Glycerol kinase (GLYK) is encoded by *glpK* and converts glycerol to glycerol-3-phosphate. Because the GLYK is the sole entrance reaction to central carbon metabolism, in the knockout mutant of *gldA* the upper bound constraints of GLYK directly suppresses the flux into central carbon metabolism, thereby decreasing succinate production and biomass formation.

Category 2: glycolysis

Pyruvate dehydrogenase (PDH) is encoded by *aceEF* and *lpdA,* and converts Pyr to acetyl-CoA. Suppression of PDH decreases NADH supply and ATP generation thereby decreasing acetate synthesis, leading to a decrease in cell growth rate and succinate production.

Category 3: anaplerotic reaction

Ppc is encoded by *ppc*, and catalyzes an anaplerotic reaction from PEP to oxaloacetate. Since Ppc reaction is an entrance of the reductive TCA cycle under microaerobic condition, the suppression of Ppc led to decrease in succinate production. Furthermore, Ppc inhibition also led to decrease in cell growth rate since malate dehydrogenase in reductive TCA cycle is a main NADH oxidation reaction.

The candidates of rate-limiting reactions for succinate production should be regulated by several factors in an actual cell. The expression of *glpK* is regulated by transcriptional repressor encoded by *glpR*. Previous report said the deletion of *glpR* increased the expression of glycerol kinase and glycerol uptake rate (95). *In vitro* analysis revealed that the activity of glycerol-3-phosphate dehydrogenase is allosterically regulated by glycerol 3-phosphate (96). Expressions of the subunits of PDH were regulated by several transcriptional factors of *arcA*, *cra*, *fnr*, *nsrR* and *pdhR* (97)(98). The multi-enzyme complex of PDH also needs $Mg^{2+}$ and vitamin B1 to catalyze the reaction (99). Ppc is allosterically activated by acetyl-CoA, fructose 1,6-bisphosphate and guanosine triphosphate, and also inhibited by malate and aspartate (100)(101). There are several ways to optimize the metabolic network for growth-coupled succinate production, since the number of predicted rate-limiting reactions was larger than of other target production process, and a lot of factors involved in the regulation of the predicted enzymes. In such case, evolution engineering such as ALE is a possible way

to increase target production coupled, since the metabolic pathway of the knockout mutant was designed for growth-coupled target production

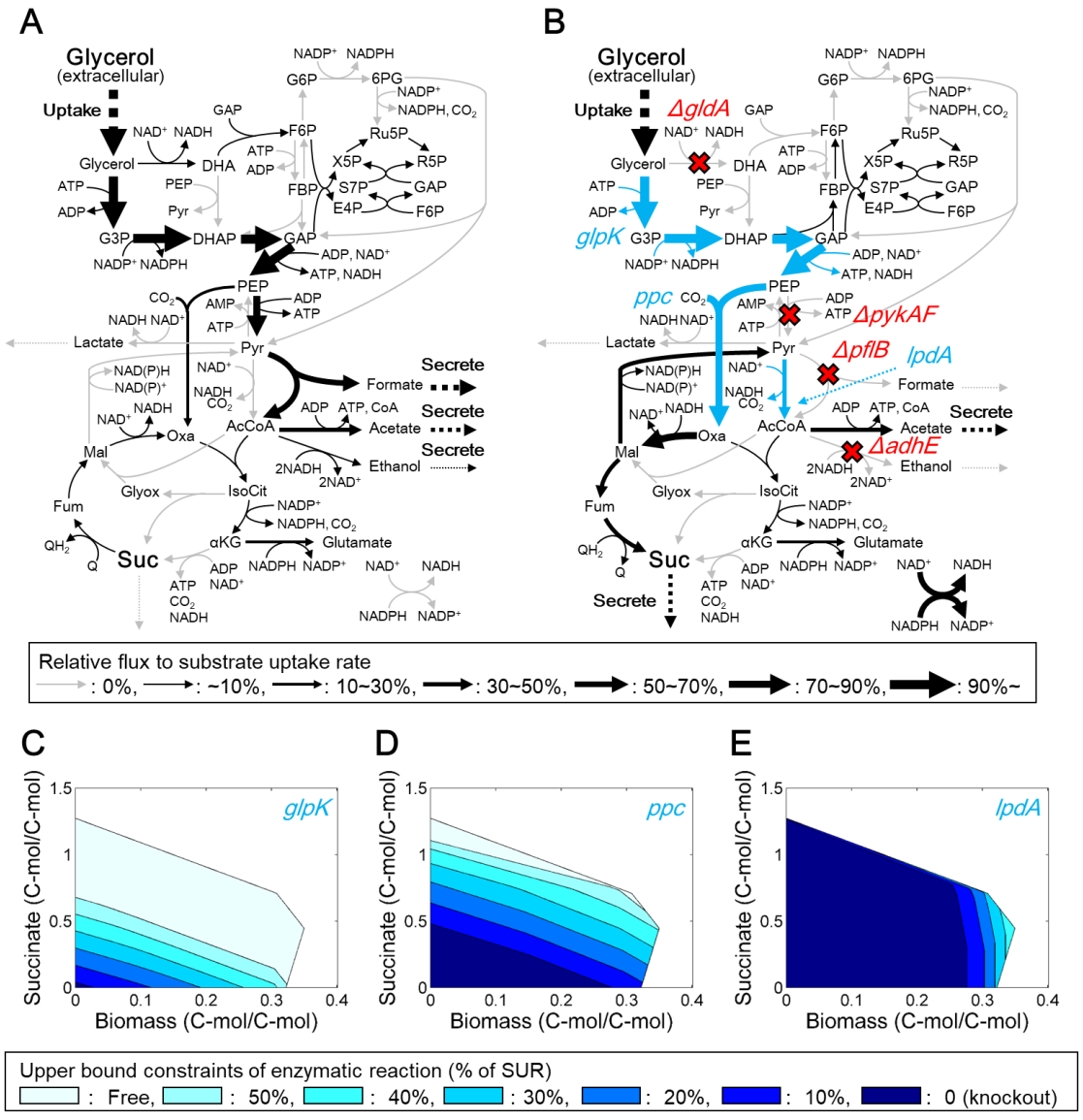



**Figure 2-10 Predicted rate-limiting reactions of growth coupled succinate production from glycerol.** (A, B) Flux distributions of wildtype (A) and *Δadh EΔpykAFΔgldAΔpflB* model (B) were calculated by using iAF1260. Blue arrows indicated predicted rate-limiting reactions. (C~E) Feasible

solution space of the knockout mutant with the upper bound constraint of glycerol kinase (C), phosphoenolpyruvate carboxylase (D) and pyruvate dehydrogenase (E) at 0-50% of SUR and free.

## 2-5 Summary

FBA has been used to predict the behavior of metabolic network and evaluate the effect of gene-manipulation on target production. However, the knockout mutant for enhancing the growth-coupled target production based on FBA prediction did not achieve the high target producing state at the optimal growth state (Table 1-1). One possible cause of these inconsistencies should be certain kinetic constraints, which hampered to optimize metabolic network towards the optimal growth state. In the present chapter, I developed simple screening method of possible rate-limiting reactions by simplifying the kinetic constraints as an upper-bound constraint of each enzymatic reaction's flux on FBA calculation. The proposed method successfully predicted the rate-limiting reactions for growth coupled production of various metabolites. Especially, in case of 1,4BDO production from glucose, the predicted rate-limiting reactions were consistency with the experimentally optimized reaction by additional genetic manipulations (58).

Conventional modeling approaches for identification of rate-limiting reactions for target production use the dynamic metabolic model such as mass action equation (78) (80). These approaches must consider the detail information of *in vivo* kinetic parameters and metabolic regulation system for prediction of metabolic behavior in accurate. Therefore, the prediction accuracy is dominant on the quantity and quality of these information. Needless to say, full kinetic representation in whole cell level is thought to be difficult at the present, since there are a lot of unknown metabolic regulations, even if in well-characterized microorganism such as *E. coli*.

Furthermore, kinetic parameters are thought to be different between culture conditions and strain backgrounds. On the other hand, the present method can screen possible rate-limiting reaction with none of experimental data and only using existing genome-scale metabolic model. Considering the complex kinetic regulation as a simple assumption of upper-bound constraint, the present approach successfully extracted the possible rate-limiting reactions from all enzymatic reactions in several microbial production processes.

Another advantage of the present approach is applicable for various microorganisms besides *E. coli*, because there are various type of genome-scale metabolic models for fermentative microorganism such as *Corynebacterium glutamicum* (49), Yeast (102), Cyanobacteria (103) and Lactic acid bacterium (104). The other advantage of this approach is small computational costs compared with the dynamic modeling approach, since the effect of the upper-bound constrain on target production is calculated as a linear programming problem. The calculation time for comprehensive analysis of single enzymatic reaction is less than 3 minutes by a single Xeon CPU (2.93 GHz).

As shown in several fermentation processes for growth-coupled target production, the present approach successfully extracted the rate-limiting reactions from all enzymatic reactions without consider the detail mechanism of kinetic regulation. These results implied that the behavior of metabolic network is dependent on the network structure rather than the regulation system. One of the disadvantages is that the regulating

mechanism on the rate-limiting reaction cannot identified by this approach. Therefore, priori knowledges of metabolism or experimental approach for understanding the mechanism are needed to build the next metabolic engineering strategy. The optimization strategy of predicted rate-limiting is discussed in the following chapters.

# Chapter 3: Strain improvement for 3-hydroxypropionic acid production from glycerol in engineered *E. coli* based on flux balance analysis

## 3-1 Highlights

- Microbial production of 3-hydroxypropionic acid has been attracted recently, due to its wide-range availability for commercial products.
- Based on *in silico* metabolic design by FBA, the effect of double gene deletion of *tpiA* and *zwf* on 3HP producing *E. coli* was evaluated by flask cultivation.
- Overflow of 1,3-propanediol (1,3PDO) from 3-hydroxypropionaldehyde is a possible reason to decrease the flux into 3-hydroxypropionaldehyde dehydrogenase, which was predicted as a rate limiting reaction.
- Additional deletion of *yqhD* encoding NADPH-dependent aldehyde reductase successfully improved 3HP production at 0.34 C-mol/C-mol by rewiring carbon flux from 1,3PDO overflow to 3HP synthesis.

## 3-2 Introduction

### 3-2-1 Introduction of 3-hydroxypropionic acid

3-hydroxypropionic acid (3HP) has recently attracted attention due to its availability as a precursor of valuable chemicals such as acrylic acid, ß-propiolactone, and malonic acid (105) (106). Its polymerized form, poly(3HP), is a promising alternative to petrochemical-derived plastic (107) (108). Because of this superior industrial availability, 3HP was designated as one of the top value-added chemicals produced by biomass, by the U.S. Department of Energy (3)(4). In the commercial bioproduction process, the substrate has a significant impact on production cost. Glycerol is a potential substrate for bioproduction considering that the recent expansion of biodiesel production has caused a surplus of glycerol as its byproduct, and a subsequent decrease in the price of glycerol (109)(110)(111).

### 3-2-2 Microbial production of 3-hydroxypropionic acid

The microbial production of 3HP from glycerol has been developed using a natural 3HP producer, *Klebsiella pneumoniae* (112)(113)(114)(93). Expression of the heterologous glycerol dehydratase and aldehyde dehydrogenase enabled 3HP to be produced in the non-natural 3HP producers *Pseudomonas denitrificans* (115) (116) and *Escherichia coli* (89)(92)(117)(118)(95). To date, various studies have reported increased 3HP production as summarized in Table 3-1. For example, Rathnasingh et al. optimized the expression level of each enzyme in this pathway in *E. coli* (92). Ashok et al. deleted *dhaT* and *yqhD* to reduce the production of byproducts

in *K. pneumoniae* (114). Kim et al. modified glycerol metabolism in 3HP-producing *E. coli* and developed fed-batch cultivation with simultaneous feeding of glycerol and glucose (118). As described above, most previous studies focused on the optimization of metabolic reactions from glycerol to 3HP and culture conditions. Considering the whole metabolic network, modification of other pathways, as well as the biosynthetic pathway of the target product, is also a key strategy for increasing the metabolic flux, leading to enhanced target production. For example, target production can be enhanced by improving the redox balance and rewiring carbon flow into biosynthetic pathway of target product via gene knockout and overexpression (114)(119).

**Table 3-1 Comparison of 3HP production using glycerol as a carbon source**

| Organism | Strategy | Yield (C-mol / C-mol) | Titer (mM) | Reference |
|---|---|---|---|---|
| *E. coli* | Expression of aldehyde dehydrogenase derived from *E. coli* and glycerol dehydratase from *K. pneumoniae* | 0.19 | 6.48 | (89) |
| *E. coli* | Optimization of culture condition | 0.39 | 50.4 | (117) |
| *E. coli* | Expression of α-ketoglutaric amialdehyde dehydrogenase derived from A. brasilense and glycerol dehydratase reactivase derived from *K. pneumoniae* | 0.40 | 31.1 | (92) |
| *E. coli* | Disrupted of byproduct synthetic pathways and activation of glycerol metabolism and fed-batch culture | 0.26 | 467.3 | (95) |
| *K. pneumoniae* | Disruption of byproduct synthetic pathway and fed-batch culture. | 0.34 | 145.2 | (114) |
| *K. pneumoniae* | Co-production of 3-hydroxypropionic acid and 1,3-propanediol in fed-batch culture | 0.41 | 542.8 | (120) |
| *P. denitrificans* | Expression of glycerol dehydratase pathway genes | 0.67 | 54.7 | (115) |
| *P. denitrificans* | Disruption of 3HP degradation genes | 0.78 | 33.1 | (116) |

### 3-2-3 Genome-scale metabolic design for 3-hydroxypropionic acid production

Recently, *in silico* metabolic simulation has been developed to consider whole metabolic networks. A genome-scale metabolic model, which includes most of the metabolic reactions of the cell (81)(103)(49)(104)(102), can estimate the flux distribution of the whole metabolic network using FBA (46)(121) by assuming the steady states of metabolic reactions and maximizing objective functions such as cell growth (67)(122). This method can be used to simulate the effects of gene modifications on target production and identify candidate genes for metabolic engineering (54)(70)(55). However, the FBA prediction always overestimates the productivities in the actual cell as shown in Table 1-1. One of the possible reasons of the inconsistency is that FBA calculation does not consider any kinetic constraints in actual cells. In the chapter 2, the screening algorithm for the rate limiting reaction was developed by imposing an upper-bound constrain of each enzymatic reaction's flux. The disadvantage of this approach could not reveal the molecular mechanism causing the rate limiting reaction. Therefore, a strain engineered on the basis of metabolic simulation should be evaluated experimentally to understand the mechanism for building the next strategy for improving target production.

### 3-2-4 Objective of this chapter

In the chapter 2, I identified the possible rate-limiting reactions for enhancing growth-coupled 3HP production by developing the novel screening method using FBA. Understanding the molecular mechanism causing

rate-limiting reaction is important for building metabolic engineering strategy of genetic manipulation for further strain improvement. The current chapter aimed to improve 3HP production based on the FBA prediction and experimental evaluation for understanding the molecular mechanism. First, I evaluated the effect of the deletion of predicted genes on 3HP production and estimated the molecular mechanism causing the rate limiting reaction. Then, I aim to releasing the metabolic limitation based on the molecular mechanism and improve target production by additional genetic manipulation.

## 3-3 Materials and methods

### 3-3-1 *In silico* screening of knockout gene targets for enhancing 3HP production

FBA is a method to estimate a metabolic flux distribution using a genome-scale metabolic model (44)(45)(46)(47). Since the genome-scale metabolic model iAF1260 does not contain biosynthetic reaction of 3HP, the reactions (Figure 2-3 and Table 2-2) were added as shown in the chapter 2. To identify the candidates for gene knockout to achieve growth-coupling production of 3HP from glycerol, FBA was performed by using the iAF1260 models containing 3HP synthetic reactions (iAF1260-3HP) as shown below:

$$\max \quad \boldsymbol{c}^T \cdot \boldsymbol{v}$$

$$\text{subject to} \quad \sum_{j \in \boldsymbol{R}} S_{i,j} \cdot v_j = 0 \quad (\forall \boldsymbol{i} \in \boldsymbol{M})$$

$$v_{substrate\_uptake} = \mathrm{SUR}$$

$$v_{oxygen\_uptake} = \mathrm{OUR}$$

$$v_{knockout\_reactions} = 0$$

where $\boldsymbol{M}$ and $\boldsymbol{R}$ are the set of metabolites and reactions, respectively. $\boldsymbol{c}$ is a vector that represents coefficients of an objective function to be maximized or minimized. Glycerol was used as the sole carbon source in metabolic simulations, and glycerol uptake rate (GUR, $v_{glycerol\_uptake}$) was set to 15 mmol $gCDW^{-1}hr^{-1}$. The oxygen uptake rate (OUR, $v_{oxygen\_uptake}$) was set to 10 mmol $gCDW^{-1}hr^{-1}$. Other external metabolites such as $CO_2$ and $NH_3$ were allowed to transport freely through the cell membrane. For determining the maximal 3HP production flux at optimal growth, the objective coefficients were set to 1 and $1\times10^{-6}$ for biomass recation and 3HP exchange reaction, respectively. After the FBA calculations for multiple reaction knockout models, the models showing 3HP production at the optimal growth solution were identified as growth-coupling target model. The calculation was implemented in Matlab (MathWorks Inc., Natick, MA, USA) with a solver for linear programming, Gurobi (http://www.gurobi.com).

### 3-3-2 Strains and plasmids

The strains and plasmids used in this study are summarized in Table 3-2 and 3-3. The MG1655(DE3) strain was constructed based on *E. coli* MG1655, using the λDE3 Lysogenization Kit (Merck KGaA, Darmstadt, Germany). The 3HP-producing strain, TK52, was constructed as published previously (92). Fragments containing *dhaB1*, *dhaB2*, and *dhaB3*, encoding for components of glycerol dehydratase, and the *gdrA* and *gdrB* genes, encoding for glycerol dehydratase reactivase, were amplified from the genomic DNA of *K. pneumoniae* subsp. *pneumoniae* (NBRC 14940), which was purchased from the National Institute of Technology and Evaluation (Tokyo, Japan). Fragments were generated by PCR using the primer pair 5′-CCGGAATTCATGAAAAGATCAAAACGATTTGCAGTACT-3′ and 5′-GTTAAGCTTGATCTCCCACTGACCAAAGCTGG-3′ for *dhaB* (*dhaB1*, *dhaB2*, *dhaB3*) and *gdrA* and the primer pair 5′-GAAAAGCTTGAGGGGGACCGTCATGTCGCTTTCACCGCCAG-3′ and 5′-GCGCTTAAGTCAGTTTCTCTCACTTAACGGC-3′ for *gdrB*. The *aldH* gene was amplified from the genomic DNA of *E. coli* MG1655 with the primer pair 5′-GGAGGATCCATGAATTTTCATCATCTGGC-3′ and 5′-TCGAAGCTTTCAGGCCTCCAGGCTTAT-3′. PCR was performed using KOD FX Neo (Toyobo Co., Ltd., Osaka, Japan). Each amplified fragment was treated with A-attachment mix (Toyobo Co., Ltd.), and then cloned into pGEM-T easy (Promega Co., Madison, WI, U.S.A.), followed by sequence confirmation by the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Inc., Foster City, CA, U.S.A.), and the 3130 Genetic Analyzer

(Applied Biosystems). The *Eco*RI-*Hin*dIII and *Hin*dIII-*Afl*II fragments carrying *dhaB-gdrA* and *gdrB*, respectively, on pGEM-T easy were cloned into the same restriction sites in pCDFDuet-1 (Merck KGaA), generating pCDFDuet/*dhaB-gdrAB*. In addition, the *Bam*HI-*Hin*dIII fragment carrying *aldH* on pGEM-T easy was cloned into the pTrc99A expression vector (Pharmacia, Stockholm, Sweden), generating pTrc99A/*aldH*.

The knockout strains were constructed using Wanner's method (25) and P1*kc*-mediated phage transduction (123). For the deletion of *yqhD*, the disruption cassette, including the tetracycline resistance gene and homologous regions upstream and downstream of *yqhD*, was amplified from pKD13tet (124) by PCR with the primer pair 5′-GCAGATCGTTCTCTGCCCTCATATTGGCCCAGCAAAGGGAGCAAGTA ATGATTCCGGGGATCCGTCGACC-3′ and 5′-CGAAAACGAAAGTTTGAGGCGTAAAAAGCTTAGCGGGCGGCTTCGTA TATTGTAGGCTGGAGCTGCTTCG-3′. The disruption cassette was introduced into the BW25113/pKD46 strain to construct BW25113*ΔyqhD*::*tet*. To delete *zwf*, *tpiA*, and *yqhD* in *E. coli* MG1655(DE3), P1 transduction was performed using P1 phage obtained from JW3890 (26), BW25113*Δzwf*::*cat* and BW25113*ΔyqhD*::*tet* strains, respectively. Finally, the plasmids pTrc99A/*aldH* and pCDFDuet/*dhaB-gdrAB*, were introduced into MG1655(DE3) and the knockout strains to construct the 3HP-producing strains (Table 3-2).

**Table 3-2 List of strains used in this study**

| Strains | Description |
|---|---|
| DH5α | $F^-$, Φ80d *lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsdR*17($rK^-$ $mK^+$), *phoA*, *supE*44, λ−, *thi*−1, *gyrA*96, *relA*1 |
| BW25113 (25) | $F^-$, $\lambda^-$, $lacI^q$ $rrnB_{T14}$ Δ$lacZ_{WJ16}$ *hsdR514* Δ*araBA-D*$_{AH33}$ Δ*rhaBAD*$_{LD78}$ |
| JW3890 (26) | The same as BW25113 but Δ*tpiA*::*kan* |
| BW25113 *Δ zwf*::*cat* (124) | The same as BW25113 but Δ*zwf*::*cat* |
| BW25113 *Δ yqhD*::*tet* | The same as BW25113 but Δ*yqhD*::*tet* |
| MG1655 | $F^-$,$\lambda^-$, *rph-1* |
| MG1655(DE3) | $F^-$, $\lambda^-$, *rph-1,* λ (DE3[*lacI lacUV5-T7 gene* 1 *indl sam7 nin5*]) |
| TK52 | MG1655(DE3) transformed with pTrc99A-*aldH* and pCDFDuet-*dhaB gdrAB* |
| TK52t | The same as TK52 but *Δtpi*::*kan* |
| TK52z | The same as TK52 but *Δzwf*::*cat* |
| TK52tz | The same as TK52 but *ΔtpiA*::*kan Δzwf*::*cat* |
| TK52tzy | The same as TK52 but *ΔtpiA*::*kan Δzwf*::*cat ΔyqhD*::*tet* |

**Table 3-3 List of plasmids used in this study.**

| Plasmids | Description |
|---|---|
| pGEM-T easy | *lacZ*α; cloning vector; pGEM 5zf(+) derivative; 3'T-overhang; $Amp^r$ |
| pKD46 | λ Red recombinase expression; R101 ori; $Amp^r$ |
| pKD13-*tet* | PCR templete; R6K-ori; FLP-*tetA*-FLP; $Amp^r$ |
| pCDFDuet-1 | *lacI*; expression vector; T7 promoter; CloDF13-ori; two sets of MCS; MCS I, His6-N; MCS II, S-tag-N; $Str^r$ |
| pTrc99A | *lacI*; expression vector; *trc* promoter; ColE1-ori; one MCS; $Amp^r$ |
| pCDFDuet-*dhaB gdrAB* | *dhaB1*, *dhaB2*, *dhaB3*, *gdrA*, and *gdrB* derived from *K. pneumoniae* in pCDFDuet-1 vector; $Str^r$ |
| pTrc99A-*aldH* | *aldH* derived from *E. coli* in pTrc99A vector; $Amp^r$ |

### 3-3-3 Medium and culture methods

Pre-cultures were grown aerobically at 37°C overnight in Lennox medium (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl, 1 g/L glucose) containing 0.05 g/L ampicillin and 0.05 g/L streptomycin. Pre-cultures were transferred to the main culture with an initial optical density at 600 nm ($OD_{600}$) of 0.02. For the main culture, M9 medium (17.1 g/L $Na_2HPO_4 \cdot 12H_2O$, 3.0 g/L $KH_2PO_4$, 2.0 g/L $NH_4Cl$, 0.5 g/L NaCl, 0.123 g/L $MgSO_4 \cdot 7H_2O$, 0.00278 g/L $FeSO_4 \cdot 7H_2O$, 0.0147 g/L $CaCl_2 \cdot 2H_2O$, 0.01 g/L thiamine HCl) supplemented with 2 g/L yeast extract, 0.02 mM cyanocobalamin, 0.1 mM IPTG, 0.05 g/L ampicillin, and 0.05 g/L streptomycin was used. Cells were cultured in 500 mL Sakaguchi flasks containing 50 mL of the M9 medium at 37°C in a shaking incubator at 120 rpm (MM-10, Taitec, Saitama, Japan).

### 3-3-4 Analytical methods

Cell growth was monitored by the measurement of $OD_{600}$ using UV-mini 1240 (Shimadzu, Kyoto, Japan). Concentrations of glycerol, 3HP, 1,3PDO, succinate, lactate, acetate, formate, and ethanol present in the supernatant of the culture broth were determined by an HPLC system (HPLC Prominence, Shimadzu) equipped with an Aminex HPX-87H column (Bio-Rad, Hercules, CA, U.S.A.), a UV/vis detector (SPD-20A), and a refractive index detector (RID-10A). The column temperature was set to 65°C, and 2 mM $H_2SO_4$ was used as the mobile phase with a flow rate of 0.5 mL/min. The flow cell temperature of the refractive index detector was set to 35°C. The supernatant of the culture broth was obtained by centrifugation at 21,500 × *g* for 5 min at 4°C, and then filtered through a Millex HV 0.45-μm filter (Merck KGaA). Methylglyoxal in the supernatant was quantified colorimetrically with 2,4-dinitrophenylhydrazine (2,4-DNPH) (125). The reaction mixture containing 67 μL of sample and 22 μL of 2,4-DNPH solution (0.1% 2,4-DNPH in 2 M HCl) was incubated for 15 min at 30°C in a 96-well microtiter plate, and then 111 μL of 10% NaOH was added. After a 15-min incubation at room temperature, the absorbance at 544 nm was measured with a microplate reader (1420 ARVO, PerkinElmer Inc., Waltham, MA, U.S.A.).

## 3-4 Results and discussion

### 3-4-1 Construction of a 3HP-producing strain in *E. coli*

The 3HP biosynthetic pathway from glycerol consists of two reactions: the dehydration of glycerol to 3-hydroxypropionaldehyde (3HPA), catalyzed by glycerol dehydratase, and the oxidation of 3HPA to 3HP, catalyzed by aldehyde dehydrogenase (89). Since *E. coli* does not possess the 3HP biosynthetic pathway, the 3HP-producing strain (TK52) was constructed by the overexpression of *dhaB* and *gdrAB*, which encode for glycerol dehydratase and glycerol dehydratase reactivase (from *K. pneumoniae*), respectively, and *aldH*, which encodes for aldehyde dehydrogenase (from *E. coli*), as described in a previous study (92). TK52 was cultivated in M9 medium in a Sakaguchi flask and 3HP was produced at 0.05 ± 0.01 C-mol/C-mol (Figure 3-1A-B and Table 3-4). Acetate was produced as a major byproduct 0.10 ± 0.02 C-mol/C-mol and small yields of 1,3PDO (0.01 ± 0.00 C-mol/C-mol) and succinate (0.01 ± 0.00 C-mol/C-mol) were also produced. Ethanol and formate were not detected in this strain. Although *E. coli* does not possess 1,3PDO biosynthetic pathways, the introduction of *dhaB* for 3HP production enabled production of 1,3PDO as follows: glycerol dehydratase converts glycerol to 3HPA, and an endogenous alcohol dehydrogenase (encoded for by *yqhD*) further converts 3HPA to 1,3PDO (126).

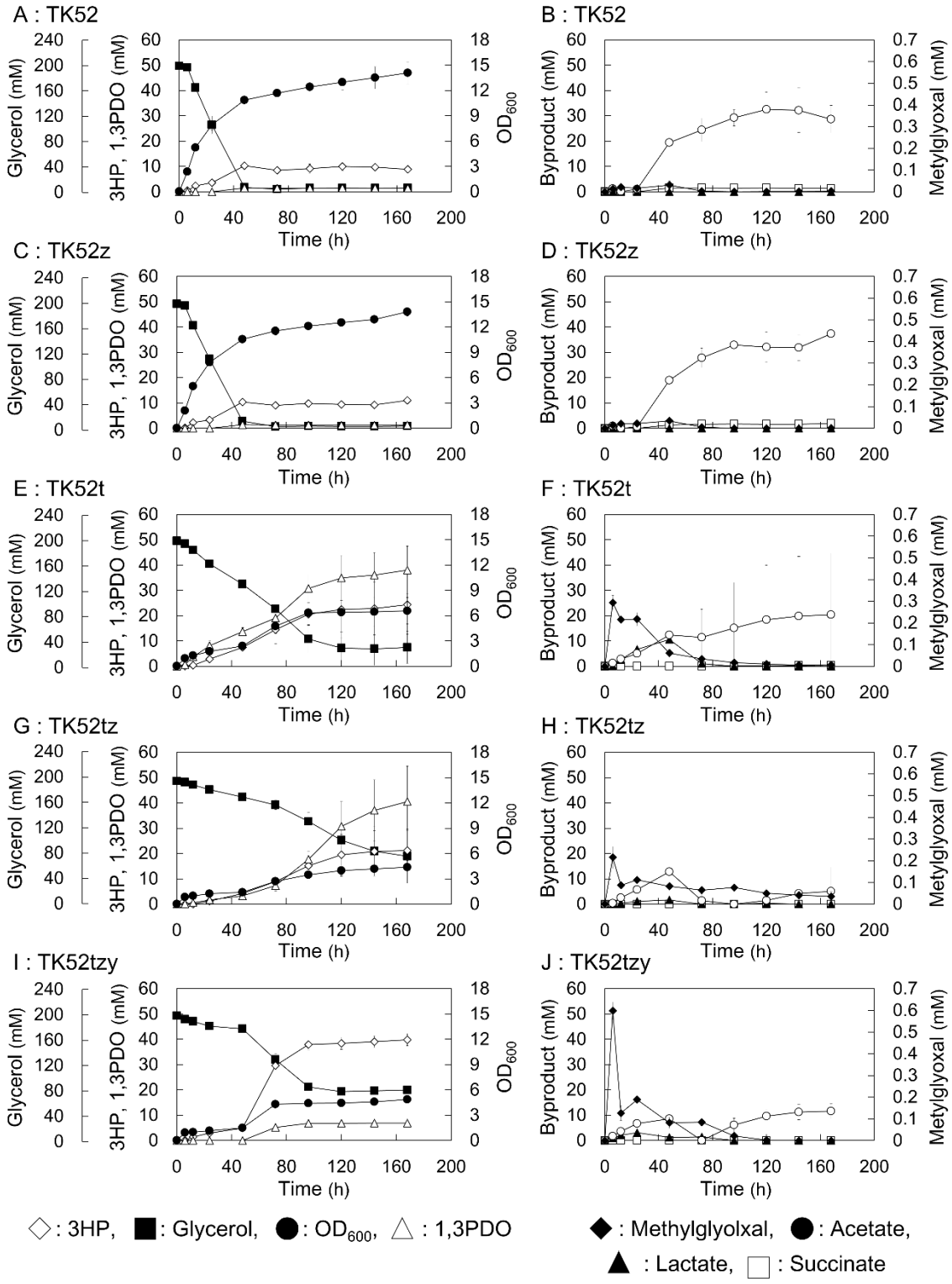



**Figure 3-1 Culture results of the 3HP-producing strains.** The culture results of the strains TK52 (A, B), TK52z (C, D), TK52t (E, F), TK52tz (G, H), and TK52tzy (I, J) are shown. The symbols indicate metabolite

concentrations and optical density as shown in the bottom. Error bars represent standard deviation of triplicate experiments in TK52 and TK52z strains, and of nine replicate experiments in other strains. Some of the error bars are smaller than the symbols.

Table 3-4 Summary of the experimental results

| | Culture result | | | | | Simulation | |
|---|---|---|---|---|---|---|---|
| | TK52 | TK52z | TK52t | TK52tz | TK52tzy | No deletion | *ΔtpiA Δzwf* |
| Specific growth rate ($h^{-1}$)*1 | 0.73 ± 0.00 | 0.71 ± 0.00 | 1st: 0.55 ± 0.01<br>2nd: 0.03 ± 0.01 | 1st: 0.54 ± 0.01<br>2nd: 0.03 ± 0.00 | 1st: 0.56 ± 0.01<br>2nd: 0.04 ± 0.01 | | |
| Maximum 3HP production rate (mmol $gCDW^{-1}hr^{-1}$)*2 | 0.08 ± 0.02 | 0.09 ± 0.01 | 0.22 ± 0.14 | 0.27 ± 0.11 | 0.94 ± 0.05 | | |
| Consumed glycerol (mM)*3 | 192.7 ± 5.0 | 193.2 ± 2.4 | 168.2 ± 24.4 | 119.0 ± 44.6 | 117.6 ± 3.9 | | |
| Biomass (C-mol/C-mol)*3 | 0.37 ± 0.04<br>(5.2 ± 0.5) | 0.36 ± 0.01<br>(5.1 ± 0.2) | 0.20 ± 0.03<br>(2.4 ± 0.6) | 0.22 ± 0.11<br>(1.6 ± 0.3) | 0.21 ± 0.01<br>(1.8 ± 0.1) | 0.48 | 0.23 |
| 3HP (C-mol/C-mol)*3 | 0.05 ± 0.01<br>(8.9 ± 1.3) | 0.06 ± 0.01<br>(11.1 ± 1.0) | 0.15 ± 0.07<br>(24.3 ± 11.7) | 0.20 ± 0.09<br>(21.2 ± 7.7) | 0.34 ± 0.01<br>(39.9 ± 2.4) | 0 | 0.71 |
| 1,3-PDO (C-mol/C-mol)*3 | 0.01 ± 0.00<br>(1.5 ± 0.8) | 0.01 ± 0.00<br>(1.3 ± 0.3) | 0.23 ± 0.05<br>(38.1± 9.4) | 0.38 ± 0.13<br>(40.5 ± 14.0) | 0.06 ± 0.01<br>(7.0 ± 0.7) | 0 | 0 |
| Succinate (C-mol/C-mol)*3 | 0.01 ± 0.00<br>(1.3 ± 0.2) | 0.01 ± 0.00<br>(1.9 ± 0.2) | 0.00 ± 0.00<br>(0.4 ± 0.3) | 0<br>(0) | 0<br>(0) | 0 | 0 |
| Acetate (C-mol/C-mol)*3 | 0.10 ± 0.02<br>(28.7 ± 5.4) | 0.13 ± 0.00<br>(37.3 ± 1.1) | 0.08 ± 0.10<br>(20.5 ± 24.3) | 0.02 ± 0.04<br>(5.1 ± 9.6) | 0.07 ± 0.02<br>(11.6 ± 3.0) | 0.27 | 0 |
| Maximum methylglyoxal (mM) | 0.03 ± 0.02 | 0.03 ± 0.02 | 0.29 ± 0.02 | 0.22 ± 0.02 | 0.60 ± 0.02 | | |

*1 Specific growth rates were calculated using $OD_{600}$ at 0–6 h for the TK52 and TK52z strains. For the TK52t, TK52tz and TK52tzy strains, specific growth rates during the 1st and 2nd growth phases were calculated using the $OD_{600}$ at 0–6 h and 48–72 h, respectively.

*2 Maximum 3HP production rates were calculated from the data at 24–48 h for the TK52 and TK52z strains, and at 48–72 h for other strains.

*3 The values in the parentheses indicate the final concentrations of biomass (g/L) and products (mM). For calculation of biomass yield, $OD_{600}$ was converted into dry cell weight using the conversion factor 0.37 g DC/L, and carbon-mol in the biomass was calculated based on the biomass composition described in the iAF1260 model (81)

### 3-4-2 Gene knockout simulation for 3HP production

Metabolic simulation was carried out to improve 3HP production by considering the whole metabolic network. A genome-scale metabolic model of *E. coli*, iAF1260 (81), which includes 2077 metabolic and transport reactions, and 1,038 unique metabolites, was employed with FBA to simulate 3HP production in *E. coli*. Since the 3HP biosynthesis pathway does not exist in *E. coli*, seven reactions involved in the 3HP biosynthesis pathway were added to the iAF1260 model (Table 2-2), which was subsequently referred to as the iAF1260-3HP model. Using the iAF1260-3HP model, multiple gene knockout simulations were performed to identify candidate genes that when deleted could enhance 3HP production under the same condition in the chapter 2.

FBA simulations predicted that no deletion *E. coli* catabolized glycerol via glycolysis, and formed acetate, ethanol, and formate at optimal growth state. This result said the 3HP production in the wildtype *E. coli* is not preferable for cell growth. Double-reaction knockout simulations identified some candidate genes that when deleted together could enhance 3HP production (Table 3-5). Among these, *ΔtpiA Δzwf*, *ΔtpiA Δpgi*,and *ΔtpiA Δedd* models displayed the highest carbon-mol yield of 3HP on glycerol (0.71 C-mol/C-mol). Because the optimal flux distributions in these knockout model were almost same and the maximum production yield of 3HP was same in those model, I focused on the double knockout of *tpiA* and *zwf* for further analysis in this study. Deletion of *tpiA* encoding triosephosphate isomerase, which converts dihydroxyacetone phosphate (DHAP) to glyceraldehyde- 3-phosphate (GAP), changed the direction of the reaction in

glycolysis from catabolism to gluconeogenesis (Figure 3-2C). Deletion of *zwf* blocked flux into the Entner-Doudoroff pathway (Figure 3-2 D), which was active in the *ΔtpiA* model. The inactivation of the Entner-Doudoroff pathway resulted in glycerol catabolism through the glycerol kinase reaction, which converts glycerol to glycerol-3-phosphate, and the methylglyoxal pathway, in which DHAP is converted to pyruvate via methylglyoxal. 3HP was produced instead of acetate in the double reaction knockout models. In the *ΔtpiA* model, acetate production was preferred since ATP was also generated. However, in the double reaction knockout models, ATP was consumed by the glycerol kinase reaction, which is why total ATP was not generated by acetate production from glycerol. When the OUR was limited, 3HP production was increased, instead of acetate production, to balance the reduced capacity of the respiratory chain due to the reduction in NADH generation by 3HP production from glycerol.

**Table 3-5 Knockout candidate gene sets for growth-coupled 3HP production**

| Knockout genes | 3HP yield (C-mol/C-mol) | Biomass yield (C-mol/C-mol) |
|---|---|---|
| – | 0 | 0.48 |
| *Δ tpiA* | 0 | 0.32 |
| *Δ tpiA Δzwf* | 0.71 | 0.23 |
| *Δ tpiA Δpgi* | 0.71 | 0.23 |
| *Δ tpiA Δedd* | 0.71 | 0.23 |

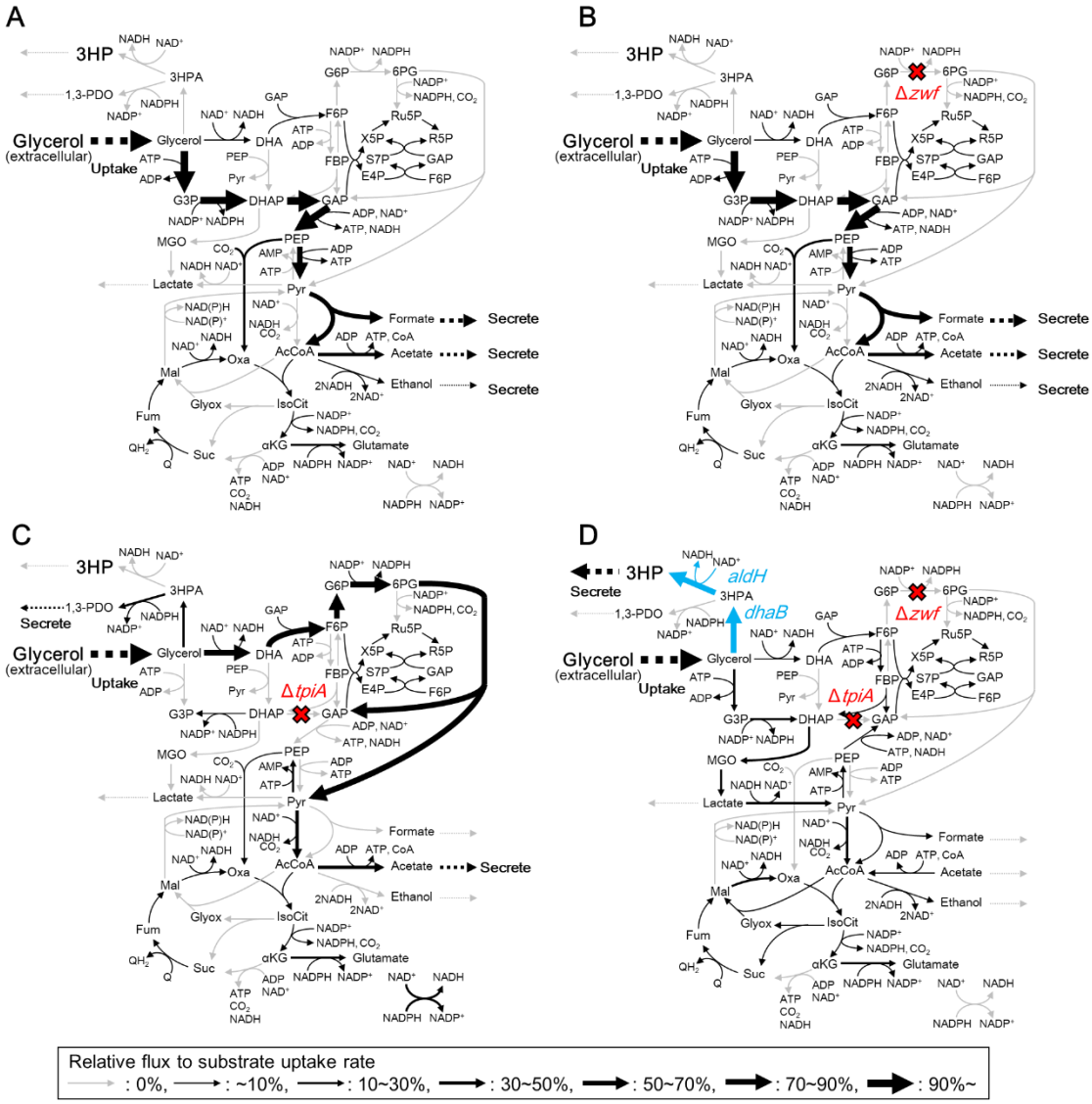



**Figure 3-2 Flux distributions for 3HP production in *E. coli*.** Flux distributions at optimal growth state of *E. coli* having 3HP synthetic pathway (A), *Δzwf* model (B), *ΔtpiA* model (C) and *ΔtpiAΔzwf* model (D). Width of the black arrow corresponds to the relative flux value of glycerol uptake rate. Gray arrows indicate the flux of the corresponding reaction was 0. Blue arrows in panel D indicate the predicted rate limiting reaction in the chapter 2. All flux distributions were calculated by FBA as same calculation condition in the chapter 2.

### 3-4-3 Evaluation of the 3HP productivity of the double knockout mutant of *tpiA* and *zwf*

Based on the results of the gene knockout simulation, both *tpiA* and *zwf* were disrupted in the TK52 strain, generating strain TK52tz *(ΔtpiA Δzwf)*. Estimated flux distributions at optimal growth state were summarized in Figure 3-2. The 3HP yield of TK52tz was successfully increased 4.4-fold relative to TK52 (0.20 ± 0.09 C-mol/C-mol) (Figure 3-1G-H and Table 1). TK52tz exhibited a two-step growth phase that was not observed in TK52, with specific growth rates of 0.54 1/h in the first growth phase (0–6 h) and 0.03 1/h in the second growth phase (48–72 h). 3HP was mainly produced in the second growth phase with the consumption of acetate and lactate that was produced prior to the second growth phase. The growth rate of TK52tz was decreased compared to TK52 (0.73 $h^{-1}$), and glycerol was not completely consumed in TK52tz. Acetate and 1,3PDO were produced as byproducts.

Based on the metabolic simulation, glycerol was predicted to be catabolized in TK52tz through the methylglyoxal pathway, which converts DHAP to pyruvate (Figure 3-2D). This pathway is not usually active in *E. coli* due to allosteric inhibition by inorganic pyrophosphate and the low activity of enzymes involved in this pathway (127)(128). The extracellular concentration of methylglyoxal, an intermediate of the pathway, was increased significantly in TK52tz (0.22 mM at maximum) compared with TK52 (0.03 mM at maximum). This suggested that flux into the methylglyoxal pathway was increased as predicted. This could result in decreased growth and incomplete glycerol consumption because

methylglyoxal is a toxic cellular electrophile that reacts with the nucleophilic centers of macromolecules such as DNA, RNA, and protein (129).

Although the 3HP yield of TK52tz was improved to 0.20 ± 0.09 C-mol/C-mol by the knockout, the yield was lower than of the predicted result of 0.71 C-mol/C-mol. In the chapter 2, two reactions in biosynthesis of 3HP from glycerol were predicted as possible-rate limiting reactions. Since glycerol dehydratase converted a large amount of glycerol to 3-hydroxypropionaldehyde at 0.58 ± 0.18 C-mol/C-mol, glycerol dehydratase was thought to be not rate-limiting reaction (Figure 3-3). On the other hand, larger amount of 3-hydroxypropionaldehyde was converted to 1,3PDO than 3HP. These result indicated that 3-hydroxypropionaldehyde dehydrogenase should be a rate-limiting reaction. Since TK52tz strongly induced the expression of these genes using the T7 RNA polymerase/promoter system, the inconsistency might be caused by low turnover rate of the enzymes and the other mechanism. Experimental evaluation revealed that 1,3PDO was main byproduct in TK52tz. The overflow of 1,3PDO synthesis should decrease the availability of 3-hydroxypropionaldehyde as a substrate for 3-hydroxypropionaldehyde dehydrogenase and limited the flux to 3HP.

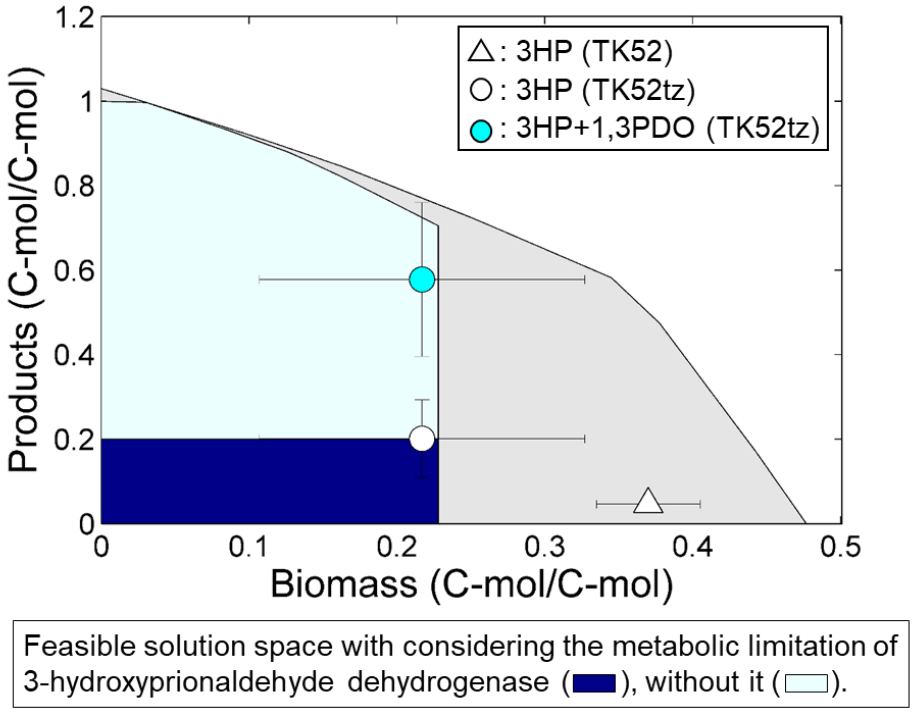



**Figure 3-3 Comparison of feasible solution spaces and culture results.** Computationally calculated feasible solution space in metabolic network in the model of iAF1260-3HP (gray), and iAF1260-3HP *Δtpi*A *Δzwf* with or without considering the metabolic limitation of 3-hydroxypropionaldehyde dehydrogenase (light blue or dark blue, respectively). For calculating the feasible solution space with the metabolic limitation, the upper bounds of 3-hydroxypropionaldehyde dehydrogenase flux was set to free or 20% of GUR. White triangle and circle indicate production yields of 3HP in TK52 and TK52tz, respectively. Blue circle indicates sum of production yield of 3HP and 1,3PDO in TK52tz.

### 3-4-4 Evaluation of the effect of gene deletion of *tpiA* and *zwf* on 3HP productivity

The single knockout strains of *tpiA* and *zwf*, TK52t (*ΔtpiA*) and TK52z (*Δzwf*), respectively, were constructed to analyze the effects of the knockout of each gene. Metabolic simulation predicted that the *zwf* knockout would not affect metabolism, but the *tpiA* knockout would alter metabolic flux into the Entner-Doudoroff pathway and decrease the growth rate. Similar culture results between TK52z and TK52 (Figure 3-1A–D, Table 3-4) suggest that the flux into the oxidative pentose phosphate pathway was small, and the deletion of *zwf* had a small impact on metabolism in this condition, as predicted by the simulation. The culture results of TK52tand TK52tz (*ΔtpiA Δzwf*) were similar (Figure 3-1E-H, Table 3-4). The 3HP yield of TK52t was increased by 3.2-fold relative to TK52 (0.15 ± 0.07 C-mol/C-mol) and slightly reduced when compared to TK52tz. TK52t produced 1,3PDO (0.23 ± 0.05 C-mol/C-mol) as a byproduct, as predicted by metabolic simulation. Acetate production by the *tpiA* knockout was also predicted, however measurements of acetate production (0.08 ± 0.10 C-mol/C-mol) in TK52t contained significant variation, thus it was difficult to compare the experimental results with the results from the metabolic simulation.

The increased 3HP yield in the single *tpiA* knockoutstrain, which was not predicted by metabolic simulation, might be due to the conversion of glycerol to glycerol 3-phosphate by glycerol kinase, which may have had a higher activity than the glycerol dehydrogenase that converts glycerol to DHA (130). This would lead to increased flux into the methylglyoxal pathway, as

indicated by elevated methylglyoxal levels (Figure 3-1 F), resulting in similar culture results for TK52tz (Figure 3-1H), such as 3HP production.

### 3-4-5 Proposed molecular mechanism of overflow of 1,3PDO

Wildtype *E. coli* does not have 1,3PDO synthetic pathway. When heterologously expressing glycerol dehydratase in *E. coli*, the engineered *E. col* can produce 1,3PDO because there are some enzymes catalyzing 3-hydroxypropionaldehyde to 1,3PDO as aldehyde reductase. YqhD is known as NADPH-dependent aldehyde reductase and the enzyme has reductase activity for a broad range of short-chain aldehydes including 3-hydroxypripionaldehyde (131). Specific activity of *E. coli* YqhD as 3-hydroxypropionaldehyde reductase is strong and several studies used YqhD for production 1,3PDO from glycerol (126)(132). Although the expression of *yqhD* in TK52tz was not directly induced by genetic manipulation, the knockout mutant produced large amount of 1,3PDO at yield of 0.38 ± 0.13 C-mol/C-mol. Previously, Ozyamak et al. reported that *yqhD* was strongly induced respond to methylglyoxal for detoxification of the toxic electrophile, which attacks macro molecules such as DNA (133). Since the knockout mutant highly accumulated methylglyoxal of 0.22 mM compared with no deletion strain, the expression of *yqhD* should be induced by the higher accumulation of methylglyoxal (Figure 3-4). This hypothesis was consistency in the lower production of 1,3PDO in TK52 and TK52z.

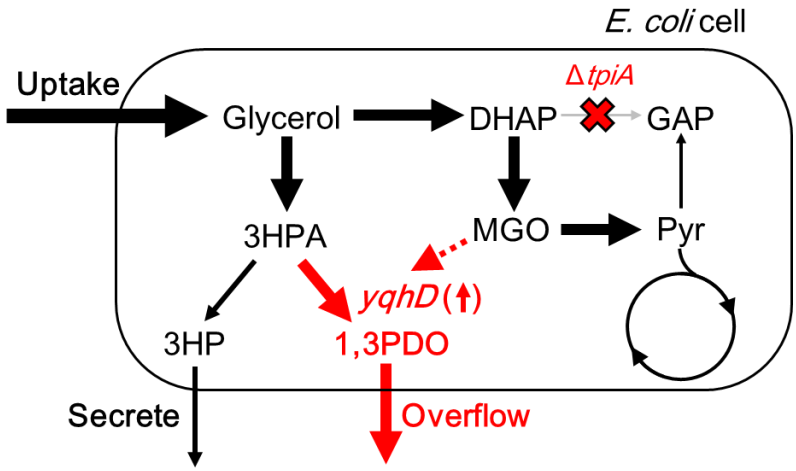



Figure 3-4 Proposed mechanism of 1,3PDO overflow in TK52tz.

### 3-4-6 Further increase in 3HP production by yqhD deletion

Based on the proposed mechanism of the rate limiting reaction, *yqhD* was deleted in TK52tz, generating the strain TK52tzy, in order to decrease 1,3PDO production and further increase 3HP production. As a result, the 3HP yield of TK52tzy was increased 1.7-fold (0.34 ± 0.01 C-mol/C-mol) and the 1,3PDO yield was drastically reduced (0.06 ± 0.01 C-mol/C-mol) relative to TK52tz (Figure 3-5 and table 3-4). Compared to the parental strain, a 7.4-fold increase in 3HP yield was achieved in TK52tzy. Despite the deletion of *yqhD*, 1,3PDO was still produced in TK52tzy (0.06 ± 0.01 C-mol/C-mol), likely due to the presence of other alcohol dehydrogenases that might convert 3HPA to 1,3PDO. The deletion of *yqhD* increased the maximum concentration of extracellular methylglyoxal in TK52tzy (0.60 mM at maximum) relative to TK52tz (0.22 mM at maximum), since YqhD also utilizes methylglyoxal as a substrate (131). The other culture results of TK52tzy, such as consumed glycerol and biomass yield, were similar to those of TK52tz (Table 3-4). The culture results of TK52t and TK52tz revealed large variations in the production of 3HP, 1,3PDO, and acetate and the consumption of glycerol (Figure 3-1E-H). On the other hand, the results from TK52tzy displayed smaller variations in these measurements (Figure 3-1I-J). Furthermore, TK52t and TK52tz produced a large amount of 1,3PDO, which accompanies NADPH oxidation. The deletion of *yqhD* in TK52tz decreased the magnitude of the error, suggesting that the high production of 1,3PDO in TK52t and TK52tz might cause redox imbalance, resulting in the large variations in the measurements.

The 3HP yield of TK52tzy (0.34 ± 0.01 C-mol/C-mol) was comparable to previous studies producing 3HP from glycerol via flask cultivation. Mohan et al. achieved 3HP yield of 0.39 ± 0.00 C-mol/C-mol by the optimization of culture conditions such as the initial culture medium pH (117). Rathnasingh et al. achieved a yield of 0.40 C-mol/C-mol by the expression of α-ketoglutaric semialdehyde dehydrogenase instead of aldehyde dehydrogenase, and periodic supplementation with vitamin $B_{12}$, a coenzyme for glycerol dehydratase (92). Jung et al. constructed an engineered *E. coli* strain by knocking out *ackA, pta*, and *yqhD* to reduce byproduct generation and knocking out *glpR* and overexpressing *glpF* to enhance glycerol metabolism (95). They achieved high 3HP production (42 g/L) by fed-batch cultivation using a jar fermenter, but the 3HP yield was lower (0.26 C-mol/C-mol) than that achieved in this study. In previous studies (92)(117), acetate was a major byproduct, as it was in this study, and higher yields of succinate, lactate, and ethanol were also produced. Succinate, lactate, and ethanol production might serve to oxidize the excess NADH that accompanies 3HP production and glycerol catabolism via glycolysis, since production of these metabolites requires NADH as a reducing agent. Conversely, the yields of these metabolites were small in TK52tzy. This might be because the deletion of *tpiA* and *zwf* prevented flux into glycolysis, reducing excess NADH production, as predicted by metabolic simulation.

In the chapter 2, there are two possible rate limiting reactions for 3HP production from glycerol in the double knockout mutant of *tpiA* and *zwf*. Since the total carbon yield of 3HP and 1,3PDO reached to 0.58 C-mol/C-mol

in the knockout mutant, the 1st step of glycerol dehydratase was thought to be sufficient to catalyze the reaction (Figure 3-3). On the other hand, the total carbon yield of 3HP and 1,3PDO decreased to 0.40 C-mol/C-mol by additional knockout of *yqhD*, and the production yield of 3HP in the knockout mutant was 0.34 ± 0.01 C-mol/C-mol lower than of the optimal solution (0.71 C-mol/C-mol, Figure 3-6). These results suggested that 3-hydroxypropionaldehyde dehydrogenase was still rate-limiting reaction and had not enough activity to converting 3-hydroxypropionaldehyde to 3HP as predicted. Changing the aldehyde dehydrogenase to the superior enzyme such as α-ketoglutaric semialdehyde dehydrogenase (92) is a possible way to increase the 3HP production. Furthermore, some amount of 1,3PDO was still secreted in the triple knockout mutant. This result said that the other aldehyde reductase catalyzed 3-hydroxypropionaldehyde to 1,3PDO. Complete disruption of 1,3PDO synthesis should improve the supply of the substrate for 3-hydroxypropionaldehyde dehydrogenase and increase the 3HP production.

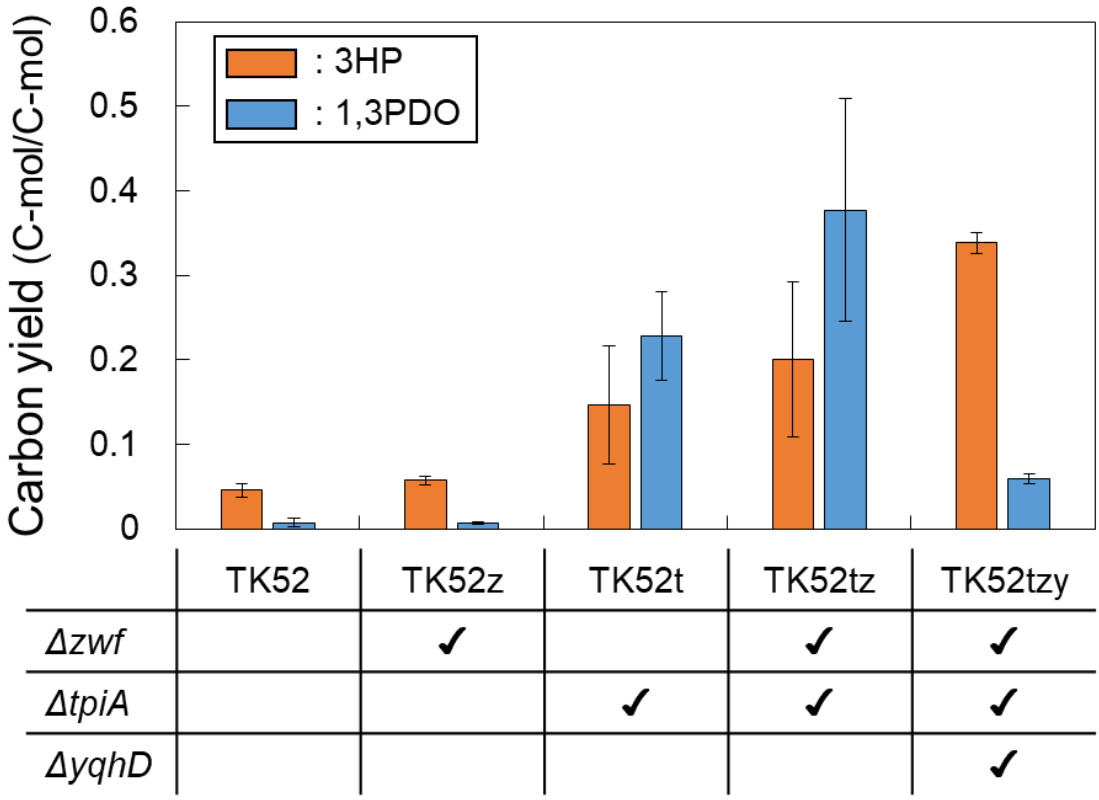



Figure 3-5 Summary of carbon yields of 3HP and 1,3PDO in the engineered strains.

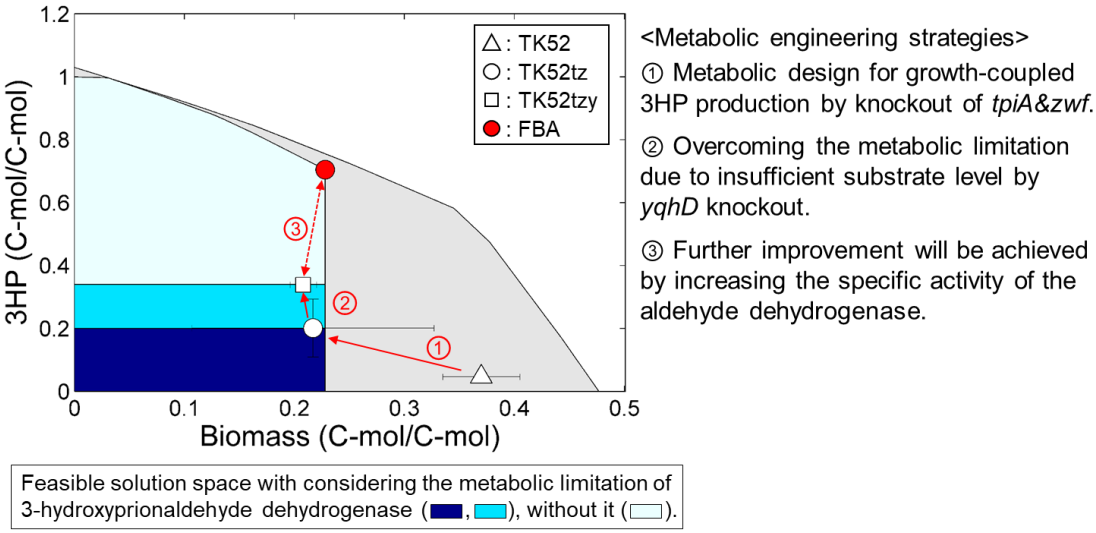



**Figure 3-6 Comparison of feasible solution spaces and culture results.** Computationally calculated feasible solution space in metabolic network in the model of iAF1260-3HP (gray), and iAF1260-3HP *ΔtpiA Δzwf* with or without considering the metabolic limitation of 3-hydroxypropionaldehyde dehydrogenase (light blue or dark blue, respectively). For calculating the feasible solution space with the metabolic limitation, the upper bounds of 3-hydroxypropionaldehyde dehydrogenase flux was set to free, 20% or 34% of GUR. White plots indicate production yields of 3HP in TK52 (triangle), TK52tz (circle) and TK52tzy (square), respectively. Red circle indicate the optimal solution of iAF1260-3HP calculated by FBA.

## 3-5 Summary

In this chapter, the effect of the predicted knockout genes on 3HP producing *E. coli* was experimentally evaluated for understanding the molecular mechanism causing the rate limiting reaction predicted in the chapter 2. Although the double knockout mutant of *tpiA* and *zwf* increased 3HP production (0.20 ± 0.09 C-mol/C-mol), which was 4.4-fold higher than of no deletion strain, a large amount of 1,3PDO (0.38 ± 0.13 C-mol/C-mol) was secreted as an unexpected result. The molecular mechanism causing the overflow of 1,3PDO formation was estimated by that activated methylglyoxal pathway should induce the expression of *yqhD* converting 3-hydroxypropanediol to 1,3PDO. Additional knockout of *yqhD* decreased 1,3PDO production to 0.06 ± 0.01 and increased 3HP production to 0.34 C-mol/C-mol by 1.7-fold.

Metabolic simulation is a powerful tool for the design of metabolic engineering strategies to improve target production and has been used successfully in many studies. Metabolic simulation using FBA is simply based on the assumption of steady state metabolism without considering the complex cellular mechanisms such as enzyme activity and regulation. Therefore, FBA cannot stop overestimating the productivity in the actual cell and causing discrepancies between the simulation and experimental results, i.e., the overflow of 1,3PDO in the present study. The *in silico* screening method developed in the chapter 2 is useful to readily screen of rate-limiting reactions causing the discrepancies. Furthermore, experimental evaluation

helps us to understand what cause the rate-limiting reactions and build next metabolic engineering strategies.

# Chapter 4: Strain improvement for succinic acid production from glycerol in engineered *E. coli* by adaptive laboratory evolution

## 4-1 Highlights

- Microbial production of succinic acid has been attracted recently, due to its wide-range availability for commercial products.
- Adaptive laboratory evolution of the knockout mutant successfully improved succinate production at 0.45 C-mol/C-mol, which was predicted value by FBA.
- All evolved strains had novel mutations in *ppc* for desensitizing Asp inhibition.
- ALE expanded the flux space of actual metabolic network of the knockout mutant.

## 4-2 Introduction

### 4-2-1 Microbial production of succinic acid

Succinic acid is widely used for balk chemical industry as a precursor for various chemicals such as 1,4-butanediol, adipic acid and polybutylene-succinate, and it has been traditionally supplied by petrochemical production. Since depletion of oil resources has been concerned recently, microbial production of succinate from biomass is alternative way to compensate the demand. The U.S. Depertment of Energy designated succinate as one of the top-value added chemicals produced by biomass like 3HP in the chapter 3 (3)(4).

The microbial production of succinate has been developed using natural succinate producers including *Mannheimia succiniciproducens* (18) and recombinant microorganisms such as *E. coli* (57)(134)(135), *C. glutamicum* (136) and *Yarrowia lipolytica* (137). Some of studies focused on fermentation of glycerol to produce succinate, since glycerol is inexpensive carbon source supplied as a byproduct of biodiesel production (109)(110)(111). The difference between carbon sources is that the succinate production from glycerol does not involve the net generation of reducing equivalents, while the fermentations of glucose or xylose to succinate is redox-generating pathway (130). Under the oxygen limited condition, wildtype *E. coli* favors to produce fermentative products generated from redox-balanced pathway. Since the ethanol production also does not generate net reducing equivalents and can generates ATP needed for cell growth, wildtype *E. coli* mainly converted glycerol to ethanol under the micro-aerobic fermentation of

glycerol. Succinic acid production from glycerol should be enhanced by a genetic manipulation for rewiring carbon flow into the biosynthetic pathway with considering redox balance and energy generation.

### 4-2-2 Genome-scale metabolic design for succinic acid production

In the chapter 2, I identified the possible rate-limiting reactions for enhancing growth-coupled succinate production by developing the novel screening method using FBA. Although overexpression of genes involved in the rate-limiting reaction is a way to overcome the metabolic limitations, the approach is not always enough to improve the enzymatic activity as expected since several factors involve the rate-limiting reaction. For example, in case of 3HP production discussed in chapter 3, the overflow of 1,3PDO in the knockout mutant was caused by activation of methylglyoxal pathway in spite of overexpression of genes involved in biosynthetic pathway of 3HP. Furthermore, the number of candidates of rate limiting reactions for succinate production was higher than of the other cases. The various factors seems to be involved in the regulation of the predicted reactions. For example, Ppc, which was one of the 9 possible rate limiting reactions, is transcriptionally regulated by Cra, and allosterically activated by acetyl-CoA, fructose 1,6-bisphosphate and guanosine triphosphate, and also inhibited by malate and aspartate (100)(101). It is still unclear that which factor mostly involved in the complex regulation *in vivo*. Even if we know the detail metabolic system, it should be difficult to release the limitation by genetic engineering due to insufficient knowledges for changing enzymatic

properties. In such case, evolution engineering is a possible approach to optimize the enzymatic functions and strains with spontaneous mutations.

### 4-2-3 Adaptive laboratory evolution

ALE is a useful experimental technique to increase cell growth rates against environmental perturbations including genetic perturbation (138)(139) and designated culture condition (140)(141)(142). In the field of metabolic engineering, ALE has been applied for latent pathway activation (143)(144) and enhanced tolerant to the target compounds (62)(64). As shown in these studies, ALE can enforce the cell growth rate for several perturbations, and the endpoint strain should maximize its growth ability for them. Since FBA calculates the metabolic state at the optimal growth state, the endpoint strain should reach to the predicted state by FBA (145)(146)(52). Previously, Ibarra et al. reported that the growth rate of wildtype *E. coli* on glycerol as the sole carbon source reached to the optimal growth state predicted by FBA (145). Fong & Palsson reported that strains of *Escherichia coli* lacking a single metabolic gene increase their growth rates (by 87% on average) during ALE, and that the endpoint growth rates were consistency with the predicted value in 39 of 50 (78%) strains tested (146). In the other study, Fong et al. reported that the knockout *E. coli* mutants designed for growth coupled lactate production successfully increased lactate production as predicted by FBA (52).

The other use of ALE is to investigate microbial adaptation process. During ALE, the cells enhance their growth rates with naturally occurring

mutations. Genome resequencing analysis of evolved strains reveals the mutations, and re-introducing them into parent's genome DNA by genome engineering can identify which mutations mostly contribute to the adaptation (141)(62)(64). Further analysis of the key mutation at function level leads to understand the molecular mechanism (147). Therefore, ALE of the knockout mutant for growth coupled succinate production designed in the chapter 2 should result in not only increasing the succinate production, but also identifying the key mutations for increased the succinate production and understanding the metabolic system causing the metabolic limitation.

### 4-2-4 Objective of this chapter

In the chapter 2, I identified the possible rate-limiting reactions for enhancing growth-coupled succinate production by developing the novel screening method using FBA. There are a lot of ways to optimize the metabolic network, since the number of predicted rate limiting reactions were larger than of other target production process, and many factors involved in the regulation of the predicted enzymes. One of the advantages of metabolic design of growth-coupled target production is that high target producing strains should be isolated by select faster growing strain via ALE, since the strain have to produce target metabolite in order to grow faster. Therefore, ALE of the knockout mutant based on FBA prediction is expected to overcome the metabolic limitation and improve the target production with increasing cell growth rate. The current chapter aimed to optimize the

metabolic network for enhancing succinate production by ALE and evaluated the rate-limiting reaction in actual cells.

First, evolved strains were obtained from parallel ALE experiment and the succinate productivities were evaluated in batch culture and continuous culture. Next, genome resequencing analysis and reverse engineering using multiplexed genome engineering (MAGE) identified the key mutations to increase the succinate production. Finally, functional analysis and multi-omics analysis revealed the molecular mechanism causing and overcoming the metabolic limitation.

## 4-3 Materials and methods

### 4-3-1 *In silico* screening of knockout gene targets for enhancing 3HP production

FBA is a method to estimate a metabolic flux distribution using a genome-scale metabolic model (44)(45)(46)(47). To identify the candidates for gene knockout to achieve growth-coupling production of succinate from glycerol, FBA was performed by using the iAF1260 models (81) as shown below:

$$\begin{aligned} &\text{max} && \boldsymbol{c}^T \cdot \boldsymbol{v} \\ &\text{subject to} && \sum_{j \in \boldsymbol{R}} S_{i,j} \cdot v_j = 0 \quad (\forall \boldsymbol{i} \in \boldsymbol{M}) \\ & && v_{substrate\_uptake} = \text{SUR} \\ & && v_{oxygen\_uptake} = \text{OUR} \\ & && v_{knockout\_reactions} = 0 \end{aligned}$$

where $\boldsymbol{M}$ and $\boldsymbol{R}$ are the set of metabolites and reactions, respectively. $\boldsymbol{c}$ is a vector that represents coefficients of an objective function to be maximized or minimized.

Glycerol was used as the sole carbon source in metabolic simulations, and glycerol uptake rate (GUR, $v_{glycerol\_uptake}$) was set to 15 mmol $gCDW^{-1}hr^{-1}$. The oxygen uptake rate (OUR, $v_{oxygen\_uptake}$) was set to 10 mmol $gCDW^{-1}hr^{-1}$. Other external metabolites such as $CO_2$ and $NH_3$ were allowed to transport freely through the cell membrane. For determining the maximal succinate production flux at optimal growth, the objective coefficients were set to 1 and

$1\times10^{-6}$ for biomass reaction and succinate exchange reaction, respectively. After the FBA calculations for multiple reaction knockout models, the models showing succinate production at the optimal growth solution were identified as growth-coupling target model. The calculation was implemented in Matlab (MathWorks Inc., Natick, MA, USA) with a solver for linear programming, Gurobi (http://www.gurobi.com).

### 4-3-2 Strains and plasmids

All strains and plasmids used in this study are listed in Table 4-2. BW25113 was used as the succinate production host. BW25113 *ΔadhE ΔpykAF ΔgldA::kan ΔpflB::tet* was constructed by Wanner's method (25) and P1 phage transduction (123). *E. coli* strain DH5α was used for plasmid construction and *E. coli* strain BL21(DE3) was used for enzymatic analysis. Ppc was expressed in pET28-a(+) (Novagen, Madison, Wis., USA).

**Table 4-2 Strains and plasmids used in this study**

| Strains and plasmids | Source |
|---|---|
| Strains | |
| BW25113 | Previous work (25) |
| BW25113*ΔadhE*::*kan* (JW1228) | Previous work (26) |
| BW25113*ΔgldA*::*kan* (JW5556) | Previous work (26) |
| BW25113*ΔpykA*::*tet* | This work |
| BW25113*ΔpykF*::*cat* | This work |
| BW25113*ΔpflB*::*tet* | This work |
| BW25113*ΔadhE*::*kanΔpykA*::*tet* | This work |
| BW25113*ΔadhE*::*kanΔpykA*::*tetΔpykF*::*cat* | This work |
| BW25113*ΔadhEΔpykAF* | This work |
| BW25113*ΔadhEΔpykAFΔgldA*::*kan* | This work |
| BW25113*ΔadhEΔpykAFΔgldA*::*kanΔpflB*::*tet* | This work |
| DH5α | Toyobo |
| BL21(DE3) | Novagen |
| Plasmids | |
| pKD46 | Previous work (25) |
| pCP20 | Previous work (25) |
| pKD13Cm | Previous work (124) |
| pKD13Tet | Previous work (124) |
| pORTMAGE2 | Previous work (148) |
| pET28-a(+) | Novagen |
| pET28-a(+) *NdeI-ppc$^{wild}$-XhoI* | This work |
| pET28-a(+) *NdeI-ppc$^{I829S}$-XhoI* | This work |
| pET28-a(+) *NdeI-ppc$^{R849S}$-XhoI* | This work |

### 4-3-3 Adaptive laboratory evolution

ALE was performed in 5 mL of M9 minimal medium (16.9 g/L $Na_2HPO_4 \cdot 12H_2O$; 3.1 g/L $KH_2PO_4$; 1.0 g/L $NH_4Cl$; 0.5 g/L NaCl; 0.493 g/L $MgSO_4 \cdot$

$7H_2O$; 0.0147 g/L $CaCl_2 \cdot H_2O$) supplemented with 10 g/L glycerol, 1% (v/v) wolfe's mineral and 1% (v/v) wolfe's vitamin at 37 °C with shaking at 20 strokes $min^{-1}$ using Bio-Photorecorder TVS062CA (Advantec Toyo Co., Tokyo, Japan). Optical density at 660 nm was automatically monitored at every 20 min during the cultivation to determine the maximum specific growth rate. From the stationary phase culture, 1 µL of broth was passaged into fresh medium. After approximately 100 generations, evolved strains, named strain A-E, were isolated from the culture broth of five parallel passaged cultivations.

### 4-3-4 Phenotype assessment

#### *Batch culture protocol*

The evolved *E. coli* strains (strain A-E) and parent strain (strain P, BW25113 *ΔadhE ΔpykAF ΔgldA::kan ΔpflB::tet*), wildtype *E. coli* (strain W, BW25113) were inoculated from the glycerol stock to 5 mL M9 medium and cultured aerobically at 37 °C overnight. The overnight cultures were transferred to fresh 50 mL M9 medium maintaining an initial optical density of 0.05 at 600 nm ($OD_{600}$). Cells were cultured in a 100 mL Erlenmeyer flask at 37 °C and incubated in a rotary shaking incubator at 150 rpm (BR-43FL, Taitec, Saitama, Japan).

#### *Continuous culture protocol*

Strain B and P were inoculated from the glycerol stock to 40 mL M9 medium and aerobically cultured at 37 °C overnight. The culture broths were

transferred to fresh 600 mL M9 medium in a 1-L jar fermenter, BMJ-P type bioreactor (ABLE, Tokyo, Japan) equipped with temperature, pH, and dissolved oxygen sensor. The temperature was maintained at 37 °C and pH was set at 7.0 using $NH_3$ solution. Air flow rate and agitation speed were set at 300 mL/min and 400 rpm, respectively. The inoculum size was set to an initial $OD_{600}$ of 0.1. After 8 hours of cultivation in batch mode, the continuous culture mode was started with a dilution rate of 0.04 $h^{-1}$. After 48 hours into the continuous mode, the agitation speed was reduced to 100 rpm, and the culture broth were collected after more than two residence times for quantitation of extracellular metabolites.

*Fed-batch culture protocol*

Strain B was inoculated from the glycerol stock to 40 mL M9 medium and aerobically cultured at 37 °C overnight. The culture broths were transferred to fresh 400 mL M9 medium in a 1-L jar fermenter, BMJ-P type bioreactor (ABLE, Tokyo, Japan) equipped with temperature, pH, and dissolved oxygen sensor. The temperature was maintained at 37 °C and pH was set at 7.0 using $NH_3$ solution. Air flow rate and agitation speed were set at 400 mL/min and 400 rpm, respectively. The inoculum size was set to an initial $OD_{600}$ of 0.1. Feed of glycerol and $NaHCO_3$ started after 6hr of main cultivation. The feeding volume was set to a constant value of 0.05 mL/min of 600 g/L glycerol and 10 g/L $NaHCO_3$. During fed-batch culture mode, glycerol was always remaining higher than of 50 g/L.

### 4-3-5 Analytical methods

Cell growth was monitored by the measurement of $OD_{600}$ using UV-mini 1240 (Shimadzu, Kyoto, Japan). Concentrations of glycerol, succinate, lactate, acetate, formate, and ethanol present in the culture supernatant were determined by high performance liquid chromatography (HPLC Prominence, Shimadzu) equipped with an Aminex HPX-87H column (Bio-Rad, Hercules, CA, U.S.A.), a UV/vis detector (SPD-20A), and a refractive index detector (RID-10A). The column temperature was set to 65 °C, and 1.5 mM $H_2SO_4$ was used as the mobile phase at a flow rate of 0.6 mL/min. The flow cell temperature of the refractive index detector was set to 40 °C. The culture supernatant was obtained by centrifugation of the broth at 21,500 × *g* for 5 min at 4 °C, followed by filtration through a Millex HV 0.45 μm filter unit (Millipore, Bedford, MA, USA).

### 4-3-6 Genome sequence analysis and mutation analysis

For genomic DNA preparation, the evolved *E. coli* strains (A ~ E) and parent strain P, wildtype *E. coli* were inoculated from the glycerol stock to 5 mL M9 medium and aerobically cultured at 37 °C overnight. Subsequently, rifampicin (300 μg/mL final concentration) was added and the culture was continued for further 3 hours to block the initiation of DNA replication. The cells were collected by centrifugation at 20,000 × *g* for 1 min, at 4 °C, and the pelleted cells were stored at -80 °C until genomic DNA preparation. The genomic DNA was extracted from cell pellets of strain P and evolved strains A~E by DNeasy Blood & Tissue Kit (Qiagen, Germany) in accordance with

the manufacturer's instructions. Genome sequence analysis were performed with MiSeq Desktop Sequencer (Illumina, Inc., San Diego, CA, US) as described previously (64). The reads obtained were aligned to the reference sequence of *E. coli* BW25113 genomic DNA (GenBank: NZ_CP009273) using breseq pipeline, version 0.28 (149). The mutations of *ppc* identified in all evolved strains were confirmed by Sanger sequence analysis of PCR products by Value Read sequencing service (Eurofins Genomics, Tokyo, Japan). Genome sequence data of *E. coli* strains in this study were deposited in the DDBJ Sequence Read Archive of the DNA Data Bank of Japan (DRA) under accession number DRA006046.

### 4-3-7 Reverse engineering of identified *ppc* mutations in strain P

Mutations identified in *ppc* were introduced into the genome of strain P by MAGE using pORTMAGE-2 vector (148). pORTMAGE-2 was a gift from Csaba Pál (Addgene plasmid # 72677). Sequences of MAGE oligos were created by MODEST tool (150) as shown in Supplementary Table 3 and the oligos were synthesized by FASMAC Co., Ltd. (Kanagawa, Japan). We have inserted the mutations of *ppc* in strain P to make its genotype the same as that of the evolved strain B or C following the procedure reported previously (148). The mutations were confirmed by MASC-PCR using KAPA2G Fast PCR kit (Nippon Genetics, Tokyo, Japan) followed by Sanger sequence analysis of PCR products by Value Read sequencing service (Eurofins).

### 4-3-8 Molecular cloning and enzyme overexpression

The *ppc* genes were amplified by PCR using KOD FX Neo (Toyobo Co., Ltd., Osaka, Japan) from the genomic DNA of strain P, B and C using the primers as shown in Supplementary Table 3. Each amplified fragment was treated with A-attachment mix (Toyobo Co., Ltd.), and then cloned into pGEM-T Easy (Promega Co., Madison, WI, U.S.A.), followed by sequence confirmation by Value Read sequencing service (Eurofins). The *Nde*I-*Xho*I fragments of pGEM-T Easy were cloned into the same restriction site of pET28-a(+) vector, generating pET28-a(+)/$ppc^{wild}$, pET28-a(+)/$ppc^{I829S}$ and pET28-a(+)/$ppc^{R849S}$. The recombinant strains of BL21(DE3) harboring the plasmids were cultured in LB medium at 37 °C overnight. Five hundred micro liters of the culture broths were transferred to 250 mL baffled Erlenmeyer flasks containing fresh 50 mL LB medium containing 0.1 mM isopropyl-ß-d-thiogalactopyranoside and cultured at 30 °C in a rotary shaking incubator at 200 rpm (BR-43FL, Taitec). After 20 hours of cultivation, the cells pellets were obtained from 10 mL of culture broth by centrifugation at 2,500 × *g* for 5 min at 4 °C.

### 4-3-9 Protein structural analysis

The protein structure of Ppc from *E. coli* (151) (PDB ID: 1FIY) was obtained from the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do). Structure visualization and introduction of identified mutations were carried out using Pymol (http://www.pymol.org).

### 4-3-10 Enzymatic assay of Ppc with L-aspartate

The enzyme activity of the purified Ppc was assayed by a coupling reaction catalyzed by malate dehydrogenase at 30 °C as previously described (152). The standard reaction mixture contained 100 mM Tris-HCl (pH 7.5), 10 mM $MnSO_4$, 10 mM $NaHCO_3$, 2 mM PEP, 0.1 mM NADH, and 1.5 U of malate dehydrogenase (Sigma-Aldrich Co, St. Louis, MI, USA). The decrease in absorbance of NADH at 340 nm was monitored by a microtiter plate reader (Varioskan Flash Multimode Reader, Thermo Fisher Scientific, USA).

### 4-3-11 Metabolome analysis

Twenty milliliters of culture broth from the batch culture experiment was sampled rapidly during the exponential growth phase and filtered through a 0.5-μm pore size filter (PTFE-type membrane, Advantec, Tokyo, Japan). Intracellular metabolites were extracted with methanol/chloroform/water (153) and dried using a SpeedVac, SPD1010 (Thermo Fisher Scientific, Waltham, MA, USA) at room temperature.

Gas chromatography–mass spectrometry (GCMS-QP2010 Ultra, Shimadzu, Japan) was performed under the following conditions. The column used was DB-5MS+DG column (30 m × 0.25 mm ID × 0.25 μm, Agilent Technologies, Santa Clara, USA). The front inlet temperature was set to 250 °C. The helium gas flow rate through the column was at 1 mL/min. The column temperature was held at 70 °C for 2 min isothermally and later raised by 3 °C/min to 280 °C and held there for 3 min isothermally. The

transfer line and ion-source temperatures were 250 °C and 200 °C, respectively. The dried extract was dissolved in 25 µL methoxyamine hydrochloride (20 mg/mL-pyridine) and incubated at 37 °C for 90 min. For *tert*-butyldimethylsilylation, 25 µL of *N*-(*tert*-butyldimethylsilyl)-*N*-methyltrifluoroacetamide containing 1% *tert*-butyldimethylchlorosilane was added and incubated at 60 °C for 30 min. After 2 h cooling, the samples were centrifuged at 21,500 × *g* for 5 min. One micro liters of the supernatant containing the derivatized samples were injected at a split injection ratio of 1:10.

### 4-3-12 Flux balance analysis with constraints of measured fluxes.

Function of optimizeCbModel in COBRA Toolbox (154) was used to calculate metabolic flux distributions on continuous cultivation. For considering glycerol assimilation pathway, the reactions of EX_glyc(e), GLYCtex, GLYK, G3PD5, GLYCDx, DHAPT and F6PA in the genome-scale metabolic model iAF1260 were added to the core *E. coli* metabolic model (http://gcrg.ucsd.edu/Downloads/EcoliCore) and the flux of the reactions corresponding to the knockout genes were set to zero. Experimentally measured GUR, cell growth rate and production rates for succinate, acetate, lactate, ethanol, and formate were used to fix their flux, and then ATP production rate was maximized as an objective function with setting allowLoops to false in the function of optimizeCbModel.

## 4-4 Results

### 4-4-1 Gene knockout simulation for design metabolic network for growth-coupled succinate production.

Genome scale metabolic model iAF1260 (81) was used to identify a candidate of gene deletions for succinate production from glycerol coupled with cell growth. Combinatorial deletion mutant of *pykAF-adhE-pflB-gldA* was predicted to produce high amount of succinate coupled with cell growth. The production yield was expected to reach 0.45 C-mol/C-mol at the optimal growth state.

FBA simulations predicted that wildtype *E. coli* catabolized glycerol via glycolysis, and formed acetate, ethanol, and formate (Figure 4-1A), whereas the knockout mutant produced succinate via Ppc and reductive TCA cycle (Figure 4-1B). Deletion of *pykAF* encoding pyruvate kinase blocked conversion of phosphoenolpyruvate (PEP) to pyruvate (Pyr) and rewired the carbon flow to Ppc. Deletion of *gldA* encoding glycerol dehydrogenase blocked supply of dihydroxyacetone (DHA) as a precursor of phosphotransferase system, which also converts PEP to Pyr. Because *adhE* and *pflB* were involved in the synthesis pathways of anaerobic fermentation production of ethanol and formate, respectively, disrupting these synthesis pathway enhanced succinate production in order to maintain the redox balance.

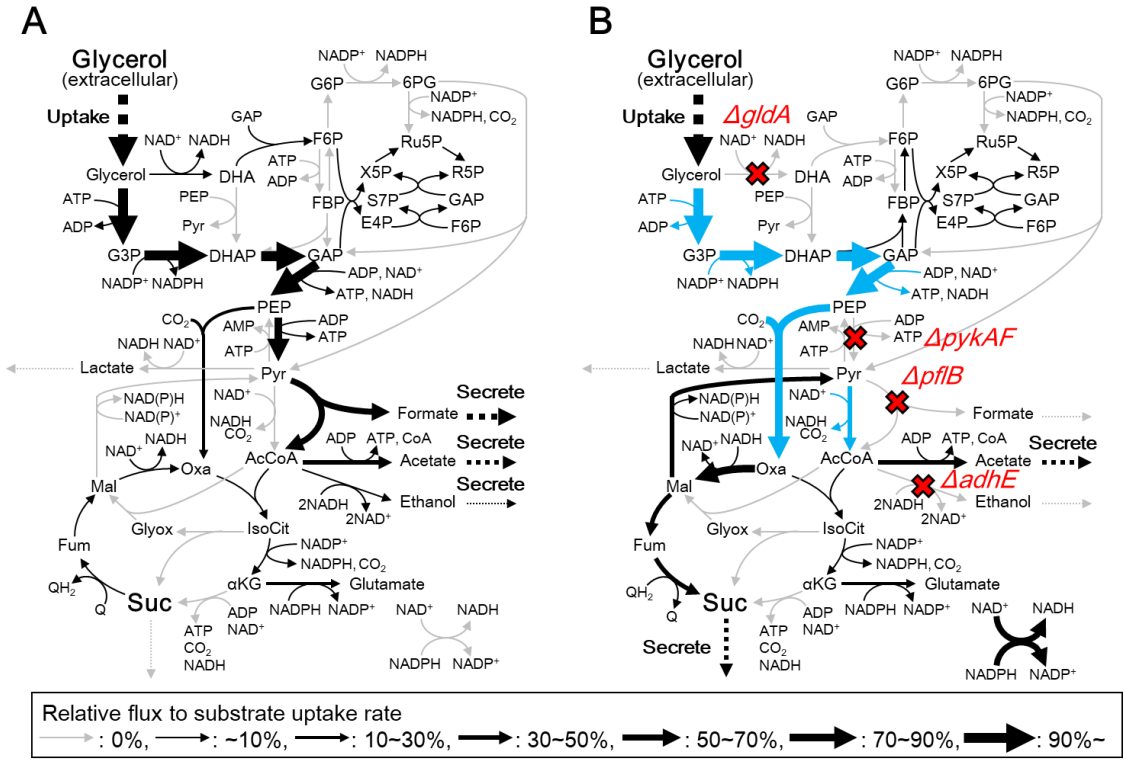



**Figure 4-1 Flux distributions for succinate production in *E. coli*.** Flux distributions at optimal growth state of wildtype *E. coli* (A) and *ΔadhEΔpykAFΔgldAΔpflB* model (B). Blue arrows in panel B indicate the predicted rate limiting reaction in the chapter 2.

### 4-4-2 Evaluation of succinic acid production of multiple knockout mutant of *pykAF, adhE, pflB* and *gldA*

Based on the gene knockout simulation using FBA, candidate genes (*pykAF-adhE-pflB-gldA*) were disrupted in wildtype *E. coli* BW25113, generating parent strain (strain P). The strain was cultivated in M9 medium using Erlenmeyer flasks. Succinate yield from the strain P (0.08 ± 0.00 C-mol/C-mol) was improved 3-fold in comparison with the wildtype strain (0.03 ± 0.00 C-mol/C-mol) (Figure 4-2A and Table 4-3).

FBA prediction is also effected by used influx values such as SUR and OUR (49). In order to accurately compare the result between experimental cultivation and FBA prediction, continuous cultivation using a bio-reactor is a better approach since it can make metabolic steady state, and measure SUR and OUR. Continuous cultivations for strain P was performed for comparing the succinate productivity of wet-experiment and FBA simulation. Although FBA predicted 0.32 C-mol/C-mol of succinate yield with setting glycerol uptake rate (GUR) and OUR to experimental values, the strain P experimentally produced only 0.01 C-mol/C-mol of succinate (Table 4-3). These gaps might be because the biomass yield of strain P (0.11 C-mol/C-mol) was lower than the predicted value (0.29 C-mol/C-mol). These results should be caused that some of rate-limiting reactions diminished the flux space of actual metabolic network of the knockout mutant as shown in the chapter 2, and the cells could not reach to the high succinate producing state.

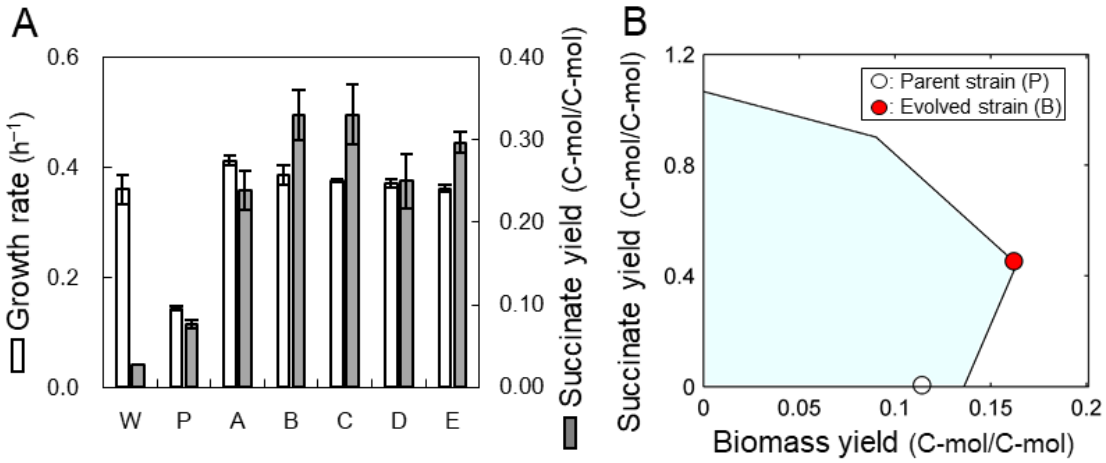



**Figure 4-2 Culture result of adaptive laboratory evolution.** (A) Specific growth rate and succinate production in a batch culture. The cells obtained after 100 generations on M9 medium were named as "strain A"-"strain E" in the descending order of growth rate. W and P indicate wildtype *E. coli* BW25113 and the knockout strain of *pykAF-adhE-pflB-gldA,* respectively. (B) Direct comparison of feasible solution space and experimental results of continuous culture. Metabolic solution space was calculated with measured values of GUR and OUR in iAF1260 (gray area) and the knockout model (blue area). GUR and OUR were set to the experimentally measured values of 3.37 mmol $gCDW^{-1}hr^{-1}$ and 2.85 mmol $gCDW^{-1}hr^{-1}$ in strain B, respectively.

**Table 4-3 Summary of the carbon-molar yield.**

| Condition & strains | Carbon yield (C-mol/C-mol) | | | | |
|---|---|---|---|---|---|
| | Biomass[*1] | Succinate | Lactate | Acetate | Ethanol |
| Batch culture | | | | | |
| (W): Wildtype | 0.20 ± 0.01 | 0.03 ± 0.00 | 0.11 ±0.00 | 0.19 ± 0.01 | 0.24 ± 0.02 |
| (P): Parent | 0.35 ± 0.01 | 0.08 ± 0.00 | 0.18 ± 0.15 | 0.12 ± 0.01 | 0 |
| (A): Evolved A | 0.44 ± 0.04 | 0.24 ± 0.02 | 0 | 0.23 ± 0.02 | 0 |
| (B): Evolved B | 0.27 ± 0.02 | 0.33 ± 0.03 | 0.08 ± 0.01 | 0.26 ± 0.02 | 0 |
| (C): Evolved C | 0.31 ± 0.04 | 0.33 ± 0.04 | 0.06 ± 0.01 | 0.28 ± 0.03 | 0 |
| (D): Evolved D | 0.25 ± 0.04 | 0.25 ± 0.03 | 0.05 ± 0.01 | 0.22 ± 0.03 | 0 |
| (E): Evolved E | 0.28 ± 0.01 | 0.30 ± 0.01 | 0.06 ± 0.01 | 0.26 ± 0.01 | 0 |
| I829S mutant[*2] | 0.28 ± 0.03 | 0.29 ± 0.01 | 0.02 ± 0.00 | 0.26 ± 0.02 | 0 |
| R849S mutant[*2] | 0.26 ± 0.02 | 0.28 ± 0.02 | 0.01 ± 0.00 | 0.27 ± 0.02 | |
| I829S-R849S mutant[*2] | 0.24 ± 0.01 | 0.28 ± 0.01 | 0.03 ± 0.00 | 0.25 ± 0.01 | |
| Continuous culture | | | 0 | | 0 |
| (P): Parent | 0.11 | 0.01 | 0 | 0.14 | 0 |
| (B): Evolved B | 0.16 | 0.45 | 0 | 0.29 | 0 |
| Simulation[*3] | | | | | |
| (P): Parent | 0.29 | 0.32 | 0 | 0.29 | 0 |
| (B): Evolved B | 0.16 | 0.44 | 0 | 0.32 | 0 |

[*1]For calculation of biomass yield, $OD_{600}$ was converted into dry cell weight using the conversion coefficient of 0.3 $gCDW \cdot L^{-1} \cdot OD_{600}^{-1}$, and carbon-mol in the biomass was calculated based on the biomass composition described in the

iAF1260 model.

*2These strains were constructed by introducing of the identified *ppc* mutations into the genome of strain P.

*3The values were calculated by FBA using the genome-scale metabolic model iAF1260 adding constraints of experimentally measured values for GUR at 4.77 and 3.37 mmol $gCDW^{-1}hr^{-1}$, and OUR at 4.14 and 2.85 mmol $gCDW^{-1}hr^{-1}$ in strain P and B, respectively.

### 4-4-3 Adaptive laboratory evolution.

Since the metabolic pathway of strain P was designed for growth-coupled succinate production, the evolved mutants with improved growth via ALE would have enhanced succinate production. Five evolved strains (strain A~E) were obtained by parallel passage of cultivations for about 100 generations (Figure 4-3). The specific growth rates were improved by more than 2-fold during ALE experiments. All of the evolved strains successfully increased the yield of succinate by more than 3.1-fold in flask cultivation (Figure 4-1A). Strain B and C were the highest producers of succinate with 0.33± 0.03 and 0.33± 0.04 C-mol/C-mol yields, respectively. The main byproduct in all the evolved strains was acetate with 0.22 to 0.28 C-mol/C-mol yield (Table 4-3). During the continuous cultivation, strain P showed almost no production of succinate, on the other hand strain B showed the higher yields of succinate with 0.45 C-mol/C-mol yield, as predicted by FBA simulation (Figure 4-1B). The biomass yield in strain B was almost the same as the predicted value. Furthermore, byproduct formations such as acetate and lactate were consistent with the FBA result (Table 4-3).

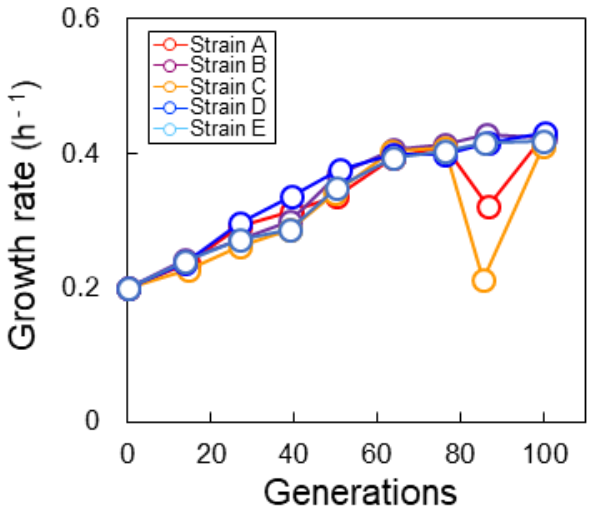



**Figure 4-3　Growth rate dynamics during adaptive laboratory evolution of the knockout mutant of *adhE-pykAF-gldA-pflB*.** Adaptive laboratory evolution was performed by five parallel passage cultivations. The evolved strains were isolated as colonies from culture broth after 100 generations. The growth rates of strain A and C were temporarily decreased around 85 generations due to aggregation.

### 4-4-4 Fed-batch fermentation of evolved strain B.

For evaluating the succinate titer, fed-batch fermentation was performed by using the highest succinate producer of evolved strain B. Feed of glycerol and $NaHCO_3$ started after 6hr of main cultivation. Final succinate titer reached to 9.7 g/L on 50hr of fed-batch fermentation (Figure 4-4). This value was 12.1-fold higher than of the culture result of the evolved strain B on 24hr of batch cultivation using Erlenmeyer flasks. Acetate was main byproduct as predicted by FBA, and the final titer of acetate reached to 7.5 g/L.

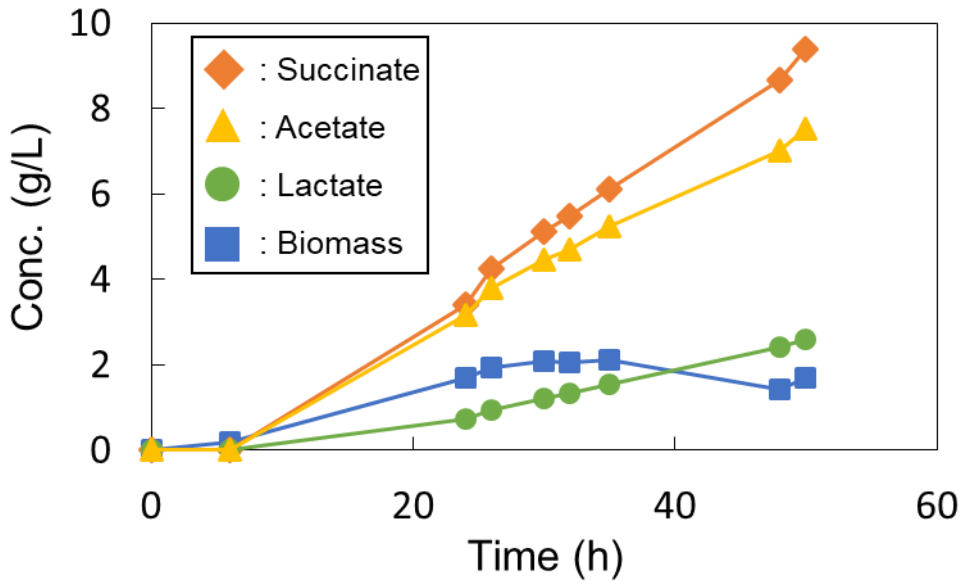



**Figure 4-4　Result of fed-batch fermentation of evolved strain B.**

4-4-5 Mutation analysis and reverse engineering.

For evaluating rate limiting reaction for succinate production in the unevolved knockout strain, genomes of all evolved strains and strain P were sequenced using high-throughput sequencer Illumina MiSeq. The raw sequences generated, 5.3-6.5×$10^8$ base pairs, showed an average coverage of approximately 100-140-fold of 4.63 ×$10^6$ *E. coli* BW25113 genome (GenBank: NZ_CP009273). Mutations were identified by the breseq pipeline (149) using the genome sequence of *E. coli* BW25113 as reference. All of the evolved strains had more than one mutation in coding regions of *ytfT, cyaA, glpK, ppc,* and *eutH* (Table 4-4). Especially, mutations of *ppc* encoding Ppc, which was the candidate of rate limiting reaction, were shared at I829S in strain A and B or R849S in strain C, D and E. Although *glpK* encoding glycerol kinase was also predicted as the candidate for rate-limiting reaction, the *ppc*

mutations must be having a dominant effect on adaptation as the highest producer of succinate (strain B and C) had mutations only in *ppc*, not in *glpK*. Comparing the flux distributions calculated by FBA, the flux of Ppc in the knockout mutant was higher than that of the wildtype *E. coli* (Figure 4-1). These results suggested that the reaction of Ppc had kinetic limitations for achieving the optimal metabolic state, and introduction of these mutations in *ppc* led to changes in the enzyme properties for enhancing the activity.

To evaluate the effect of these identified mutations on succinate production, mutations in *ppc* were introduced into the genome of strain P by multiplexed genome engineering. Introducing mutations of I829S or R849S in strain P improved the cell growth rate by about 2.0-fold, and increased the yield of succinate to 0.29 ± 0.01 or 0.28 ± 0.02 C-mol/C-mol, respectively, which were 89% or 84% of the yield from the highest producers, strain B and C (Table 4-3 and Figure 4-5). These results indicated that the mutations identified in *ppc* were important for succinate production in strain P. On the other hand, the engineered mutant containing both the I829S and R849S mutations displayed the same amount of succinate production as the single mutants (Figure 4-5). Since there was no synergistic effect of the two mutations on succinate production, these mutations likely affect the same enzymatic characteristics.

**Table 4-4 Identified mutations in the evolved strains.**

| Strain | Gene | Mutation | Type |
|---|---|---|---|
| A | *ytfT* | C393T | I130I |
| | *cyaA* | +C (1901/2547t) | Frameshift |
| | *glpK* | G762T | Q254H |
| | *ppc* | T2486G | I829S |
| B | *ppc* | T2486G | I829S |
| C | *ppc* | C2545A | R849S |
| D | *ppc* | C2545A | R849S |
| E | *ppc* | C2545A | R849S |
| | *eutH* | C936A | N312K |

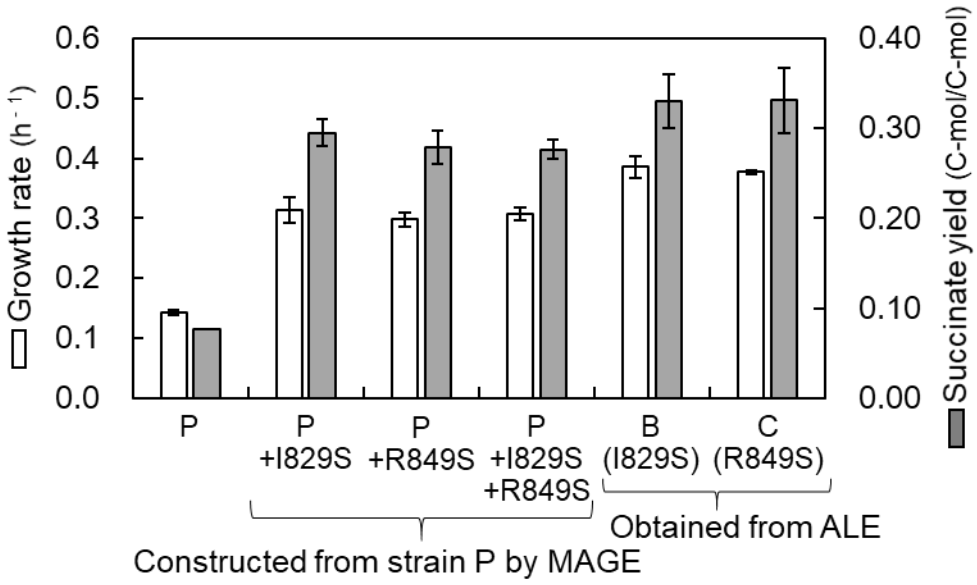



**Figure 4-5 Batch culture results for engineered *E. coli* after introducing the mutations of Ppc$^{I829S}$, Ppc$^{R849S}$, or both by MAGE.**

4-4-6 Functional analysis of the mutations of Ppc.

For characterization of the mutated amino acid residues of Ppc, multiple sequence alignment was performed with clustalW (Kyoto University

Bioinformatics Center: http://clustalw.genome.jp, Figure 4-6). The isoleucine residue at position 829 in *E. coli* Ppc is conserved in *Mannheimia succinicproducens* and *Thermus thermophilus* HB8, whereas other microorganisms including *Corynebacterium glutamicum* have a valine residue at the equivalent position of Ile829 in *E. coli*. The arginine residue at position 849 is highly conserved in various microorganisms. Functional analysis of the mutants of Ppc was performed to understand how the identified mutations in *ppc* changed enzymatic properties. Protein structure of Ppc has been reported previously by Kai et al. (151). The model displayed that the mutated positions at Ile829 and Arg849 residues were in the same α-helix strand (Figure 4-7A, yellow). Four residues, Lys773, Arg832, Arg587, and Asn881, are involved in L-aspartate binding, and among these, Arg832 is located in the same α-helix strand that contains Ile829 and Arg849 (Figure 4-7B). Because L-aspartate allosterically inhibits Ppc, it was hypothesized that the I829S and R849S mutations should decrease the inhibition by aspartate, leading to an increase in the flux of Ppc for succinate synthesis.

The effect of the replacement of Ile829 or Arg849 by serine in Ppc was examined by *in vitro* enzymatic analysis. The enzymatic activities of Ppc were measured by a coupling reaction catalyzed by malate dehydrogenase and measuring the absorbance of NADH at 340 nm. The activity of wildtype Ppc ($Ppc^{Wild}$) was drastically decreased by ca. 78% in the presence of 1 mM L-aspartate with 4 mM PEP (Figure 4-7F). On the other hand, Ppc mutants ($Ppc^{I829S}$ and $Ppc^{R849S}$) displayed lower levels of the inhibition compared to $Ppc^{wild}$ (Figure 4-7G and 4-7H). Especially, $Ppc^{I829S}$ showed no sensitivity to

L-aspartate inhibition with different concentrations of PEP tested from 0.2 mM to 4 mM.

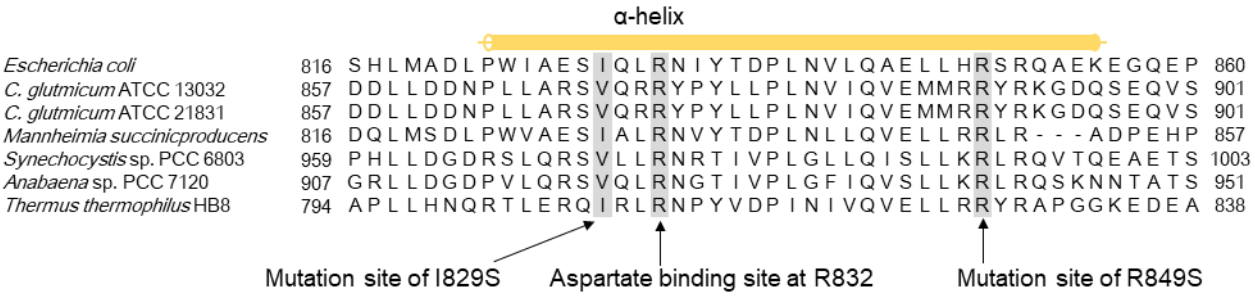



**Figure 4-6 Multiple sequence alignment of Ppc.**

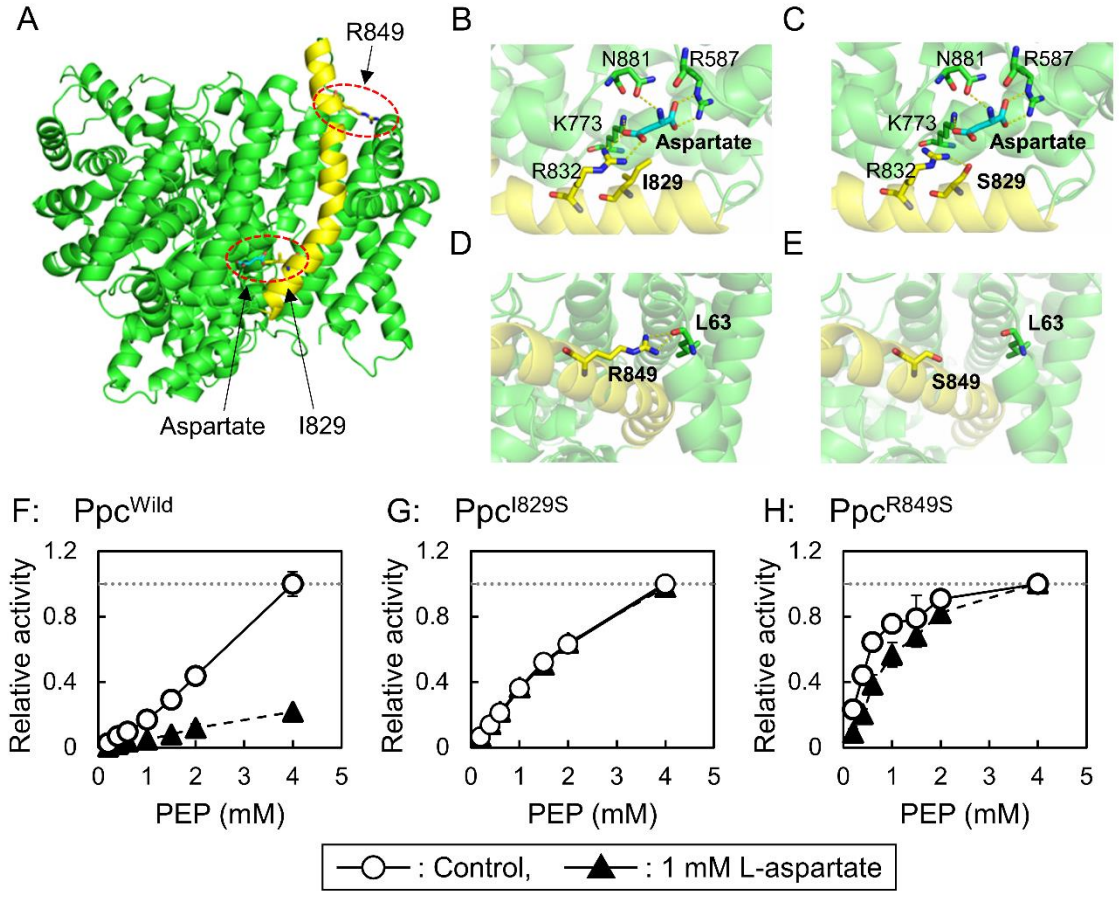



**Figure 4-7 Functional analysis of identified mutations in Ppc.** (A) Protein structure of Ppc wildtype monomer. (B-E) Close view of mutated site for I829S (B, C), and R849S (D, E). Yellow dashed line indicated hydrogen bonds (< 3.2 Å). (F-H) Specific activity of Ppc of wildtype (F), I829S mutant (G) and R849S mutant (H) with (black triangle) or without (open circle) 1 mM aspartate. The values indicate the relative activities against the activity with 4 mM PEP and in the absence of L-aspartate.

### 4-4-7 Metabolic profiling analysis.

For verification of *in vivo* effect of allosteric inhibition of L-aspartate to Ppc, intracellular metabolites in central carbon metabolism and amino acids

were quantified by gas chromatography–mass spectrometry (Figure 4-8). Aspartate concentration in strain P was 3.05 ±0.35 mM, which was three times higher than the concentration that could cause 78% inactivation of Ppc activity *in vitro*. Furthermore, PEP in strain P was 12.5-fold higher than of wildtype *E. coli* (6.04 ± 0.20 mM and 0.48 ± 0.06 mM, respectively). Compared with the accumulations of these metabolites in strain P, aspartate concentration in strain B had no significant change (2.72 ±0.35 mM), on the other hand, PEP concentrations was drastically decreased to 0.75 ± 0.10 mM. These results indicated that the evolved strains could overcome this metabolic limitation by introducing single amino acids substitution of Ile829Ser or Arg849Ser in Ppc.

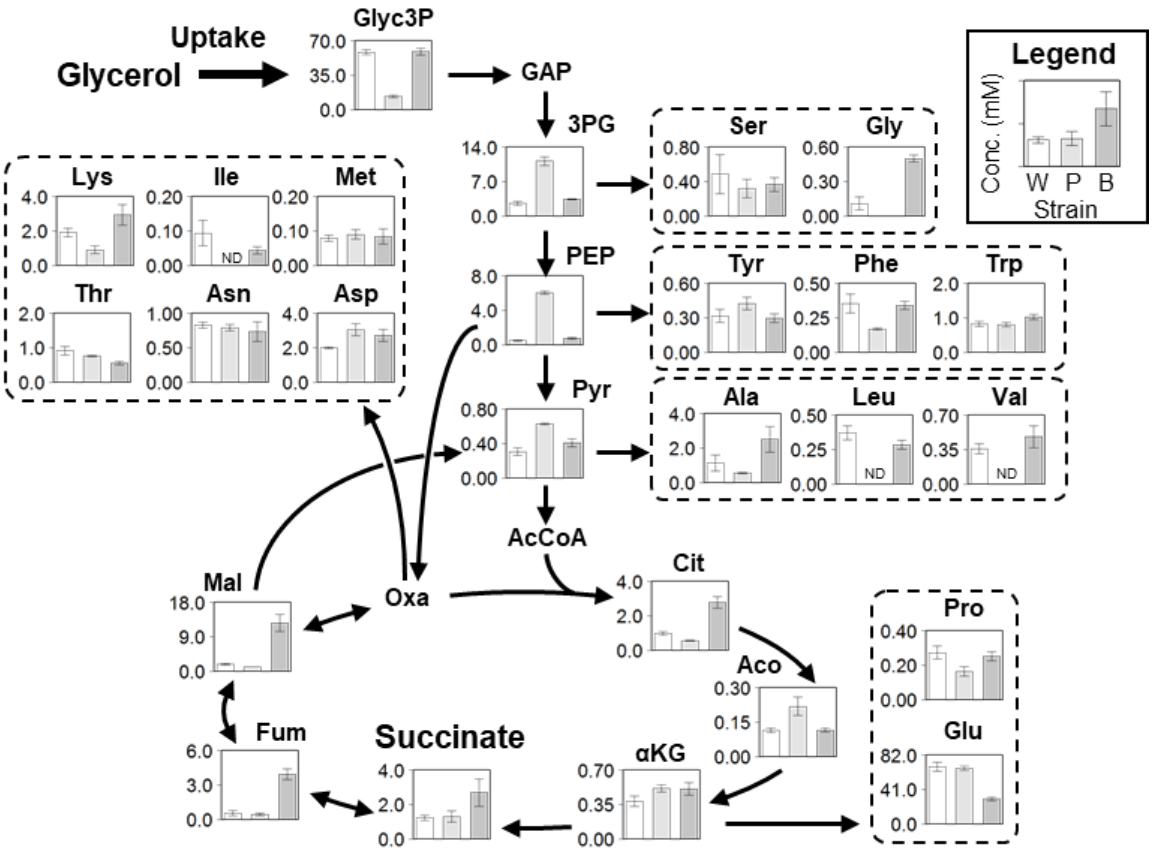



**Figure 4-8 Metabolic profiling of central carbon metabolism and amino acids**. The values indicate absolute concentration of intracellular metabolites of wildtype, strain P and B as quantified by GC-MS.

## 4-5 Discussion

Although there are many reports that the enhanced production of target metabolites using pathway engineering by gene deletion based on *in silico* design with FBA, most of these applications did not achieve the predicted value at the optimal growth state (Table 1-1). One possible cause of these inconsistencies should be that some of metabolic reaction exist as rate-limiting reactions due to enzymatic capacity or regulation. In the chapter 2, I revealed that the knockout mutant of *pykAF-adhE-pflB-gldA* have 9 possible rate-limiting reactions involved in glycerol assimilation, glycolysis and anaplerotic reaction for succinate production. In case of the 3HP production discussed in the chapter 3, the mechanism causing the rate-limiting step was revealed by experimental evaluation, and this led the further strain improvement. Although experimental evaluation is a powerful approach to directly identify the rate-limiting reactions *in vivo*, it is difficult to apply the larger number of target reaction due to experimental cost and technical limitations. In the present chapter, the mechanism causing the rate-limiting step on the succinate production was revealed through ALE experiment of the knockout mutant based on FBA.

Mutation analysis of high-succinate producing strains obtained from ALE displayed that the all evolved strains have more than one mutation into their chromosomal DNA. The mutations included coding regions of *ppc* and *glpK,* which were predicted as representative genes involved in the possible rate-limiting reactions. The Ppc was experimentally confirmed to be mediating a dominant rate-limiting step. Mutation of *glpK* encoding glycerol

kinase was only observed in evolved strain A. Previous reports demonstrated that a point mutation in the *glpK* (*glpK* 218a>t) gene promoted the growth of wildtype *E. coli* on glycerol medium by decreasing inhibitory effect of fructose 1,6-bisphosphate on GlpK (147)(155). Since the mutation site of *glpK* 762 g>t occurred in strain A was different with the previous one. Although the *glpK* mutation might increase the cell growth by increasing glycerol catabolism, the *ppc* mutations are much effective in their metabolic adaptation. The other candidates such as glycolytic reactions were not observed in all of the evolved strains. Previous report mentioned that glycolytic fluxes in *E. coli* are mainly controlled by pyruvate kinase (156)(157), and the reversible glycolytic reactions to enolase are governed by PEP consuming irreversible reaction (88). Since the knockout mutant lacked pyruvate kinase, and Ppc was a main PEP consuming reaction, the candidates in glycolytic reaction turned out to be non-rate limiting steps.

Ppc in *E. coli* is tightly regulated by L-aspartate, which bind to its four residues (Lys773, Arg832, Arg587, and Asn881) (151). Protein structure model of I829S mutant of Ppc suggests that the mutation pushes the residue of Arg832 away from L-aspartate binding by forming a hydrogen bond (2.9 Å) between the hydroxyl residue of Ser829 and the amino group of Arg832 (Figure 4-7B and 4-7C). Arg849 residue has a hydrogen bond with the residue of Leu63 (Figure 4-7D). Because the mutated residue of Ser849 cannot bind the residue of Leu63, the R849S mutant should change the tertiary structure of Ppc and decrease the affinity for L-aspartate (Figure 4-7E). *In vitro* enzyme assay displayed that the activity of $Ppc^{wild}$ was

inhibited to ca. 78% by 1 mM L-aspartate (Figure 4-7G), and the intracellular concentration of aspartate was high at 3.05 ± 0.35 mM in strain P (Figure 4-7). These results suggest that Ppc in unevolved knockout mutant was mostly suppressed. Flux distribution calculated by FBA with constraints of measured fluxes indicated that Ppc flux of strain P on continuous cultivation was 2.8% of GUR, whereas strain B showed higher Ppc flux value at 94.5% of GUR (Figure 4-9). Higher accumulation of PEP in strain P (6.04 ± 0.20 mM) than in strain B (0.75 ± 0.10 mM) was also consistence with these results (Figure 4-8). By setting the flux of Ppc to 2.8% of GUR, the knockout mutant model cannot achieve the optimal growth state without any constraints of Ppc (Figure 4-10A). Therefore, evolved strains alleviate the kinetic limitation of Ppc by introducing single amino acid substitutions generating $Ppc^{I829S}$ or $Ppc^{R849S}$, and optimized their metabolic network to improve succinate production as expected by FBA (Figure 4-10B). In other microorganisms such as *C. glutamicum*, Ppc is also thought to be possible rate limiting reaction since the enzyme is tightly regulated by L-aspartate (158)(152)(159). Multiple sequence alignment displayed that the amino acid residues of Ile829 and Arg849 are highly conserved in these microorganisms (Figure 4-6). This indicated that the introduction of mutations identified in Ppc should improve the flux and target production in other microorganisms also.

Quantitative evaluation of the strength of the constraint of rate-limiting reaction is useful to increase the prediction accuracy of metabolic state by FBA. Without considering the kinetic limitation of Ppc in the genome-scale

metabolic model of *E. coli* iAF1260, the model overestimated the yields of biomass and succinate (Figure 4-11, right bars). On the other hand, the predicted values were close to the culture result (Figure 4-11, left bars) by considering the additional constraint by setting the upper-bound of Ppc's flux to 2.8% of GUR (Figure 4-11, center bars), which is the measured value by constraint based metabolic flux analysis. Further evaluation of the metabolic limitations in actual cells designated for several fermentation process should improve the prediction accuracy and develop the metabolic design for strain improvement.

Metabolic regulation existing in wildtype microorganism is a main obstacle for enhancing target production. Although stoichiometric models are normally used without any kinetic constraints, the effect of metabolic regulation on target production can be evaluated by considering it as an additional constrain. Recent studies have added the additional constraints such as enzymatic regulation (32), expression regulation (33) or thermodynamic feasibility (34) on flux balance modeling for improvement in simulation accuracy. Adding such constraints on our approach should improve prediction accuracy of rate limiting reactions and enable identification of target genes for genetic manipulation. The next question to consider would be on how to modify the enzymatic properties. Since small molecules can affect the activity of metabolic enzymes, engineering of enzymes with desired catalytic properties is important for enhancing the metabolic flow. Nevertheless, its application in metabolic engineering is still very challenging. Evolution engineering such as ALE enables to construct

ideal mutant enzymes or strains without having any prior knowledge on the properties of catalytic enzymes. In the present study, we revealed that Ppc is one of the rate limiting reactions in the actual metabolic network, and identified effective mutations of $Ppc^{I829S}$ or $Ppc^{R849S}$ for overcoming the metabolic limitation. Further identification of beneficial mutations with integrated use of *in silico* strain design and ALE will help us to understand the genotype-metabolic phenotype relationship for genome-scale metabolic design.

Evolved strain B show the highest production yield of succinate in the batch cultures of all evolved strains and produced 9.7 g/L succinate on 50hr of fed-batch fermentation (Figure 4-4). This value was higher than of the previous study using an engineered strain disrupting *maeA* on high succinate producing strain obtained from ALE of pyruvate kinase deficient *E. coli* (6.0 ± 0.19 g/L) (139). The other study achieved the higher succinate titer of 14 g/L in 72hr of cultivation of an engineered strain disrupting *adhE, pta, poxB, ldhA* and *ppc* and overexpressing *pyc* on a plasmid vector. The one of advantages using the evolved strain B for industrial fermentation process is that the strain B does not have any plasmid vector for metabolic engineering, since expressing genes in plasmid vector is thought to lead unstable production on the commercial scale plant.

Acetate was confirmed as main byproduct in the fed-batch fermentation of evolved strain B (7.5 g/L). FBA predicted that succinate production in the knockout mutant increased with decreasing oxygen uptake rate, while the acetate production was decreased with it (Figure 4-12). Thus, optimization of

culture condition such as changing aerobic condition should contribute further development of succinate production process from glycerol by using evolved strain B.

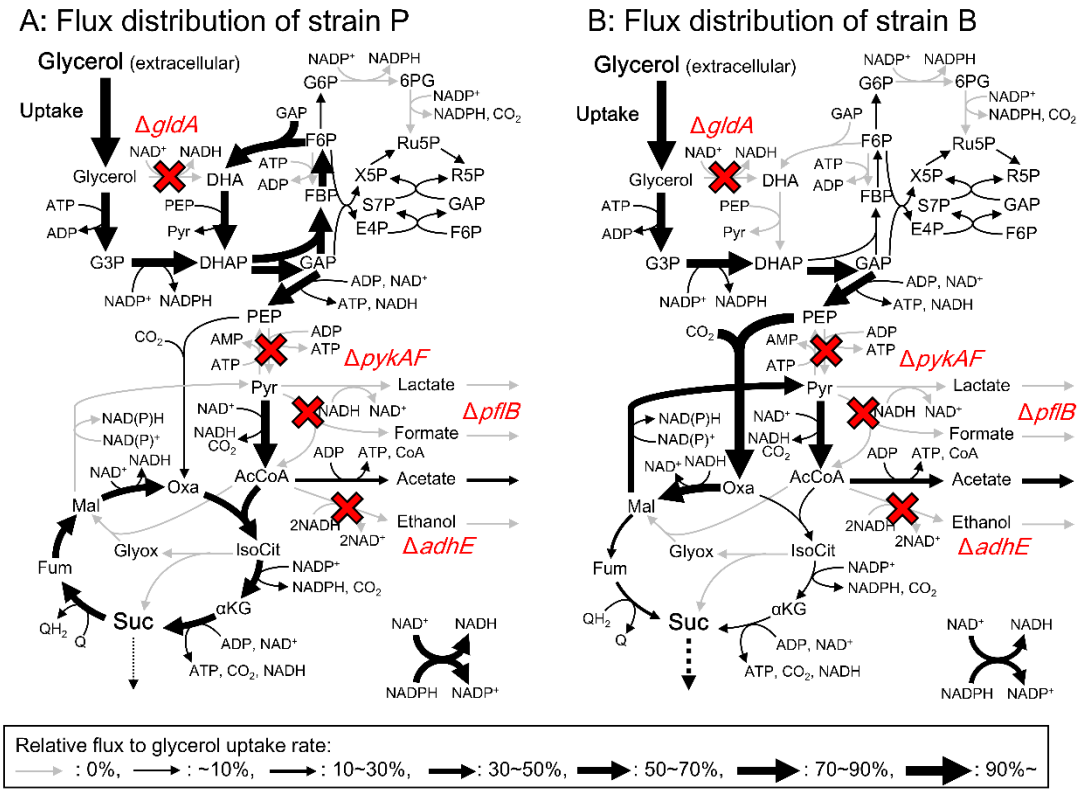



**Figure 4-9 Flux distributions of strain P (A) and B (B).** Estimated fluxes were normalized to a GUR of 100. The flux distributions were calculated as linear programming by optimizeCbModel function in COBRA toolbox. Briefly, the experimentally measured GUR, cell growth rate, and secretion rates of succinate, acetate, lactate and formate were used as constraints, and then ATP production rate was maximized as the objective function. The stoichiometric model used was constructed from the core *E. coli* metabolic model by adding glycerol assimilation pathway and disruption of the reactions corresponding to the knockout genes.

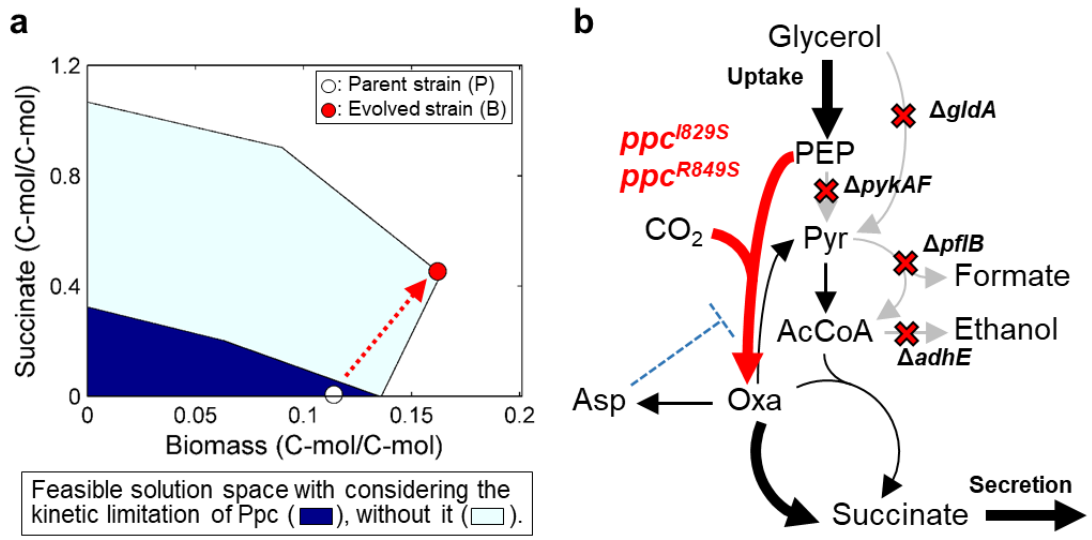



**Figure 4-10 Adaptation trajectory of the knockout mutant for improving succinate production**. (A) Computationally predicted feasible solution space in metabolic network with or without considering the kinetic limitation of Ppc. Experimentally measured GUR and OUR of strain B were considered as constraints, and the upper bounds of Ppc flux was set to free or 2.8% of GUR. White and red circles indicate the experimental results of strain P and B. (B) Overview of metabolic evolution for succinate production demonstrated in this study.

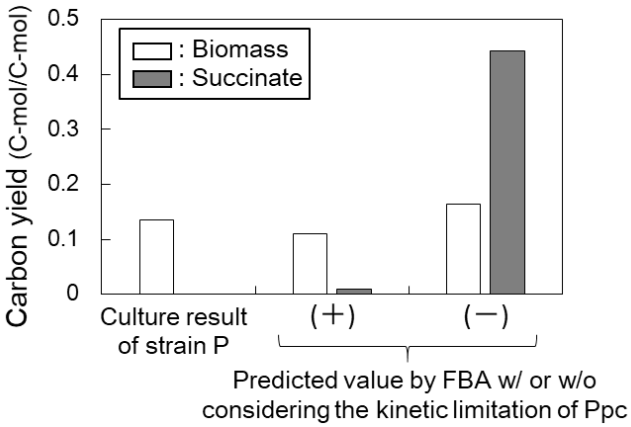



**Figure 4-11 Comparison of the experimental result of the knockout mutant and optimal values by FBA.** Left bars indicate the culture result of strain P on continuous culture experiment. Center bars indicate the predicted values by FBA using revised metabolic model with setting the upper-bound of Ppc's flux to the measured flux. Right bars indicate the predicted values by FBA without any metabolic constraint (i.e. original metabolic model).

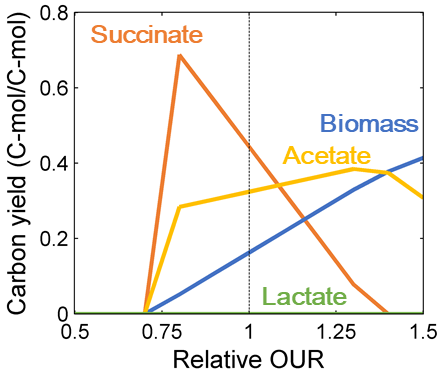



**Figure 4-12 *In silico* evaluation of the effect of OUR on succinate production by evolved strain B.** Carbon yield of biomass and each metabolites at the optimal growth state were calculated by using FBA and iAF1260. GUR was set to the experimentally measured value of 3.37 mmol $gCDW^{-1}hr^{-1}$. OUR was set to the relative values of the experimentally measured value of 2.85 mmol $gCDW^{-1}hr^{-1}$.

## 4-6 Summary

When there are many or unknown options to improve the rate limiting reactions, evolution engineering is powerful approach to screen the more effective strategy for enhancing the target production. Since the metabolic system of the knockout strain was designated for growth-coupled target production, ALE is enable to enhance the target production. In this chapter, ALE successfully improved the growth-coupled succinate production in the knockout mutant by 3-fold, and the productivity of the evolved strain reached to the predicted value by FBA (0.45 C-mol/C-mol). The other advantage of ALE is easy to identify the effective mutation on increased fitness for the perturbation of gene deletion, because mutation rate during ALE is lesser than of random mutagenesis approach. Analyzing the effect of identified mutations on the enzymatic function can contribute to understand that the molecular mechanism causing and overcoming the metabolic limitation. Therefore, further identification of beneficial mutations using evolution engineering will accelerate the genome-scale metabolic design for improving target production with compensating our lacking knowledges of the genotype-metabolic phenotype relationship.

# Chapter 5: General conclusion and future perspective

## 5-1 General conclusion

Microbial production of valuable chemicals has been attracted as an alternative way to compensate the increasing demand for commercial use, since oil resources are limited and unsustainable. Strain improvement for enhancing target production is an important for development of microbial production process. Recent development of computational approaches enable us to design rational metabolic network for enhancing target production *in silico*.

FBA is one of the most widely used approach to predict metabolic behavior and build metabolic engineering strategies such as gene knockout. Although microbial production of various target metabolites have been improved based on FBA, the target productivities were always lower than of the predicted values. These inconsistencies should be arisen from that FBA only considers mass-balance equation on metabolic network for prediction of metabolic behavior, while metabolism in actual cell is regulated by complicated interactions of various factors.

The general objective of this study was to reveal the reason why FBA overestimates the productivity of growth-coupled target production in engineered strains. The hypothesis causing the inconsistency was validated through development of *in silico* modeling approach and experimental strain improvements.

In the chapter 2, in order to validate the hypothesis *in silico*, simple screening algorithm of key enzyme for metabolic engineering was developed by addition of a constraint on conventional FBA calculation. The proposed method simplify the complex metabolic regulation as an upper-bound constraint on enzymatic flux and evaluate changing productivity at optimal growth state. When a limited reaction decreased target productivity, the reaction was thought to be a possible rate-limiting reaction. The predicted reactions in growth-coupled 1,4BDO production were consistency with the optimized enzymatic reactions for the experimental strain improvement performed in the previous study. Furthermore, the proposed method successfully screened possible rate-limiting reactions in various fermentation process. The biggest advantage of the proposed approach use only the existing genome-scale metabolic model for the prediction without any experimental data sets. On the other hand, since this approach is simple screening method and impossible to know the reason why the reactions are limited and which reaction mostly effected target production *in vivo*, understanding the molecular mechanism causing rate-limiting reactions is needed for build metabolic engineering strategy.

In the chapter 3, experimental evaluation of double knockout mutant of *tpiA* and *zwf* for growth-coupled 3HP production was performed to understand the molecular mechanistic causing the rate-limiting step. Overflow of 1,3PDO was thought to decrease precursor supply for 3-hydroxypropionaldehyde dehydrogenase, which was predicted as a possible rate-limiting reaction. Increased activity of methylglyoxal bypass pathway

should induce the expression of *yqhD* encoding NADPH-dependent aldehyde reductase, and this led to the overflow of 1,3PDO. Additional disruption of *yqhD* based on the experimental evaluation successfully improved 3HP production from 0.20 C-mol/C-mol to 0.34 C-mol/C-mol. These results indicated that experimental evaluation is a useful approach to reveal the mechanism causing rate-limiting reaction and build next metabolic engineering strategy.

In the chapter 4, ALE of knockout mutant of *adhE, pykAF, gldA* and *pflB* was performed for increasing growth-coupled succinate production, and understanding the mechanism causing the rate-limiting reaction. All of the evolved strain obtained from five parallel passaged cultivation successfully increased succinate production yield by more than 3-fold. The highest producing strain achieved the production yield of 0.45 C-mol/C-mol, which was almost same as the predicted value by FBA. Experimental analysis of evolved strains displayed that the allosteric inhibitor L-aspartate tightly regulated the activity of Ppc, which was one of the predicted possible rate-limiting reactions, and decreased target production in the unevolved knockout mutant. The evolved strains overcame the kinetic limitation for increasing succinate production by introducing novel mutations in *ppc*. These results indicated that ALE is a powerful approach not only for strain improvement of growth-coupled target production, but also for identifying key mutations overcoming the metabolic limitation.

In conclusion, this study first verified the hypothesis that some of kinetic constraints caused the overestimation of the effect of gene knockout on the

target productivity through development of novel computational approach and experiments of strain improvement.

### 5-2 Contribution of the present results for strain improvement

The present results in this doctor thesis verified the hypothesis that kinetic constraints diminish the flux space of actual metabolic network and thereby decrease the target productivity on growth-coupled target production based on FBA design. Here, the contributions of the present results for strain improvement were discussed.

In general, the strain improvement for optimizing the production of valuable chemicals is an overarching challenge in biotechnology. Recent development of computational modeling approach of metabolism enable to identify the effective knockout mutant for the growth-coupled target production *in silico*. The present results in this thesis give the lesson that considering the flux space of actual metabolic network is important for further strain improvement for growth-coupled target production based on the computational approach. Recently, Kamp & Klamt reported that rational design of the metabolic network for the growth-coupled target production by gene deletion is capable for almost all metabolites including carbon-based foods, bio-plastic, bio-fuel and pharmaceutical in five major production organisms of *E. coli*, *S. cerevisiae*, *C. glutamicum*, *A. niger* and *Synecocystis* sp. PCC 6803 (160). Because the screening method of possible rate-limiting reaction developed in the chapter 2 does not need any experimental data sets and only use existing genome-scale metabolic models, the method is easily

available for various microorganisms including them. The integrated use of the conventional gene knockout simulation and the screening method of possible rate-limiting reaction developed in the chapter 2 will accelerate to development of the various microbial processes of growth-coupled target production.

The proposed screening method in the chapter 2 suggested that Ppc is a possible rate-limiting step in the production of industrially produced metabolites such as 1,4BDO and succinate. As shown in the chapter 4, experimental analysis of intracellular metabolism and protein function revealed that Ppc was tightly regulated by aspartate *in vivo*, since aspartate was highly accumulated independent on their genetic background (Figure 4-8). Novel mutants of Ppc were identified by mutation analysis of the evolved strains by ALE in the chapter 4. The mutants of Ppc are insensitive to the inhibition by aspartate, and have enough enzymatic activity to increase the flux of Ppc. In the other industrial microorganism such as *C. glutamicum*, Ppc is also allosterically regulated by aspartate, and the amino acid residues of the mutated sites are highly conserved in them (Figure 4-6). Introducing the same mutations on the coding region of *ppc* of the industrially used strains is expected to increase the productivities.

### 5-3 Future perspective for strain improvement

The proposed screening method is available for screening possible rate-limiting reactions in various host cell factories, if only there are existing genome-scale metabolic models. To understand the mechanism causing the

rate-limiting reaction is needed for build metabolic engineering strategy by referring to the priori knowledges or using experimental evaluation. Optimization of the expression level of corresponding gene can be made by disrupting transcriptional repressor, changing the sequence of 5' untranslated region, or using synthetic biology tools such as promoter library (161). The next question to consider would be on how to modify the enzymatic properties. When enzymatic property causes the rate-limiting step, protein engineering for desirable catalytic properties is required for enhancing the metabolic flow. Nevertheless, its application in metabolic engineering is still very challenging. Evolution engineering such as ALE enables to construct ideal mutant enzymes or strains without having any prior knowledge on the properties of catalytic enzymes. The result of this study displayed that Ppc is the rate-limiting reactions in the actual metabolic network, especially in industrially produced metabolites of 1,4BDO and succinate, and identified effective mutations of $Ppc^{I829S}$ or $Ppc^{R849S}$ for overcoming the metabolic limitation. The approaches of overcoming the metabolic limitations were summarized in Table 5-1. Further identification of beneficial mutations will help us to develop genome-scale metabolic design for microbial cell factories.

**Table 5-1 Metabolic engineering strategies for overcoming the metabolic limitation predicted in the chapter 2.**

| Optimization target | Metabolic engineering strategy | Reference |
|---|---|---|
| Local optimization | | |
| Phosphoenolpyruvate carboxylase | ● The replacement of Ile829 or Arg849 in *ppc* by serine. | Chapter 4 |
| Pyruvate dehydrogenase | ● Using $lpdA^{D354K}$ mutant derived from *K. pneumoniae.* | (58) |
| | ● Disrupting transcriptional regulator of *arcA* | |
| Succinyl-CoA synthase | ● Disrupting transcriptional regulator of *arcA* | (58) |
| Succinate semialdehyde dehydrogenase | ● Disrupting transcriptional regulator of *arcA* | (58) |
| Glycerol dehydratase | ● Expressing glycerol dehydratase reactivase encoded by *grdAB* derived from *K. pneumoniae* | (92) |
| | ● Addition of enough amount of coenzyme $B_{12}$ | (89) |
| 3-hydroxypropionaldehyde dehydrogenase | ● Disrupting *yqhD* encoding NADPH-dependent aldehyde reductase | Chapter 3 |
| | ● Expressing superior enzymes derived from other microorganism, rather than of using *E. coli aldH*. | (92) |
| Pathway activity | | |
| Lower glycolysis | ● Increasing activity of PEP consuming reaction. | Chapter 4, (88), (58) |
| Reductive TCA cycle | ● Disrupting transcriptional regulator of *arcA* | (58) |
| Glycerol assimilation pathway | ● The replacement of Asp72 in *glpK* by valine | (147)(155) |

## Acknowledgements

This study was carried out in Metabolic Engineering Laboratory, Graduate School of Information Science and Technology, Osaka University. I am extremely thankful for my supervisor, Professor Hiroshi Shimizu for his guidance, passion and scientific lead. I really appreciate for his giving the opportunity for me to study about strain improvement in his laboratory. I am also thankful to Professor Dr. Fumio Matsuda, Symbiotic Network Design Laboratory, Graduate School of Information Science and Technology, Osaka University for given constructive comments, advices and encouragements with drinking together so many times.

I am also thankful to each member of my thesis advisory committee: Professor Dr. Taro Maeda, Professor Dr. Naoki Wakamiya and Professor Dr. Hideo Matsuda, Department of Bioinformatic Engineering, Graduate School of Information Science and Technology, Osaka University for all their constructive comments and advices on my study.

I would like to thank Dr. Yoshihiro Toya, Metabolic Engineering Laboratory, Department of Bioinformatic Engineering, Graduate School of Information Science and Technology, Osaka University for his many helps and constructive comments, and Dr. Katsunori Yoshikawa for his kind guidance on my study. And I am also especially thankful to all members in Dr. Shimizu's laboratory for every supports in my student life.

I would like to thank Prof. Chikara Furusawa and Dr. Takaaki Horinouchi, Quantitative Biology Center, RIKEN for technical supports on

genome resequencing analysis and constructive comments on the study of adaptive laboratory evolution.

I would like to thank Dr. Kazufumi Hosoda, Graduate School of Information Science and Technology, Osaka University for technical supports on enzymatic assay experiment, and Associate Professor Dr. Takashi Hirasawa, Department of Bioengineering, Tokyo Institute of Technology for technical supports on genetic manipulation.

I am thankful to the financial supports by a Grant-in-Aid for JSPS Fellows (DC1) and "Program for Leading Graduate Schools" of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

I am thankful for Professor Dr. Mattheos Koffas, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute in U.S.A. for giving me the great opportunity of my visiting study in Dr. Koffas's laboratory. I am also thankful for Dr. Bready Cress, Dr. Rufeng Wang and Dr. Trevor John Simmons in Rensselaer Polytechnic Institute for their passions, helps and kindness.

I am thankful for my friends in the Humanware-innovation program, Osaka University for their passions, helps and kindness.

Finally, I would like to thank my parents for their support, love and continuous encouragements throughout many years.

## Reference


1. McGovern PE, et al. (2004) Fermented beverages of pre- and proto-historic China. *Proc Natl Acad Sci U S A* 101(51):17593–17598.
2. Sano C (2009) History of glutamate production. *American Journal of Clinical Nutrition*, p 728S–732S.
3. Werpy T, Petersen G (2004) Top value added chemicals from biomass. vol. I: Results of screening for potential candidates from sugars and synthesis gas. *US Dep energy*.
4. Bozell JJ, Petersen GR (2010) Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy's "Top 10" revisited. *Green Chem* 12:539–554.
5. Rowlands RT (1984) Industrial strain improvement: mutagenesis and random screening procedures. *Enzyme Microb Technol* 6(1):3–10.
6. Connor MR, Cann AF, Liao JC (2010) 3-Methyl-1-butanol production in Escherichia coli: random mutagenesis and two-phase fermentation. *Appl Microbiol Biotechnol* 86(4):1155–1164.
7. Ditchburn P, Giddings B, Macdonald KD (1974) Rapid Screening for the Isolation of Mutants of Aspergillus nidulans with Increased Penicillin Yields. *J Appl Bacteriol* 37(4):515–523.
8. Schrumpf B, Eggeling L, Sahm H (1992) Isolation and prominent characteristics of an L-lysine hyperproducing strain of Corynebacterium glutamicum. *Appl Microbiol Biotechnol* 37(5):566–571.
9. Teshiba S, Furuya A (1982) Mechanisms of 5'-inosinic acid accumulation by permeability mutants of brevibacterium ammoniagenes. I. genetical improvement of 5'-IMP productivity of a permeability mutant of B. ammoniagenes. *Agric Biol Chem* 46(9):2257–2263.
10. Cohen SN, Chang ACY, Boyer HW, Helling RB (1973) Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proc Natl Acad Sci U S A* 70(11):3240–3244.
11. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with

chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74(12):5463–5467.

12. Russell CB, Thaler DS, Dahlquist FW (1989) Chromosomal transformation of Escherichia coli recD strains with linearized plasmids. *J Bacteriol* 171(5):2609–2613.
13. Stephanopoulos G (1999) Metabolic Fluxes and Metabolic Engineering. *Metab Eng* 1(1):1–11.
14. Stephanopoulos G, Aristidou A, Nielsen J (1999) *Metabolic Engineering: Principles and Methodologies* (Academic Press).
15. Becker J, Zelder O, Häfner S, Schröder H, Wittmann C (2011) From zero to hero-Design-based systems metabolic engineering of Corynebacterium glutamicum for l-lysine production. *Metab Eng* 13(2):159–168.
16. Qian ZG, Xia XX, Lee SY (2009) Metabolic engineering of Escherichia coli for the production of putrescine: A four carbon diamine. *Biotechnol Bioeng* 104(4):651–662.
17. Lee J, et al. (2009) Metabolic engineering of a reduced-genome strain of Escherichia coli for L-threonine production. *Microb Cell Fact* 8(2):1–12.
18. Lee SJ, Song H, Lee SY (2006) Genome-Based Metabolic Engineering of Mannheimia succiniciproducens for Succinic Acid Production. *Appl Environ Microbiol* 72(3):1939–1948.
19. Ikeda M, Ohnishi J, Hayashi M, Mitsuhashi S (2006) A genome-based approach to create a minimally mutated Corynebacterium glutamicum strain for efficient L-lysine production. *J Ind Microbiol Biotechnol* 33(7):610–615.
20. Ajikumar PK, et al. (2010) Isoprenoid pathway optimization for taxol precursor overproduction in Escherichia coli. *Science* 330(6000):70–74.
21. Ro D-K, et al. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440(7086):940–943.
22. Mazumdar S, Clomburg JM, Gonzalez R (2010) Escherichia coli strains engineered for homofermentative production of D-lactic acid from glycerol. *Appl Environ Microbiol* 76(13):4327–4336.
23. Liu C, Wang Q, Xian M, Ding Y, Zhao G (2013) Dissection of

malonyl-coenzyme A reductase of Chloroflexus aurantiacus results in enzyme activity improvement. *PLoS One* 8(9):e75554.

24. Liu C, et al. (2015) Functional balance between enzymes in malonyl-CoA pathway for 3-hydroxypropionate biosynthesis. *Metab Eng* 34:104–111.
25. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* 97(12):6640–6645.
26. Baba T, et al. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2:2006.0008.
27. Li Y, et al. (2015) Metabolic engineering of Escherichia coli using CRISPR-Cas9 meditated genome editing. *Metab Eng* 31:13–21.
28. Jakočinas T, et al. (2015) Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. *Metab Eng* 28:213–222.
29. Li H, et al. (2016) CRISPR-Cas9 for the genome engineering of cyanobacteria and succinate production. *Metab Eng* 38(June):293–302.
30. San K-Y, et al. (2002) Metabolic Engineering through Cofactor Manipulation and Its Effects on Metabolic Flux Redistribution in Escherichia coli. *Metab Eng* 4(2):182–192.
31. Son HF, Park S, Yoo TH, Jung GY, Kim K-J (2017) Structural insights into the production of 3-hydroxypropionic acid by aldehyde dehydrogenase from Azospirillum brasilense. *Sci Rep* 7(February):46005.
32. Machado D, Herrgård MJ, Rocha I (2015) Modeling the Contribution of Allosteric Regulation for Flux Control in the Central Carbon Metabolism of E. coli. *Front Bioeng Biotechnol* 3:154.
33. Covert MW, Schilling CH, Palsson BØ (2001) Regulation of gene expression in flux balance models of metabolism. *J Theor Biol* 213(1):73–88.
34. Hamilton JJ, Dwivedi V, Reed JL (2013) Quantitative assessment of thermodynamic constraints on the solution space of genome-scale metabolic models. *Biophys J* 105(2):512–522.

35. Walsh K, Koshland DE (1984) The Determination of Flux through the Branch Point of Two Metabolic Cycles. *J Biol Chem* 259(15):9646–9655.

36. Walsh K, Koshland DE (1985) Branch point control by the phosphorylation state of isocitrate dehydrogenase: A quantitative examination of fluxes during a regulatory transition. *J Biol Chem* 260(14):8430–8437.

37. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28(1):56–59.

38. Schomburg I, et al. (2004) BRENDA, the enzyme database: updates and major new developments. *Nucleic Acids Res* 32:D431-4333.

39. Edwards JS (1999) Systems Properties of the Haemophilus influenzae Rd Metabolic Genotype. *J Biol Chem* 274(25):17410–17416.

40. Edwards JS, Palsson BØ (2000) The Escherichia coli MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proc Natl Acad Sci U S A* 97(10):5528–5533.

41. Forster J, Famili I, Palsson BØ, Nielsen J (2003) Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. *Genome Res* (13):244–253.

42. Feist AM, Palsson BØ (2008) The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. *Nat Biotechnol* 26(6):659–667.

43. Monk J, Nogales J, Palsson BØ (2014) Optimizing genome-scale network reconstructions. *Nat Biotechnol* 32(5):447–452.

44. Varma A, Palsson BØ (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Appl Environ Microbiol* 60(10):3724–3731.

45. Kauffman KJ, Prakash P, Edwards JS (2003) Advances in flux balance analysis. *Curr Opin Biotechnol* 14(5):491–496.

46. Orth JD, Thiele I, Palsson BØ (2010) What is flux balance analysis? *Nat Biotechnol* 28(3):245–248.

47. Chen X, Alonso AP, Allen DK, Reed JL, Shachar-Hill Y (2011) Synergy

between 13C-metabolic flux analysis and flux balance analysis for understanding metabolic adaption to anaerobiosis in E. coli. *Metab Eng* 13(1):38–48.

48. Segre D, Vitkup D, Church GM (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci U S A* 99(23):15112–15117.

49. Shinfuku Y, et al. (2009) Development and experimental verification of a genome-scale metabolic model for Corynebacterium glutamicum. *Microb Cell Fact* 8:43.

50. Zabriskie DW, Arcuri EJ (1986) Factors influencing productivity of fermentations employing recombinant microorganisms. *Enzyme Microb Technol* 8(12):706–717.

51. Schneider J, Eberhardt D, Wendisch VF (2012) Improving putrescine production by Corynebacterium glutamicum by fine-tuning ornithine transcarbamoylase activity using a plasmid addiction system. *Appl Microbiol Biotechnol* 95(1):169–178.

52. Fong SS, et al. (2005) In silico design and adaptive evolution of Escherichia coli for production of lactic acid. *Biotechnol Bioeng* 91(5):643–648.

53. Jantama K, et al. (2008) Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of Escherichia coli C that produce succinate and malate. *Biotechnol Bioeng* 99(5):1140–1153.

54. Burgard AP, Pharkya P, Maranas CD (2003) Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng* 84(6):647–657.

55. Ohno S, Shimizu H, Furusawa C (2014) FastPros: Screening of reaction knockout strategies for metabolic engineering. *Bioinformatics* 30(7):981–987.

56. Feist AM, et al. (2010) Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli. *Metab Eng* 12(3):173–186.

57. Lee SJ, et al. (2005) Metabolic engineering of Escherichia coli for enhanced production of succinic acid, based on genome comparison and

in silico gene knockout simulation. *Appl Environ Microbiol* 71(12):7880–7887.

58. Yim H, et al. (2011) Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. *Nat Chem Biol* 7(7):445–452.
59. Ng CY, Jung M, Lee J, Oh M-K (2012) Production of 2,3-butanediol in Saccharomyces cerevisiae by in silico aided metabolic engineering. *Microb Cell Fact* 11(1):68.
60. Scandalios JG (2002) Oxidative stress responses - what have genome-scale studies taught us? *Genome Biol* 3(7):1.
61. Zaldivar J, Martinez A, Ingram LO (1999) Effect of selected aldehydes on the growth and fermentation of ethanologenic Escherichia coli. *Biotechnol Bioeng* 65(1):24–33.
62. Mundhada H, et al. (2016) Increased production of L-serine in Escherichia coli through Adaptive Laboratory Evolution. *Metab Eng* 39:141–150.
63. Stadtman, E. R.; Cohen, G. N.; LeBras, Gisele; De Robichon-Szulmajster H (1961) Feed-back Inhibition and Repression of Aspartokinase Activity in Escherichia coli and Saccharomyces cerevisiae. *J Biol Chem* 236(7).
64. Horinouchi T, Sakai A, Kotani H, Tanabe K, Furusawa C (2017) Improvement of isopropanol tolerance of Escherichia coli using adaptive laboratory evolution and omics technologies. *J Biotechnol* 255:47–56.
65. Long CP, Gonzalez JE, Sandoval NR, Antoniewicz MR (2016) Characterization of physiological responses to 22 gene knockouts in Escherichia coli central carbon metabolism. *Metab Eng* 37:102–113.
66. Oberhardt MA, Palsson BØ, Papin JA (2009) Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* 5:320.
67. Feist AM, Palsson BØ (2008) The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. *Nat Biotechnol* 26(6):659–667.
68. McCloskey D, Palsson BØ, Feist AM (2013) Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli. *Mol Syst Biol* 9:661.

69. Alper H, Jin Y-S, Moxley JF, Stephanopoulos G (2005) Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli. *Metab Eng* 7(3):155–164.

70. Ohno S, Furusawa C, Shimizu H (2013) In silico screening of triple reaction knockout Escherichia coli strains for overproduction of useful metabolites. *J Biosci Bioeng* 115(2):221–228.

71. Chowdhury A, Zomorrodi AR, Maranas CD (2014) k-OptForce: Integrating Kinetics with Flux Balance Analysis for Strain Design. *PLoS Comput Biol* 10(2):e1003487.

72. Lu W, et al. (2015) Identification and elimination of metabolic bottlenecks in the quinone modification pathway for enhanced coenzyme Q10 production in Rhodobacter sphaeroides. *Metab Eng* 29:208–216.

73. Pitera DJ, Paddon CJ, Newman JD, Keasling JD (2007) Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli. *Metab Eng* 9(2):193–207.

74. Cremer J, Eggeling L, Sahm H (1991) Control of the lysine biosynthesis sequence in Corynebacterium glutamicum as analyzed by overexpression of the individual corresponding genes. *Appl Environ Microbiol* 57(6):1746–1752.

75. Niederberger P, Prasad R, Miozzari G, Kacser H (1992) A strategy for increasing an in vivo flux by genetic manipulations. The tryptophan system of yeast. *Biochem J* 287:473–479.

76. Teusink B, et al. (2000) Can yeast glycolysis be understood terms of vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem* 267(17):5313–5329.

77. Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M (2002) Dynamic modeling of the central carbon metabolism of Escherichia coli. *Biotechnol Bioeng* 79(1):53–73.

78. Rizk ML, Liao JC (2009) Ensemble modeling for aromatic production in Escherichia coli. *PLoS One* 4(9):e6903.

79. Khodayari A, Zomorrodi AR, Liao JC, Maranas CD (2014) A kinetic model of Escherichia coli core metabolism satisfying multiple sets of mutant flux data. *Metab Eng* 25:50–62.

80. Andreozzi S, et al. (2016) Identification of metabolic engineering targets for the enhancement of 1,4-butanediol production in recombinant E. coli using large-scale kinetic models. *Metab Eng* 35:148–159.

81. Feist AM, et al. (2007) A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 3:121.

82. Snoep JL, et al. (1993) Differences in sensitivity to NADH of purified pyruvate dehydrogenase complexes of Enterococcus faecalis, Lactococcus lactis, Azotobacter vinelandii and Escherichia coli: implications for their activity in vivo. *FEMS Microbiol Lett* 114(3):279–283.

83. Kim Y, Ingram LO, Shanmugam KT (2008) Dihydrolipoamide dehydrogenase mutation alters the NADH sensitivity of pyruvate dehydrogenase complex of Escherichia coli K-12. *J Bacteriol* 190(11):3851–3858.

84. Iuchi S, Lin EC (1988) arcA (dye), a global regulatory gene in Escherichia coli mediating repression of enzymes in aerobic pathways. *Proc Natl Acad Sci U S A* 85(6):1888–1892.

85. Cunningham L, Guest JR (1998) Transcription and transcript processing in the sdhCDAB-sucABCD operon of Escherichia coli. *Microbiology* 144:2113–2123.

86. Yu BJ, et al. (2006) sucAB and sucCD are mutually essential genes in Escherichia coli. *FEMS Microbiol Lett* 254(2):245–250.

87. Hua Q, Joyce AR, Fong SS, Palsson BØ (2006) Metabolic analysis of adaptive evolution for in silico-designed lactate-producing strains. *Biotechnol Bioeng* 95(5):992–1002.

88. Kochanowski K, et al. (2013) Functioning of a metabolic flux sensor in Escherichia coli. *Proc Natl Acad Sci U S A* 110(3):1130–1135.

89. Raj SM, Rathnasingh C, Jo J-E, Park S (2008) Production of 3-hydroxypropionic acid from glycerol by a novel recombinant Escherichia coli BL21 strain. *Process Biochem* 43(12):1440–1446.

90. Mori K, Toraya T (1999) A Reactivating Factor for Coenzyme B12-dependent Diol Dehydratase. *J Biol Chem* 274(6):3372–3377.

91. Toraya T, Shirakashi T, Kosuga T, Fukui S (1976) Substrate specificity of coenzyme B12-dependent diol dehydrase: Glycerol as both a good substrate and a potent inactivator. *Biochem Biophys Res Commun* 69(2):475–480.

92. Rathnasingh C, Raj SM, Jo J-E, Park S (2009) Development and evaluation of efficient recombinant Escherichia coli strains for the production of 3-hydroxypropionic acid from glycerol. *Biotechnol Bioeng* 104(4):729–739.

93. Luo LH, et al. (2013) Identification and characterization of Klebsiella pneumoniae aldehyde dehydrogenases increasing production of 3-hydroxypropionic acid from glycerol. *Bioprocess Biosyst Eng* 36(9):1319–1326.

94. Chu HS, et al. (2015) Metabolic engineering of 3-hydroxypropionic acid biosynthesis in Escherichia coli. *Biotechnol Bioeng* 112(2):356–364.

95. Seok Jung W, Ho Kang J, Su Chu H, Suk Choi I, Myung Cho K (2014) Elevated production of 3-hydroxypropionic acid by metabolic engineering of the glycerol metabolism in Escherichia coli. *Metab Eng* 23:116–122.

96. Bell RM, Cronan JE (1975) Mutants of Escherichia coli defective in membrane phospholipid synthesis. Phenotypic suppression of sn-glycerol-3-phosphate acyltransferase Km mutants by loss of feedback inhibition of the biosynthetic sn-glycerol-3-phosphate dehydrogenase. *J Biol Chem* 250(18):7153–7158.

97. Perrenoud A, Sauer U, Perrenoud A, Sauer U (2005) Impact of Global Transcriptional Regulation by ArcA , ArcB , Cra , Crp , Cya , Fnr , and Mlc on Glucose Catabolism in Escherichia coli. *J Bacteriol* 187(9):3171–3179.

98. Ogasawara H, Ishida Y, Yamada K, Yamamoto K, Ishihama A (2007) PdhR (pyruvate dehydrogenase complex regulator) controls the respiratory electron transport system in Escherichia coli. *J Bacteriol* 189(15):5534–5541.

99. Kale S, Arjunan P, Furey W, Jordan F (2007) A dynamic loop at the active center of the Escherichia coli pyruvate dehydrogenase complex E1 component modulates substrate utilization and chemical

communication with the E2 component. *J Biol Chem* 282(38):28106–28116.

100. Corwin LM, Fanning GR (1967) Studies of parameters affecting the allosteric nature of phosphoenolpyruvate carboxylase of Escherichia coli. *J Biol Chem* 243(12):3517–3525.
101. Lin H, Vadali R V, Bennett GN, San K-Y (2004) Increasing the acetyl-CoA pool in the presence of overexpressed phosphoenolpyruvate carboxylase or pyruvate carboxylase enhances succinate production in Escherichia coli. *Biotechnol Prog* 20(5):1599–1604.
102. Mo ML, Palsson BØ, Herrgård MJ (2009) Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC Syst Biol* 3(1):37.
103. Yoshikawa K, et al. (2011) Reconstruction and verification of a genome-scale metabolic model for Synechocystis sp. PCC6803. *Appl Microbiol Biotechnol* 92(2):347–358.
104. Flahaut NAL, et al. (2013) Genome-scale metabolic model for Lactococcus lactis MG1363 and its application to the analysis of flavor formation. *Appl Microbiol Biotechnol* 97(19):8729–8739.
105. Valdehuesa KNG, et al. (2013) Recent advances in the metabolic engineering of microorganisms for the production of 3-hydroxypropionic acid as C3 platform chemical. *Appl Microbiol Biotechnol* 97(8):3309–3321.
106. Kumar V, Ashok S, Park S (2013) Recent advances in biological production of 3-hydroxypropionic acid. *Biotechnol Adv* 31(6):945–961.
107. Wang Q, et al. (2013) Biosynthesis of poly(3-hydroxypropionate) from glycerol by recombinant Escherichia coli. *Bioresour Technol* 131:548–551.
108. Andreessen B, Steinbüchel A (2010) Biosynthesis and biodegradation of 3-hydroxypropionate-containing polyesters. *Appl Environ Microbiol* 76(15):4919–4925.
109. da Silva GP, Mack M, Contiero J (2009) Glycerol: a promising and abundant carbon source for industrial microbiology. *Biotechnol Adv* 27(1):30–39.
110. Ito T, Nakashimada Y, Senba K, Matsui T, Nishio N (2005) Hydrogen

and ethanol production from glycerol-containing wastes discharged after biodiesel manufacturing process. *J Biosci Bioeng* 100(3):260–265.

111. Mattam AJ, Clomburg JM, Gonzalez R, Yazdani SS (2013) Fermentation of glycerol and production of valuable chemical and biofuel molecules. *Biotechnol Lett* 35(6):831–842.
112. Ashok S, Raj SM, Rathnasingh C, Park S (2011) Development of recombinant Klebsiella pneumoniae ΔdhaT strain for the co-production of 3-hydroxypropionic acid and 1,3-propanediol from glycerol. *Appl Microbiol Biotechnol* 90(4):1253–1265.
113. Ashok S, et al. (2013) Effect of puuC overexpression and nitrate addition on glycerol metabolism and anaerobic 3-hydroxypropionic acid production in recombinant Klebsiella pneumoniae ΔglpKΔdhaT. *Metab Eng* 15:10–24.
114. Ashok S, et al. (2013) Production of 3-hydroxypropionic acid from glycerol by recombinant Klebsiella pneumoniae ΔdhaTΔyqhD which can produce vitamin B(12) naturally. *Biotechnol Bioeng* 110(2):511–524.
115. Zhou S, Catherine C, Rathnasingh C, Somasundar A, Park S (2013) Production of 3-hydroxypropionic acid from glycerol by recombinant Pseudomonas denitrificans. *Biotechnol Bioeng* 110(12):3177–3187.
116. Zhou S, Ashok S, Ko Y, Kim D-M, Park S (2014) Development of a deletion mutant of Pseudomonas denitrificans that does not degrade 3-hydroxypropionic acid. *Appl Microbiol Biotechnol* 98(10):4389–4398.
117. Raj SM, Rathnasingh C, Jung W-C, Park S (2009) Effect of process parameters on 3-hydroxypropionic acid production from glycerol using a recombinant Escherichia coli. *Appl Microbiol Biotechnol* 84(4):649–657.
118. Kim K, Kim S-K, Park Y-C, Seo J-H (2014) Enhanced production of 3-hydroxypropionic acid from glycerol by modulation of glycerol metabolism in recombinant Escherichia coli. *Bioresour Technol* 156:170–175.
119. Ida Y, Hirasawa T, Furusawa C, Shimizu H (2013) Utilization of Saccharomyces cerevisiae recombinant strain incapable of both ethanol and glycerol biosynthesis for anaerobic bioproduction. *Appl Microbiol*

*Biotechnol* 97(11):4811–4819.

120. Huang Y, Li Z, Shimizu K, Ye Q (2013) Co-production of 3-hydroxypropionic acid and 1,3-propanediol by Klebseilla pneumoniae expressing aldH under microaerobic conditions.pdf. 505–512.
121. Oberhardt MA, Palsson BØ, Papin JA (2009) Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* 5(320):1–15.
122. Price ND, Reed JL, Palsson BØ (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat Rev Microbiol* 2(11):886–897.
123. Thomason LC, Costantino N, Court DL (2007) E. coli genome manipulation by P1 transduction. *Curr Protoc Mol Biol* Chapter 1:Unit 1.17.
124. Nakahigashi K, et al. (2009) Systematic phenome analysis of Escherichia coli multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Mol Syst Biol* 5(306):1–14.
125. Gilbert RP, Brandt RB (1975) Spectrophotometric determination of methyl glyoxal with 2,4-dinitrophenylhydrazine. *Anal Chem* 47(14):2418–2422.
126. Tang X, Tan Y, Zhu H, Zhao K, Shen W (2009) Microbial conversion of glycerol to 1,3-propanediol by an engineered strain of escherichia coli. *Appl Environ Microbiol* 75(6):1628–1634.
127. Cooper R (1984) Metabolism of methylglyoxal in microorganisms. *Annu Rev Microbiol* 38:49–68.
128. Schuetz R, Kuepfer L, Sauer U (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Mol Syst Biol* 3(119):1–15.
129. Tötemeyer S, Booth N a, Nichols WW, Dunbar B, Booth IR (1998) From famine to feast: the role of methylglyoxal production in Escherichia coli. *Mol Microbiol* 27(3):553–562.
130. Durnin G, et al. (2009) Understanding and harnessing the microaerobic metabolism of glycerol in Escherichia coli. *Biotechnol Bioeng* 103(1):148–161.
131. Jarboe LR (2011) YqhD: a broad-substrate range aldehyde reductase

with various applications in production of biorenewable fuels and chemicals. *Appl Microbiol Biotechnol* 89(2):249–257.

132. Rujananon R, Prasertsan P, Phongdara A (2014) Biosynthesis of 1,3-propanediol from recombinant E. coli by optimization process using pure and crude glycerol as a sole carbon source under two-phase fermentation system. *World J Microbiol Biotechnol* 30(4):1359–1368.
133. Ozyamak E, De Almeida C, De Moura APS, Miller S, Booth IR (2013) Integrated stress response of Escherichia coli to methylglyoxal: Transcriptional readthrough from the nemRA operon enhances protection through increased expression of glyoxalase I. *Mol Microbiol* 88(5):936–950.
134. Millard CS, Chao YP, Liao JC, Donnelly MI (1996) Enhanced production of succinic acid by overexpression of phosphoenolpyruvate carboxylase in Escherichia coli. *Appl Environ Microbiol* 62(5):1808–1810.
135. Zhang X, Shanmugam KT, Ingram LO (2010) Fermentation of glycerol to succinate by metabolically engineered strains of escherichia coli. *Appl Environ Microbiol* 76(8):2397–2401.
136. Okino S, et al. (2008) An efficient succinic acid production process in a metabolically engineered Corynebacterium glutamicum strain. *Appl Microbiol Biotechnol* 81(3):459–464.
137. Gao C, et al. (2016) Robust succinic acid production from crude glycerol using engineered Yarrowia lipolytica. *Biotechnol Biofuels* 9(1):179.
138. Charusanti P, et al. (2010) Genetic basis of growth adaptation of Escherichia coli after deletion of pgi, a major metabolic gene. *PLoS Genet* 6(11):e1001186.
139. Soellner S, Rahnert M, Siemann-Herzberg M, Takors R, Altenbuchner J (2013) Evolution of pyruvate kinase-deficient Escherichia coli mutants enables glycerol-based cell growth and succinate production. *J Appl Microbiol* 115(6):1368–1378.
140. Lenski RE, et al. (1998) Evolution of competitive fitness in experimental populations of E. coli: What makes one genotype a better competitor than another? *Antonie Van Leeuwenhoek* 73(1):35–47.
141. Sandberg TE, et al. (2014) Evolution of Escherichia coli to 42 °C and

Subsequent Genetic Engineering Reveals Adaptive Mechanisms and Novel Mutations. *Mol Biol Evol* 31(10):2647–2662.

142. Herring CD, et al. (2006) Comparative genome sequencing of Escherichia coli allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* 38(12):1406–1412.
143. Lee D-H, Palsson BØ (2010) Adaptive evolution of Escherichia coli K-12 MG1655 during growth on a Nonnative carbon source, L-1,2-propanediol. *Appl Environ Microbiol* 76(13):4158–4168.
144. Fong SS, Nanchen A, Palsson BO, Sauer U (2006) Latent pathway activation and increased pathway capacity enable Escherichia coli adaptation to loss of key metabolic enzymes. *J Biol Chem* 281(12):8024–8033.
145. Ibarra RU, Edwards JS, Palsson BO (2002) Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* 420(6912):186–189.
146. Fong SS, Palsson BØ (2004) Metabolic gene-deletion strains of Escherichia coli evolve to computationally predicted growth phenotypes. *Nat Genet* 36(10):1056–1058.
147. Cheng K-K, et al. (2014) Global metabolic network reorganization by adaptive mutations allows fast growth of Escherichia coli on glycerol. *Nat Commun* 5:3233.
148. Nyerges Á, et al. (2016) A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc Natl Acad Sci U S A* 113(9):2502–2507.
149. Deatherage DE, Barrick JE (2014) Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol* 1151:165–188.
150. Bonde MT, et al. (2014) MODEST: A web-based design tool for oligonucleotide-mediated genome engineering and recombineering. *Nucleic Acids Res* 42(Web server issue):W408–W415.
151. Kai Y, et al. (1999) Three-dimensional structure of phosphoenolpyruvate carboxylase : A proposed mechanism for allosteric inhibition. *Proc Natl Acad Sci U S A* 96(3):823–828.
152. Chen Z, Bommareddy RR, Frank D, Rappert S, Zeng AP (2014)

Deregulation of feedback inhibition of phosphoenolpyruvate carboxylase for improved lysine production in Corynebacterium glutamicum. *Appl Environ Microbiol* 80(4):1388–1393.

153. Nishino S, Okahashi N, Matsuda F, Shimizu H (2015) Absolute quantitation of glycolytic intermediates reveals thermodynamic shifts in Saccharomyces cerevisiae strains lacking PFK1 or ZWF1 genes. *J Biosci Bioeng* 120(3):280–286.
154. Schellenberger J, et al. (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat Protoc* 6(9):1290–1307.
155. Applebee MK, Joyce AR, Conrad TM, Pettigrew DW, Palsson B (2011) Functional and metabolic effects of adaptive glycerol kinase (GLPK) mutants in Escherichia coli. *J Biol Chem* 286(26):23150–23159.
156. Emmerling M, Bailey JE, Sauer U (1999) Glucose catabolism of Escherichia coli strains with increased activity and altered regulation of key glycolytic enzymes. *Metab Eng* 1(2):117–127.
157. Emmerling M, Bailey JE, Sauer U (2000) Altered regulation of pyruvate kinase or co-overexpression of phosphofructokinase increases glycolytic fluxes in resting Escherichia coli. *Biotechnol Bioeng* 67(5):623–627.
158. Nakamura T, Minoguchi S, Izui K (1996) Purification and characterization of recombinant phosphoenolpyruvate carboxylase of Thermus sp. *J Biochem* 120(3):518–524.
159. Takeya M, Hirai MY, Osanai T (2017) Allosteric Inhibition of Phosphoenolpyruvate Carboxylases is Determined by a Single Amino Acid Residue in Cyanobacteria. *Sci Rep* 7:41080–41088.
160. von Kamp A, Klamt S (2017) Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms. *Nat Commun* 8:15956.
161. Ng CY, Farasat I, Maranas CD, Salis HM (2015) Rational Design of a Synthetic Entner-Doudoroff Pathway for Improved and Controllable NADPH Regeneration. *Metab Eng* 29:86–96.

## Appendix

List of abbreviations

| Abbreviation | Full name |
|---|---|
| ALE | Adaptive laboratory evolution |
| CDW | Cell dry weight |
| C-mol/C-mol | Carbon molar yield |
| FBA | Flux balance analysis |
| GLYK | Glycerol kinase |
| GUR | Glycerol uptake rate |
| MAGE | Multiplexed automated genome engineering |
| OD | Optical density |
| OUR | Oxygen uptake rate |
| PDB | Protein data bank |
| PDH | Pyruvate dehydrogenase |
| Ppc | Phosphoenolpyruvate carboxylase |
| SUR | Substrate uptake rate |

List of metabolite abbreviations

| Abbreviation | Metabolite name |
|---|---|
| 1,3PDO | 1,3-Propanediol |
| 1,4BDO | 1,4-Butanediol |
| 3HP | 3-hydroxypropionic acid |
| 3HPA | 3-hydroxypropionic aldehyde |
| 3PG | 3-phosphoglycerate |
| 6PG | 6-phospho-D-gluconate |
| AcCoA | Acetyl-coenzyme A |
| Aco | Aconitate |
| ADP | Adenosine diphosphate |
| Ala | Alanine |
| AMP | Adenosine monophosphate |
| Arg | Arginine |
| Asn | Asparagine |
| Asp | Aspartic Acid |
| ATP | Adenosine triphosphate |
| Cit | Citrate |
| CO2 | Carbo dioxide |
| CoA | Coenzyme A |
| Cys | Cysteine |
| DHA | Dihydroxyacetone |
| DHAP | Dihydroxyacetone phosphate |
| E4P | D-Erythrose-4-phosphate |
| F6P | D-Fructose-6-phosphate |
| FBP | D-Fructose-1,6-bisphosphate |
| Fum | Fumarate |
| G6P | D-Glucose-6-phosphate |
| GAP | Glyceraldehyde-3-phophate |
| Gln | Glutamine |
| Glu | Glutamic Acid |
| Gly | Glycine |
| Glyc3P | Glycerol-3-phosphate |
| Glyox | Glyoxylic acid |
| His | Histidine |

| | |
|---|---|
| Ile | Isoleucine |
| IsoCit | Isocitrate |
| KDPG | 2-keto-3-deoxy-6-phosphogluconate |
| Leu | Leucine |
| Lys | Lysine |
| Mal | Malate |
| Met | Methionine |
| MGO | Methylglyoxal |
| $NAD^+$ | Nicotinamide adenine dinucleotide (oxidized) |
| NADH | Nicotinamide adenine dinucleotide (reduced) |
| $NADP^+$ | Nicotinamide adenine dinucleotide phosphate (oxidized) |
| NADPH | Nicotinamide adenine dinucleotide phosphate (reduced) |
| $NH_4$ | Ammonia |
| Oxa | Oxaloacetate |
| PEP | Phosphoenolpyruvate |
| Phe | Phenylalanine |
| Pro | Proline |
| Pyr | Pyruvate |
| Q | Quinone (oxidized) |
| $QH_2$ | Quinone (reduced) |
| Ru5P | D-Ribulose-5-phosphate |
| S7P | D-Sedoheptulose-7-phosphate |
| Ser | Serine |
| Suc | Succinate |
| SucCoA | Succinyl-coenzyme A |
| Thr | Threonine |
| Trp | Tryptophan |
| Tyr | Tyrosine |
| Val | Valine |
| Xu5P | D-Xylose-5-phophate |
| αKG | α-Ketoglutarate |