Metabolic Engineering 24 (2014) 117–128

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

Optimal cofactor swapping can increase the theoretical yield for chemical production in *Escherichia coli* and *Saccharomyces cerevisiae*

Zachary A. King [a], Adam M. Feist [a,b,*]

[a] Department of Bioengineering, University of California, 9500 Gilman Drive #0412, San Diego, La Jolla, CA 92093-0412, USA
[b] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Lyngby, Denmark

ARTICLE INFO

Article history:
Received 20 November 2013
Received in revised form
6 March 2014
Accepted 6 May 2014
Available online 14 May 2014

Keywords:
Cofactor balancing
Escherichia coli
Saccharomyces cerevisiae
Constraint-based modeling
MILP
Theoretical yield
Metabolic engineering

ABSTRACT

Maintaining cofactor balance is a critical function in microorganisms, but often the native cofactor balance does not match the needs of an engineered metabolic flux state. Here, an optimization procedure is utilized to identify optimal cofactor-specificity "swaps" for oxidoreductase enzymes utilizing NAD(H) or NADP(H) in the genome-scale metabolic models of *Escherichia coli* and *Saccharomyces cerevisiae*. The theoretical yields of all native carbon-containing molecules are considered, as well as theoretical yields of twelve heterologous production pathways in *E. coli*. Swapping the cofactor specificity of central metabolic enzymes (especially GAPD and ALCD2x) is shown to increase NADPH production and increase theoretical yields for native products in *E. coli* and yeast—including L-aspartate, L-lysine, L-isoleucine, L-proline, L-serine, and putrescine—and non-native products in *E. coli*—including 1,3-propanediol, 3-hydroxybutyrate, 3-hydroxypropanoate, 3-hydroxyvalerate, and styrene.

1. Introduction

Division of the roles of currency metabolites is a well-conserved feature of metabolism in microorganisms. In *Escherichia coli* and *Saccharomyces cerevisiae*, the cofactors NAD(H) and NADP(H) are both responsible for transferring reducing equivalents between metabolic subsystems. NAD(H) is primarily generated by glycolytic enzymes and transfers reducing equivalents to the electron transport chain or to fermentation products (Russell and Cook, 1995; Gottschalk, 1986; Sauer et al., 2004). NADP(H) is produced primarily by the pentose phosphate pathway and transhydrogenase enzymes, and it transfers reducing equivalents to provide energy for biosynthesis (Gottschalk, 1986; Sauer et al., 2004). While this separation is not complete (for example, fungi utilize NADPH for pentose catabolism, Verho et al., 2003), the functional separation of these electron carriers and the specificity of oxidoreductase enzymes to a specific electron carrier allow the cell to precisely partition resources between ATP production and anabolism.

When growing in a steady state, microorganisms coordinate the production of reduced cofactors to match cofactor consumption, and their metabolic network structures and regulatory systems are organized to carry out this balancing act in common environments (Sauer et al., 2004; Lunzer et al., 2005; Zhu et al., 2005). Consequently, the cofactor balance in microorganisms is poorly optimized for many synthetic cellular objectives (Ghosh et al., 2011; Hyung Lim et al., 2013; Jan et al., 2013). Thus, one should consider how cofactor balance can be optimized when formulating new cellular objectives for metabolic engineering and synthetic biology. Cofactor balance optimization is especially important for introducing non-native production pathways that are driven by cofactor concentration (Shen et al., 2011).

A number of experimental methods have been developed to increase the availability of the reduced cofactors NADH and NADPH to enzymes in production pathways that are cofactor-driven, and thereby increase yield of high-value byproducts (reviewed by Lee et al., 2013). One strategy to increase cofactor availability is overexpression of genes that generate cofactor-producing enzymes. Overexpression of NADH-producing formate dehydrogenase (*fdh1* from *Candida boidinii*) in *E. coli* was shown to increase production of ethanol during anaerobic fermentation and to cause production of fermentation byproducts during aerobic growth (Berríos-Rivera et al., 2002, 2004, 2002b, 2002a). Similarly, the manipulation of transhydrogenase enzymes in *E. coli* can shift byproduct yield. The overexpression of *sthA*, which encodes the soluble transhydrogenase enzyme, was shown to increase yield of both (S)-2 chloropropionate and poly(3-hydroxybutyrate), two byproducts produced in *E. coli* by anabolic reactions that utilize NADPH (Sanchez et al., 2006; Jan et al., 2013). For (S)-2 chloropropionate production, the deletion of *pntAB*,

* Corresponding author at: Department of Bioengineering, University of California, 9500 Gilman Drive #0412, San Diego, La Jolla, CA 92093-0412, USA. Fax: 858 822 3120.
E-mail address: afeist@ucsd.edu (A.M. Feist).

http://dx.doi.org/10.1016/j.ymben.2014.05.009
1096-7176/

which encodes the membrane-bound transhydrogenase enzyme, also increased product yield (Jan et al., 2013).

A second strategy to increase cofactor availability is to replace the native enzyme with a non-native oxidoreductase with specificity for the opposite cofactor. For example, the NAD(H)-dependent glyceraldehyde-3-phosphate dehydrogenase (GAPD) in *E. coli* (encoded by the gene *gapA*) was replaced with the NADP(H)-dependent GAPD from *Clostridium acetobutylicum* (encoded by the gene *gapC*) to increase the production of lycopene and to increase the NADPH yield to drive a bioprocessing reaction (cyclohexanone to ε-caprolactone) where *E. coli* acts only as a source of reducing equivalents (Martínez et al., 2008). In another study, the native NAD(H)-dependent GAPD of *S. cerevisiae* (encoded by the genes *TDH1–3*) was supplemented with the NADP(H)-dependent GAPD from *Kluyveromyces lactis* (encoded by the gene *GDP1*) to increase fermentation of D-xylose to ethanol (Verho et al., 2003). These studies show that experimental implementation of such cofactor "swaps" is feasible and can result in promising increases in product yield. However, the question that remains to be answered is, which enzymes should be modified for maximum yield?

Computational studies have also investigated cofactor balancing. Most studies to date have utilized constraint-based modeling, which represents the metabolic network by formulating the stoichiometry of metabolic reactions as a linear system of equations. Thermodynamic constraints (e.g., reaction irreversibility) and environmental parameters (e.g., nutrient availability) can be included in the formulation, and, by assuming that the system is in a mass-balanced steady state, linear optimization techniques can be used to identify optimal metabolic flux states and modifications in well-understood organisms (Price et al., 2004). In a previous study, the authors reported the development of a bilevel optimization method called OptSwap to identify growth-coupled designs using modifications of oxidoreductase specificity and knockouts (King and Feist, 2013). Similarly, Chung et al. (2013) presented a method called cofactor modification analysis (CMA) which optimized modifications of oxidoreductase specificity to improve the yield of terpenoids in yeast, and Lakshmanan et al. (2013) used the method to identify growth-coupled bioprocessing designs. Ghosh et al. (2011) used constraint-based modeling to analyze cofactor balancing for the specific case of yeast producing ethanol from L-arabinose and D-xylose. Chin et al. (2009) utilized the constraint-based model of *E. coli* to calculate theoretical yields of xylitol with various knockouts that affected cofactor balance, but they did not report a strategy to improve the cofactor balance. Despite the success of these targeted studies, a comprehensive analysis of the effects of changing cofactor specificity on a system-wide scale does not exist.

In this work, constraint-based modeling is utilized to identify optimal cofactor-specificity swaps for increasing theoretical yield in the genome-scale metabolic models of *E. coli* and *S. cerevisiae*. This work presents a global analysis of cofactor swapping for a large number of products across two important production organisms, and the optimizations identify the minimal cofactor swaps necessary to maximize theoretical yield in the metabolic network.

2. Methods

2.1. Models and parameters

The *i*JO1366 metabolic reconstruction of *E. coli* K-12 MG1655 (Orth et al., 2011) and the *i*MM904 metabolic reconstruction of *S. cerevisiae* (Mo et al., 2009) were used for the simulations in this work. Flux balance analysis (FBA) (Kauffman et al., 2003) and parsimonious flux balance analysis (pFBA) (Lewis et al., 2010) were implemented in MATLAB as reported. As described previously, the *i*JO1366 oxidative stress reactions CAT, SPODM, and SPODMpp and the FHL reaction were constrained to zero (Orth et al., 2011), and the *i*JO1366 POR5 (pyruvate:ferredoxin oxidoreductase) reaction was made irreversible, as supported by biochemical data (Blaschkowski et al., 1982). In *i*MM904 simulations, free exchange of six sterols and fatty acids—ergosterol, zymosterol, hexadecenoate, and octadecanoate (saturated, monounsaturated, and polyunsaturated)—was allowed under anaerobic conditions, as reported by Mo et al. (2009).

For cofactor swap optimizations, the substrate uptake rates (SURs) for the solitary carbon substrates in each simulation were constrained to a maximum uptake rate of 10 mmol gDW^{-1} h^{-1}. For aerobic simulations, the oxygen uptake rate was set to a maximum of 10 mmol gDW^{-1} h^{-1}. For these simulations, the minimum flux through the biomass objective function was set to 0.1 h^{-1}. For simulations with *S. cerevisiae* under anaerobic conditions with D-xylose substrate and *E. coli* under anaerobic conditions with glycerol substrate, the *in silico* growth rate was less than 0.1 h^{-1}. To explore the effect of cofactor swaps in these cases, the minimum biomass requirement was set to 10% of the maximum growth rate.

For substrate uptake, only native gene content was considered. For growth of *S. cerevisiae* on D-xylose, the preexisting *i*MM904 model includes a D-xylose catabolism pathway consisting of xylose reductase (XR, EC 1.1.1.307) and xylitol dehydrogenase (XDH, EC 1.1.1.10), which were included in the model based on annotation for the genes GRE3 and XYL2, respectively (Mo et al., 2009). However, under experimental conditions, D-xylose catabolism in yeast cannot support growth, and recombinant XR and XDH are necessary for fermentation of D-xylose (Bettiga et al., 2008; Ghosh et al., 2011; Bengtsson et al., 2009). Therefore, these reactions were considered to be native in the simulations, but their usage requires heterologous gene expression. Xylose isomerase (XI, EC 5.3.1.5), an alternative enzyme for D-xylose uptake which is not in *i*MM904, was also considered in the analysis of *S. cerevisiae* xylose catabolism (Bettiga et al., 2008).

Theoretical maximum yield ($Y_{P/S}$) is reported as the percentage of carbon consumed during substrate uptake that is converted to production of the target byproduct, subject to a minimum growth rate requirement at steady state (Feist et al., 2010):

$$Y_{P/S} = \frac{\text{byproduct production rate (mmol carbon gDW}^{-1}\text{ h}^{-1})}{\text{substrate uptake rate (mmol carbon gDW}^{-1}\text{ h}^{-1})}$$

2.2. Non-native pathways

To simulate cofactor swapping in realistic production pathways for non-native compounds, a literature search was performed to identify the most recent and successful experimentally-validated strain designs for production of non-native compounds in *E. coli*. Pathways were reconstructed by creating *in silico* reactions corresponding to the genes used in these experiments (Table 3, Supplementary Tables). *In silico* cofactor specificities were determined based on reports in the literature. Transport was assumed to be non-energy-coupled unless otherwise specified in the *i*JO1366 reconstruction or in the literature.

2.3. Selection of the reaction sets for cofactor-specificity swaps

The sets of oxidoreductase reactions available for modification during optimizations were determined by locating central, high-flux oxidoreductase reactions, as previously reported (King and Feist, 2013). In the metabolic models *i*JO1366 and

*i*MM904, all reactions utilizing NAD(H) or NADP(H) as a substrate were located. The reactions were sorted by flux magnitude after pFBA optimization for flux through the biomass objective function under conditions of aerobic and anaerobic growth on glucose and D-xylose minimal media. The reactions with highest flux under all conditions were selected. Lactate dehydrogenase, malic enzymes, and lactaldehyde dehydrogenase were added to the set for *E. coli* based on interest in the literature (Stols and Donnelly, 1997; Wang et al., 2011; Zhang et al., 2007). Under anaerobic conditions, the NADH:oxidoreductase I reaction in *i*JO1366 was removed from the pool of enzymes during model reduction. Thus, 21 oxidoreductase reactions were chosen for *E. coli* under anaerobic conditions and 22 oxidoreductase reactions under aerobic conditions (Table 1), and 22 reactions were chosen for *S. cerevisiae* under both aerobic and anaerobic conditions (Table 2).

2.4. MILP formulation

A mixed integer linear programming (MILP) problem was formulated to find the set of cofactor-specificity swaps that maximize the theoretical yield of chemical production with a minimum biomass requirement (Fig. 1). The MILP formulation is functionally identical to the one used by Chung et al. (2013), but here it is implemented differently. As described in our previous report, the non-native oxidoreductase enzymes are added to the system and coupled so that either the native enzyme or the non-native enzyme is active (King and Feist, 2013). The final formulation of the MILP problem can be stated follows, where S is the stoichiometric matrix, I is the set of metabolites, J is the set of reactions, D is the set of oxidoreductase reactions, s and t are decision variables defining the state of each swap, and L is the maximum number of swaps allowed:

$$\begin{aligned}
&\text{maximize} && v_{chemical} && \\
&\text{subject to} && \sum_{j \in J} S_{ij} v_j = 0 && \forall i \in I \\
& && LB_j \le v_j \le UB_j && \forall j \in J \\
& && v_{biomass} \ge min_biomass && \\
& && s_d \in \{0,1\} && \forall d \in D \\
& && t_d \in \{0,1\} && \forall d \in D \\
& && s_d + t_d = 1 && \forall d \in D \\
& && s_d LB_{x(d)} \le v_{x(d)} \le s_d UB_{x(d)} && \forall d \in D \\
& && t_d LB_{y(d)} \le v_{y(d)} \le t_d UB_{y(d)} && \forall d \in D \\
& && \sum_{d \in D} (1 - s_d) \le L &&
\end{aligned}$$

2.5. Non-unique solutions

In many cases, the solution to the MILP problem was non-unique. To investigate the diversity of cofactor swaps that could produce the same results, an exhaustive search method was utilized. For each solution, all swaps that improved maximum theoretical yield to within 99% of the optimal solution were found. These were identified as solution "groups" (Figs. 2–4). The same procedure was repeated to investigate secondary swaps, after first swapping a reaction from the first-swap group. Similar results could be achieved utilizing MILP solvers that enumerate solutions, such as the Solution Pool feature in CPLEX 11+ (IBM, Armonk, NY, USA).

Table 1
Oxidoreductase enzymes in the *E. coli* *i*JO1366 model chosen for cofactor balance optimizations.

Key oxidoreductase enzymes in *E. coli*	Gene name	Model reaction abbreviations	Native electron carrier in *E. coli*	pFBA, relative flux scaled to carbon uptake[a]			
				Glucose aerobic	D-xylose aerobic	Glucose anaerobic	D-xylose anaerobic
NADH:ubiquinone oxidoreductase I	*nuoF, nuoAC, nuoE, nuoGN*	NADH16pp	NADH	1.98	1.98	0.0	0.0
Glyceraldehyde-3-phosphate dehydrogenase	*gapA*	GAPD	NADH	1.68	1.38	1.89	1.59
Glucose-6-phosphate dehydrogenase	*zwf*	G6PDH2r	NADPH	0.45	0.40	0.15	0.11
6-Phosphogluconate dehydrogenase	*gnd*	GND	NADPH	0.45	0.40	0.15	0.11
Pyruvate dehydrogenase	*lpd, aceE, aceF*	PDH	NADH	0.37	0.66	0.0	0.0
FAD reductase	*fre*	FADRx	NADH	0.16	0.14	0.06	0.04
Phosphoglycerate dehydrogenase	*serA*	PGCD	NADH	0.12	0.11	0.04	0.03
Isocitrate dehydrogenase	*icd*	ICDHyr	NADPH	0.08	0.07	0.03	0.02
Aspartate-semialdehyde dehydrogenase	*asd*	ASAD	NADPH	0.08	0.07	0.03	0.02
Acetohydroxy acid isomeroreductase (2-Acetolactate)	*mdh*	MDH	NADH	0.08	0.07	0.02	0.01
Methylene tetrahydrofolate dehydrogenase	*folD*	MTHFD	NADPH	0.06	0.06	0.02	0.02
Acetohydroxy acid isomeroreductase	*ilvC*	KARA1	NADPH	0.06	0.06	0.02	0.02
Homoserine dehydrogenase	*metL* OR *thrA*	HSDy	NADPH	0.05	0.05	0.02	0.01
3-Isopropylmalate dehydrogenase	*leuB*	IPMD	NADH	0.03	0.03	0.01	0.01
Shikimate dehydrogenase	*aroE*	SHK3Dr	NADPH	0.03	0.02	0.01	0.01
Dihydrodipicolinate reductase	*dapB*	DHDPRy	NADPH	0.03	0.02	0.01	0.01
Acetaldehyde dehydrogenase	*adhE* OR *mhpF*	ACALD	NADH	0.0	0.0	0.93	0.78
Alcohol dehydrogenase (ethanol)	*adhP* OR *adhE*	ALCD2x	NADH	0.0	0.0	0.93	0.78
L-1,2-Propanediol oxidoreductase	*fucO*	LCARR	NADH	0.0	0.0	0.0	0.0
D-Lactate dehydrogenase	*ldhA*	LDH_D	NADH	0.0	0.0	0.0	0.0
Malic enzyme (NADH)	*maeA*	ME1	NADH	0.0	0.0	0.0	0.0
Malic enzyme (NADPH)	*maeB*	ME2	NADPH	0.0	0.0	0.0	0.0

[a] The flux through each reaction for maximum growth simulation using pFBA, relative to carbon uptake rate.

Table 2
Oxidoreductase enzymes in the *i*MM904 model of *S. cerevisiae* chosen for cofactor balance optimizations.

Key oxidoreductase enzymes in *S. cerevisiae*	Gene symbol	Model reaction abbreviations	Native electron carrier in *S. cerevisiae*	pFBA, relative flux scaled to carbon uptake[a]			
				Glucose aerobic	D-xylose aerobic	Glucose anaerobic	D-xylose anaerobic
Glyceraldehyde-3-phosphate dehydrogenase	TDH1 or TDH2 or TDH3	GAPD	NADH	1.68	1.20	1.88	1.63
Alcohol dehydrogenase (ethanol)	ADH1 or ADH4 or ADH5 or SFA1	ALCD2x[b]	NADH	1.01	0.78	1.67	1.57
NADH dehydrogenase	NDE1 or NDE2	NADH2-u6cm	NADH	0.69	1.44	0.03	0.01
Isocitrate dehydrogenase	IDP2	ICDHy	NADPH	0.30	0.11	0.02	0.01
Glutamate dehydrogenase	GDH1 or GDH3	GLUDyi	NADPH	0.26	0.22	0.00	0.00
Phosphoglycerate dehydrogenase	SER3 or SER33	PGCD	NADH	0.03	0.03	0.01	0.00
Acetaldehyde dehydrogenase	ALD6	ALDD2y	NADPH	0.02	0.02	0.01	0.00
6-Phosphogluconate dehydrogenase	GND12	GND	NADPH	0.02	0.59	0.00	0.00
Glucose-6-phosphate dehydrogenase	ZWF1	G6PDH2	NADPH	0.02	0.59	0.00	0.00
Aspartate-semialdehyde dehydrogenase	HOM2	ASADi	NADPH	0.02	0.02	0.01	0.00
Homoserine dehydrogenase	HOM6	HSDxi	NADH	0.02	0.02	0.01	0.00
3-Isopropylmalate dehydrogenase	LEU2	IPMD	NADH	0.02	0.01	0.01	0.00
L-Aminoadipate-semialdehyde dehydrogenase	LYS2, LYS5	AASAD2	NADH	0.02	0.01	0.01	0.00
Saccharopine dehydrogenase (L-glutamate forming)	LYS9	SACCD1	NADPH	0.02	0.01	0.01	0.00
Saccharopine dehydrogenase (L-lysine forming)	LYS1	SACCD2	NADH	0.02	0.01	0.01	0.00
Shikimate dehydrogenase	ARO1	SHK3D	NADPH	0.01	0.01	0.01	0.00
Glycerol-3-phosphate dehydrogenase	GPD1	G3PD1ir	NADH	0.00	0.00	0.00	1.00
Xylose reductase	GRE3	XYLR	NADPH	0.00	1.00	0.00	1.00
Xylitol dehydrogenase	XYL2	XYLTD_D	NADH	0.00	1.00	0.00	1.00
Glutamate synthase	GLT1	GLUSx	NADH	0.00	0.00	0.10	0.03
Malate dehydrogenase	MDH2	MDH	NADH	0.00	0.00	0.08	0.03
Glycerol dehydrogenase	GCY1	GLYCDy	NADPH	0.00	0.00	0.00	1.00

[a] The flux through each reaction for maximum growth simulation using pFBA, relative to carbon uptake rate.
[b] The irreversible reaction ALCD2ir was removed from the model in favor of the equivalent but reversible ALCD2x.

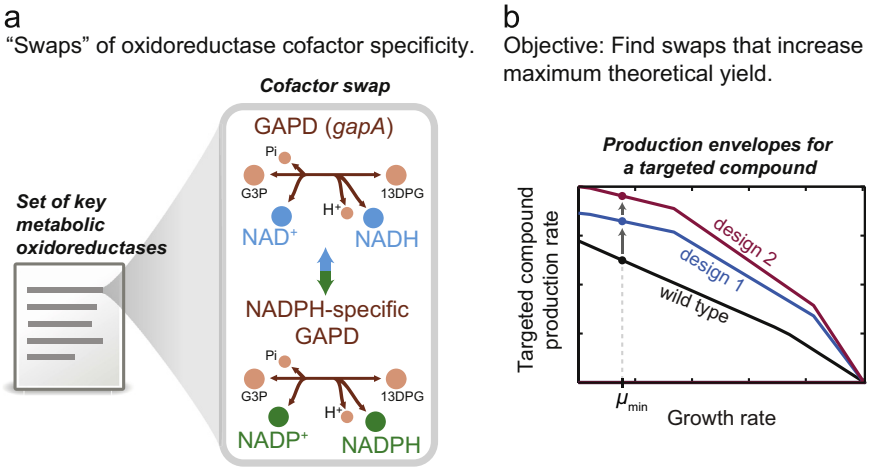

Fig. 1. Cofactor specificity modifications of oxidoreductase reactions can improve the maximum theoretical yield of metabolic byproducts. In this work, (a) oxidoreductase reactions in the metabolic model are identified as candidates for *in silico* cofactor "swaps" and (b) an optimization procedure identifies swaps that maximize the theoretical yield of a targeted compound.

2.6. Sensitivity analysis

A number of parameters can effect the simulation of theoretical yields; these include the oxygen uptake rate (OUR), the minimum growth rate (μ_{min}), and the SUR. A sensitivity analysis was performed for aerobic (OUR = 10 mmol gDW^{-1} h^{-1}) and anaerobic (OUR = 0) conditions. For a range of SUR values from 5 to 20 mmol gDW^{-1} h^{-1} (corresponding to a realistic range for glucose uptake in *E. coli*, Covert et al., 2004) and μ_{min} values from 0 to maximum growth, the optimal single swap was identified and the theoretical yield improvement was calculated.

2.7. Determining cofactor usage

To determine the NADPH usage of each production pathway, pFBA simulations were run optimizing for flux through each production pathway with zero growth ($\mu_{min} = 0$) and with the

a

		Anaerobic						Aerobic					
		Glucose			Xylose			Glucose			Xylose		
Product	No. of Carbons	wt	1 swap	2 swaps	wt	1 swap	2 swaps	wt	1 swap	2 swaps	wt	1 swap	2 swaps
L-Cysteine	3	0.10	0.18 GAPD		0.06	0.11 GAPD		0.39	0.43 GAPD	0.45 Group 8	0.41	0.44 GAPD	0.46 PGCD
L-Homoserine	4	0.26	0.45 GAPD	0.48 3+ rxns	0.15	0.31 GAPD		0.79	0.81 3+ rxns	0.83 GAPD	0.77	0.78 3+ rxns	0.80 Group 9
L-Threonine	4	0.23	0.40 GAPD		0.13	0.24 GAPD		0.77	0.79 3+ rxns	0.80 GAPD	0.75	0.76 3+ rxns	0.77 Group 9
1,5-Diaminopentane	5	0.21	0.37 GAPD		0.12	0.23 GAPD		0.59	0.60 GAPD	0.61 3+ rxns	0.57	0.58 GAPD	0.59 3+ rxns
L-Lysine	6	0.26	0.45 GAPD		0.15	0.27 GAPD		0.70	0.72 GAPD	0.73 3+ rxns	0.68	0.70 GAPD	0.70 3+ rxns
L-Isoleucine	6	0.28	0.47 GAPD	0.54 PDH	0.16	0.38 GAPD	0.39 3+ rxns	0.67	0.69 GAPD	0.70 3+ rxns	0.65	0.66 GAPD	0.67 3+ rxns
L-Leucine	6	0.40	0.56 GAPD	0.58 Group 1	0.25	0.40 GAPD		0.70	0.72 GAPD	0.72 3+ rxns	0.67	0.69 GAPD	0.70 3+ rxns
Glycine	2	0.17	0.24 GAPD	0.26 3+ rxns	0.10	0.15 GAPD		0.66	0.71 GAPD	0.74 3+ rxns	0.69	0.74 GAPD	0.78 PGCD
5-Methylthio-D-ribose	6	0.07	0.10 GAPD		0.04	0.06 GAPD		0.24	0.26 GAPD	0.26 3+ rxns	0.24	0.25 GAPD	0.25 3+ rxns
Spermidine	7	0.08	0.12 GAPD		0.05	0.07 GAPD		0.28	0.30 GAPD	0.30 3+ rxns	0.28	0.29 GAPD	0.29 3+ rxns
L-Proline	5	0.28	0.41 GAPD		0.16	0.25 GAPD		0.76	0.79 GAPD	0.79 3+ rxns	0.73	0.76 GAPD	0.76 3+ rxns
Cys-Gly	5	0.09	0.12 GAPD		0.05	0.08 GAPD		0.37	0.42 GAPD	0.43 Group 8	0.38	0.43 GAPD	0.44 3+ rxns
L-Aspartate	4	0.43	0.60 GAPD		0.25	0.36 GAPD		1.11	1.15 GAPD		1.08	1.11 GAPD	
Reduced glutathione	10	0.14	0.19 GAPD		0.08	0.11 GAPD		0.59	0.67 GAPD	0.68 3+ rxns	0.62	0.68 GAPD	0.68 3+ rxns
L-Serine	3	0.28	0.38 GAPD		0.16	0.23 GAPD		0.95	0.97 Group 6	0.97 3+ rxns	0.93	0.95 Group 6	0.95 3+ rxns
Thymine	5	0.13	0.17 GAPD		0.07	0.10 GAPD		0.60	0.67 GAPD	0.69 Group 8	0.62	0.70 GAPD	0.70 3+ rxns
Putrescine	4	0.17	0.22 GAPD		0.10	0.14 GAPD		0.58	0.60 GAPD		0.56	0.58 GAPD	
Glycerol	3	0.28	0.38 GAPD		0.16	0.23 GAPD		0.70	0.71 3+ rxns	0.71 3+ rxns	0.68	0.68 3+ rxns	0.69 3+ rxns

Percent increase in $Y_{P/S}$ from wildtype: 0% 10% 20% 30% 40% 50% 60% 70% >80%

b

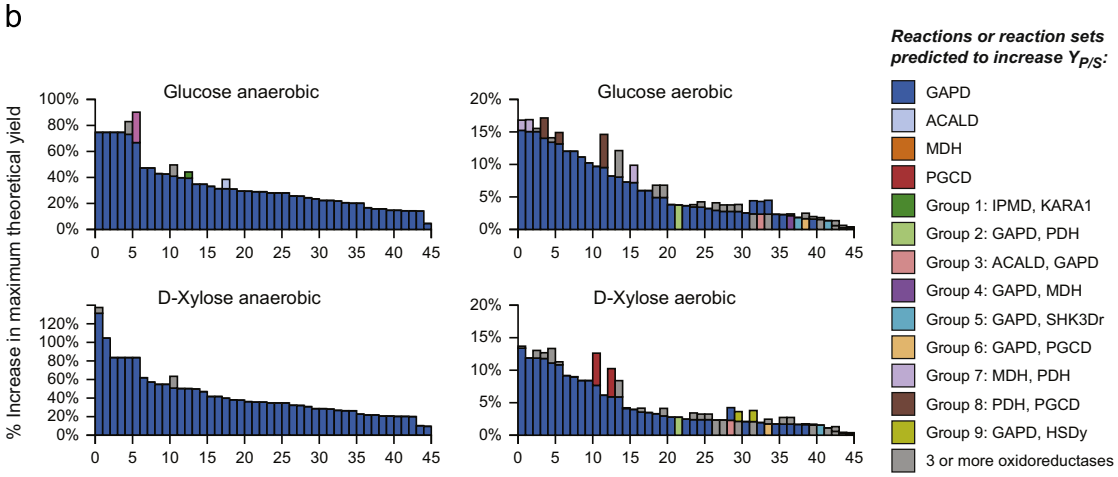

Fig. 2. Results from optimizing the cofactor specificity of oxidoreductases in the *E. coli* iJO1366 model to increase the maximum theoretical yield of native metabolic compounds. (a) Maximum theoretical yield for wildtype (wt) metabolic content and after one oxidoreductase swap (1 swap) and two oxidoreductase swaps (2 swap). Colors indicate the percent increase in maximum theoretical yield compared to wildtype. (b) Reactions found to most influence production of products are ranked and plotted by increase in theoretical maximum yield under each condition, and each color indicates a reaction or group of reactions that can be swapped to reach the optimal theoretical maximum yield.

transhydrogenase enzymes constrained to zero flux. Then, the flux through all reactions consuming NADPH was summed. This procedure was repeated for all carbon-containing metabolites that could be exported by the metabolic model.

All simulations were performed using MATLAB (The MathWorks Inc., Natick, MA, USA) and the COBRA Toolbox (Becker et al., 2007) software packages with TOMLAB /CPLEX (Tomlab Optimization Inc., San Diego, CA, USA; IBM, Armonk, NY, USA) and Gurobi (Gurobi Optimization, Inc., Houston, TX, USA) MILP solvers.

3. Results

3.1. Native pathways

To determine the effect of cofactor swaps on product yield, simulations were performed to optimize production of 81 and 154 target compounds in *E. coli* and *S. cerevisiae*, respectively, while allowing one and two swaps of oxidoreductase specificity out of the pool of oxidoreductase reactions (Tables 1 and 2). All carbon-containing molecules that can be exported by the metabolic

a

		Anaerobic				Aerobic			
		Glucose		Xylose*		Glucose		Xylose	
Product	No. of carbons	wt	swap	wt	swap	wt	swap	wt	swap
L-Aspartate	4	0.33	0.63 Group 2	0.20	0.70 ALCD2x	0.94		0.93	
L-Valine	5	0.41	0.78 Group 2	0.26	0.83 Group 1	0.75	0.78 Group 2	0.70	0.77 ALCD2x
L-Alanine	3	0.50	0.94 Group 2	0.31	1.00 Group 2	0.90	0.94 Group 2	0.84	0.93 ALCD2x
L-Serine	3	0.17	0.29 Group 1	0.10	0.50 ALCD2x	0.87	0.94 4+ rxns	0.76	0.93 4+ rxns
L-Cysteine	3	0.07	0.11 Group 1	0.04	0.18 ALCD2x	0.30	0.39 ALCD2x	0.29	0.42 ALCD2x
L-Lactate	3	0.17	0.26 Group 1	0.10	0.45 ALCD2x	0.77	0.80 ALCD2x	0.72	0.78 ALCD2x
L-Methionine	5	0.09	0.14 Group 1	0.06	0.23 ALCD2x	0.38	0.52 ALCD2x	0.37	0.50 ALCD2x
L-Phenylalanine	9	0.20	0.30 Group 1	0.13	0.51 ALCD2x	0.74	0.75 4+ rxns	0.69	0.73 ALCD2x
Pyruvate	3	0.40	0.53 Group 1	0.31	0.78 ALCD2x	0.94		0.93	
Thymine	5	0.08	0.10 Group 1	0.05	0.18 Group 1	0.36	0.39 Group 3	0.33	0.43 ALCD2x
sn-Glycero-3-phosphocholine	8	0.05	0.06 Group 1	0.03	0.11 Group 1	0.20	0.22 Group 3	0.20	0.24 ALCD2x
Glycine	2	0.11	0.15 Group 1	0.08	0.25 ALCD2x	0.53	0.76 ALCD2x	0.48	0.81 ALCD2x
L-Lysine	6	0.10	0.13 Group 2	0.06	0.23 Group 1	0.56	0.67 4+ rxns	0.50	0.65 ALCD2x
L-Proline	5	0.17	0.22 Group 2	0.10	0.38 Group 1	0.75	0.79 4+ rxns	0.70	0.77 ALCD2x
L-Asparagine	4	0.13	0.18 Group 2	0.08	0.30 Group 1	0.56	0.60 4+ rxns	0.55	0.65 Group 1
episterol	28	0.00		0.00		0.24	0.32 Group 1	0.21	0.36 ALCD2x
fecosterol	28	0.00		0.00		0.24	0.32 Group 1	0.21	0.36 ALCD2x
zymosterol	27	0.00		0.00		0.24	0.32 Group 1	0.21	0.36 ALCD2x
Ergosterol	28	0.00		0.00		0.23	0.30 Group 1	0.20	0.33 ALCD2x

Percent increase in $Y_{P/S}$ from wildtype: 0% 10% 20% 30% 40% 50% 60% 70% >80%

b

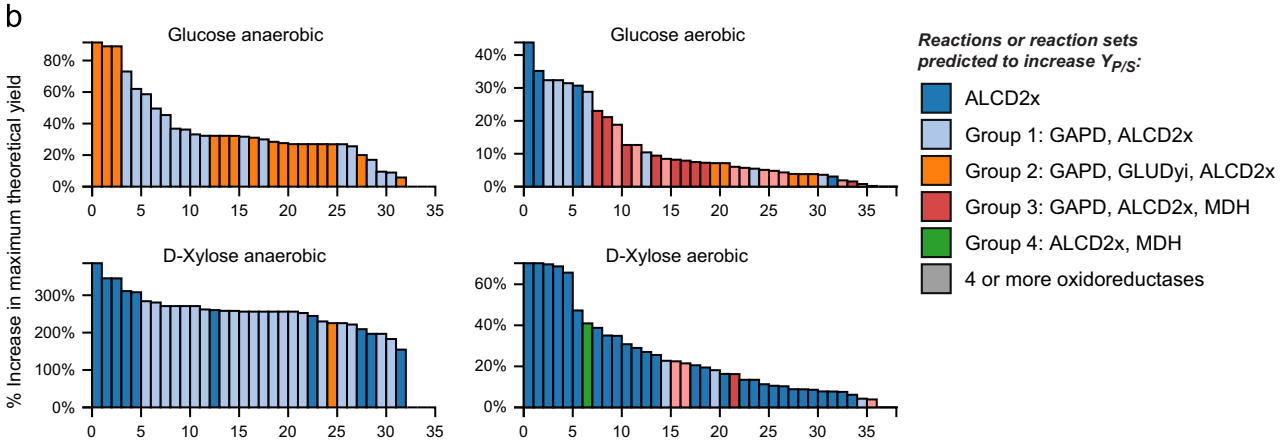

Fig. 3. Results from optimizing the cofactor specificity of oxidoreductases in the *S. cerevisiae* *i*MM904 model to increase the maximum theoretical yield of native metabolic compounds. (a) Carbon yield for wildtype (wt) metabolic content and after one oxidoreductase swap (swap). Colors indicate the percent increase in maximum theoretical yield compared to wildtype. (b) Reactions found to most influence production of products are ranked and plotted by increase in theoretical maximum yield under each condition, and each color indicates a reaction or group of reactions that can be swapped to reach the optimal theoretical maximum yield. Note that the D-xylose uptake reactions (XR/XDH) exist in the *i*MM904 model and thus in these simulations, but this pathway is not active in wildtype *S. cerevisiae* (see Methods). *Minimum growth rate set to 10% of the maximum growth rate.

models were considered. Increases in theoretical maximum yield were seen for many native products after one swap in both the *E. coli* *i*JO1366 and *S. cerevisiae* *i*MM904 models (Figs. 2 and 3). The theoretical yields for these pathways are dependent on the specific values selected for SUR, OUR, and μ_{min}. For these optimizations, the parameters were chosen to represent a possible scenario for a bioprocessing strain (maximum values of SUR = 10 mmol gDW^{-1} h^{-1} and OUR = 10 mmol gDW^{-1} h^{-1}, and a minimum value of $\mu_{min} = 0.1$ h^{-1}). Studies have reported specific growth rates in the range of 0.03–0.43 h^{-1} for production strains of *E. coli* (Martínez et al., 2008; Murarka et al., 2008; Bettiga et al., 2008; Qian et al., 2009; Rathnasingh et al., 2009), and the SUR was

	Anaerobic									Aerobic								
	Glucose			Xylose			Glycerol*			Glucose			Xylose			Glycerol		
Product	wt	1 swap	2 swaps	wt	1 swap	2 swaps	wt	1 swap	2 swaps	wt	1 swap	2 swaps	wt	1 swap	2 swaps	wt	1 swap	2 swaps
1,3-Propanediol	0.29	0.39 GAPD	0.46 PDH	0.17	0.31 GAPD	0.32 PDH	0.69	0.75 Group 1	0.78 PDH	0.63	0.63 4+ rxns	0.63 4+ rxns	0.61	0.61 4+ rxns	0.61 4+ rxns	0.71	0.72 4+ rxns	0.72 4+ rxns
1,4-Butanediol	0.51			0.43			0.64			0.64			0.62			0.68		
R,R-2,3-Butanediol	0.63			0.51			0.33			0.66			0.64			0.65		
meso-2,3-Butanediol	0.63			0.51			0.33			0.66			0.64			0.65		
R-3-Hydroxybutyrate	0.53	0.72 GAPD	0.73 4+ rxns	0.40	0.60 GAPD		0.36	0.40 GAPD		0.78	0.79 Group 2	0.81 GAPD	0.75	0.77 Group 2	0.79 GAPD	0.80	0.82 GAPD	0.83 4+ rxns
S-3-Hydroxybutyrate	0.63	0.74 GAPD		0.49	0.60 GAPD		0.40			0.79	0.81 Group 2	0.81 4+ rxns	0.77	0.78 Group 2	0.79 4+ rxns	0.81	0.83 4+ rxns	
3-Hydroxypropanoate	0.32	0.55 GAPD	0.61 PDH	0.19	0.41 GAPD		0.39			0.84	0.87 Group 1	0.89 PDH	0.81	0.84 Group 1	0.86 PDH	0.91	0.92 4+ rxns	0.92 4+ rxns
R-3-Hydroxyvalerate	0.48	0.61 GAPD		0.28	0.40 GAPD		0.34			0.75	0.76 Group 1		0.72	0.73 Group 1		0.78		
S-3-Hydroxyvalerate	0.59	0.61 Group 1		0.36	0.40 GAPD		0.34			0.76	0.76 4+ rxns		0.73	0.73 4+ rxns		0.78		
Styrene	0.24	0.29 GAPD		0.14	0.18 GAPD		0.15			0.67	0.67 Group 3		0.65	0.66 Group 4		0.71		
p-Hydroxystyrene	0.23	0.27 GAPD		0.13	0.17 GAPD		0.14			0.69	0.70 Group 3		0.67	0.68 Group 3		0.72	0.72 4+ rxns	
Lycopene	0.36	0.41 GAPD		0.00			0.22			0.66	0.67 4+ rxns		0.00			0.70		

Percent increase in $Y_{P/S}$ from wildtype: 0% 10% 20% 30% 40% 50% 60% 70% >80%

Reaction sets predicted to increase $Y_{P/S}$:

Group 1: GAPD, PDH
Group 2: ACALD, GAPD, PDH
Group 3: GAPD, SHK3Dr
Group 4: GAPD, PDH, SHK3Dr

Fig. 4. The effect of swapping cofactor specificity of oxidoreductases for the production of non-native compounds in *E. coli*. The table shows the maximum theoretical yield for wildtype (wt) and after one oxidoreductase swap (1 swap) and two oxidoreductase swaps (2 swaps), and the selected reaction or groups of reactions that can be swapped to reach the optimal theoretical maximum yield. Colors indicate the percent increase in maximum theoretical yield compared to wildtype. *Minimum growth rate set to 10% of the maximum growth rate.

chosen to match observed glucose uptake under aerobic conditions (Covert et al., 2004). To further justify the selection of these parameters for the global analysis, a sensitivity analysis was performed, discussed in Section 3.3.

3.1.1. *E. coli*

During anaerobic growth on glucose, cofactor optimizations of *i*JO1366 showed an increase in theoretical yield greater than 20% for 29 native products after one cofactor swap. The effect of a second swap was small in most simulations, and further swaps (>2 swaps) had no effect on product yield (Fig. 2). For anaerobic growth with D-xylose as a substrate, the increases in maximum theoretical yield were slightly greater in magnitude but qualitatively similar to those with glucose as a substrate. Aerobically, smaller increases in maximum theoretical yield were observed for all native targets and cofactor swaps. A 10–15% increase in theoretical yields was seen for 15 products (e.g., thymine, agmatine, L-arginine, D-alanine) with glucose as a substrate (Fig. 2A), and similar increases were seen with D-xylose as a substrate.

In *i*JO1366, converting the cofactor specificity of GAPD (EC 1.2.1.13) from production of NADH to production of NADPH has a global effect on the theoretical yield of native products. When considering the 45 metabolites with the greatest change in yield after cofactor swapping, simulated under 4 media conditions, it was observed that in all 180 cases swapping the native GAPD for the NADPH-dependent GAPD results in the greatest increase in yield, and for 155 of 180 cases no other cofactor swap can produce the theoretical maximum yield (the solutions are unique) (Fig. 2B). In order to investigate the relationship between cofactor swapping and the NADPH usage of production pathways, pFBA simulations (Lewis et al., 2010) were run maximizing production for each target molecule with zero growth. Then, for each simulation, the sum of the fluxes through all reactions that consume NADPH was calculated. This sum is an indication of the NADPH usage for producing a particular metabolite at the theoretical limit of production at steady state. A correlation was observed between the NADPH usage of production pathways and the improvement in theoretical yield after swapping GAPD (Pearson's $r=0.76$, Supplementary Fig. 1). This correlation points to the finding that producing more NADPH can lead to a higher maximum theoretical yield for a number of desired production targets.

3.1.2. *S. cerevisiae*

Cofactor swapping optimizations for *S. cerevisiae* also showed increases in theoretical yield after one cofactor swap (Fig. 3). *S. cerevisiae* does not contain an active pathway for D-xylose catabolism. However, many studies have reported the introduction of a D-xylose uptake pathway containing the enzymes XR, which oxidizes NADPH, and XDH, which reduces NAD^+ to NADH, and high D-xylose consumption has been achieved with the pathway from *Pichia stipitis* (Ghosh et al., 2011; Bengtsson et al., 2009). These reactions were utilized in the analysis of native yeast metabolism because they are included in *i*MM904 (as discussed in the Methods). Maximum theoretical yields of many products increase 250–350% after one swap when this pathway is modeled and D-xylose is present as the substrate. XI is an alternative, cofactor-balanced pathway for xylose uptake (Bettiga et al., 2008). When this enzyme is simulated, the predicted flux distribution utilizes XI in preference to XR/XDH. Using XI increases yield to the maximum theoretical yield achieved by cofactor swaps. Thus, modifying the D-xylose uptake pathway is an alternative to implementing the optimal cofactor swap.

Under aerobic conditions, increases in maximum theoretical yield were smaller than under anaerobic conditions, as in the *E. coli* metabolic model. However, the theoretical yield increases were still significant. For example, maximum theoretical yields of episterol, fecosterol, zymosterol, and ergosterol increased 29–32% with glucose as a substrate and 66–70% with D-xylose as a substrate after swapping the cofactor specificity of the alcohol dehydrogenase reaction ALCD2x (EC 1.1.1.1).

In the yeast metabolic model, swapping the cofactor specificity of a second oxidoreductase had no impact on theoretical yield. Also, in comparison to the results for *E. coli*, more of the yeast yield optimizations resulted in non-unique solution groups (Fig. 3B). For instance, under anaerobic conditions with glucose substrate, the L-serine yield reached a maximum of 0.47 Cmol/Cmol substrate after swapping the cofactor specificity of either oxidoreductase enzyme in the solution group (Group 1: GAPD or ALCD2x).

In contrast to *E. coli*, the reaction that appeared most often in solutions is ALCD2x, and GAPD appeared secondarily.

3.2. Non-native pathways in E. coli

Major building block molecules that have been heavily studied as non-native products of *E. coli* were selected for the cofactor swapping analysis. These products include (1) the polyester building-blocks 1,3-propanediol, 1,4-butanediol, 2,3-butanediol, 3-hydroxybutyrate, and 3-hydroxyvalerate, and (2) the polyvinyl building blocks styrene and hydroxystyrene (Adkins et al., 2012). Production of lycopene, a red-colored carotenoid sold commercially as a colorant and a nutritional supplement, was also simulated. When applicable, production pathways for alternate sterioisomers were considered, so that, in total, twelve non-native pathways were reconstructed in the *E. coli* model (Table 3).

Optimization of cofactor swapping increased the theoretical maximum yield of 9 of 12 non-native pathways (Fig. 4). The effects were much greater anaerobically; under aerobic conditions, less than 10% increases were observed. For three products (1,4-butanediol, R,R-2,3-butanediol, and meso-2,3-butanediol), the optimization did not find any solution, indicating that modifying cofactor specificity is not predicted to increase the maximum yield of these products. For these products, a sensitivity analysis was performed (Section 3.3).

3.3. NADPH yield and parameter sensitivity

The effect of cofactor swaps was observed to depend on the values of SUR, OUR, and μ_{min} selected for the simulation. To better understand the relationship between these parameters and the observed improvements in theoretical yield, a simple scenario where *E. coli* is used to regenerate reducing equivalents of NADPH for the conversion of cyclohexanone to ε-caprolactone was explored. Martínez et al. (2008) performed a GAPD cofactor swap in *E. coli* and observed an increase in the yield of NADPH reducing equivalents for this chemical conversion. By examining this simple case, the effect of the cofactor swaps on NADPH yield can be isolated from other factors such as shifts in carbon metabolism.

Cofactor optimizations were performed for a range of SUR and μ_{min} values, under aerobic and anaerobic conditions. The NADPH yield in *E. coli* was observed to decrease as the biomass production approached a maximum (Fig. 5A,D). However, swapping the cofactor specificity of the GAPD reaction resulted in increased maximum theoretical yield compared to the wildtype (Fig. 5C,F). It was also observed that the selection of GAPD as the optimal swap was consistent across parameters for aerobic and anaerobic conditions (see Discussion for further analysis).

For the non-native compounds 1,4-butanediol and 2,3-butanediol, running the cofactor swap optimization with the default parameters did not identify any opportunities to improve theoretical yield. The same sensitivity analysis was performed for these target molecules to determine whether, under any other conditions, a cofactor swap might be beneficial (Supplementary Fig. 2). It was observed that under a certain range of SUR and μ_{min} values, when growth is near the maximum growth rate, a swap does improve theoretical yield for 1,4-butanediol, under aerobic and anaerobic conditions. Specifically, under anaerobic conditions, it was found that there were 790 feasible parameter sets based on the parameters of the sensitivity analysis (see Methods). Out of these, the cofactor swap was beneficial for 292 of the 790 feasible parameter sets (i.e., 37%, and the increase can be seen in Supplementary Fig. 2). Furthermore, when examining both 2,3-butanediol isomers during the sensitivity analysis, the cofactor swapping did not improve theoretical yield at any values of SUR and μ_{min}.

Finally, the sensitivity analysis was performed for L-cysteine production (on glucose minimal media in *E. coli*) to further examine whether the optimizations for a single parameter set (Figs. 2–4) are sensitive to the selection of SUR, OUR, and μ_{min}. It was observed that theoretical yields were improved by the GAPD cofactor swap across a broad range of SUR, OUR, μ_{min} values (Supplementary Fig. 3). Thus, the parameter set chosen for the global analysis, where SUR and μ_{min} are relatively small, leads to accurate estimations of the maximum effect of a cofactor swap (e.g., L-cysteine anaerobically), or even underestimations of the maximum effect of a cofactor swap (e.g., L-cysteine aerobically, 1,4-butanediol), and the general trends observed in the global analysis are a meaningful representation of the total impact of cofactor swapping on theoretical yields.

4. Discussion

4.1. Cofactor swaps for certain enzymes have a global impact on theoretical yields

This study presents a computational analysis to determine optimal cofactor-specificity swaps of oxidoreductase enzymes in genome-scale metabolic models, specifically for the production organisms *E. coli* and *S. cerevisiae*. Increases in theoretical yield were observed for many products of *E. coli* and *S. cerevisiae* metabolism after one cofactor swap, and swapping certain reactions (esp. GAPD, ALCD2x) was seen to have a global benefit for

Table 3
Non-native pathways reconstructed in *E. coli* for cofactor balance optimization.

Product	Number of reactions (+number of transporters) added to model	Reducing equivalents consumed in non-native pathway	Reference
1,3-Propanediol	2 (+2)	1 NADPH	(Tang et al., 2009)
1,4-Butanediol	7 (+2)	3 NADH	(Yim et al., 2011)
Meso-2,3-Butanediol	3 (+1)	1 NADH	(Ui et al., 2004)
R,R-2,3-Butanediol	3 (+1)	1 NADH	(Yan et al., 2009)
R-3-Hydroxybutyrate	4 (+2)[a]	1 NADPH	(Tseng et al., 2009)
S-3-Hydroxybutyrate	4 (+2)[a]	1 NADH	(Tseng et al., 2009)
3-Hydroxypropanoate	2 (+2)	2 NADPH	(Rathnasingh et al., 2012)
R-3-Hydroxyvalerate	5 (+2)	1 NADPH	(Tseng et al., 2010)
S-3-Hydroxyvalerate	5 (+2)	1 NADH	(Tseng et al., 2010)
Styrene	2 (+1)	None	(McKenna and Nielsen, 2011)
p-Hydroxystyrene	2 (+1)	None	(Qi et al., 2007)
Lycopene	3 (+1)	8 NADPH	(Martínez et al., 2008)

[a] Two alternate pathways were reconstructed for these cases, according to the literature.

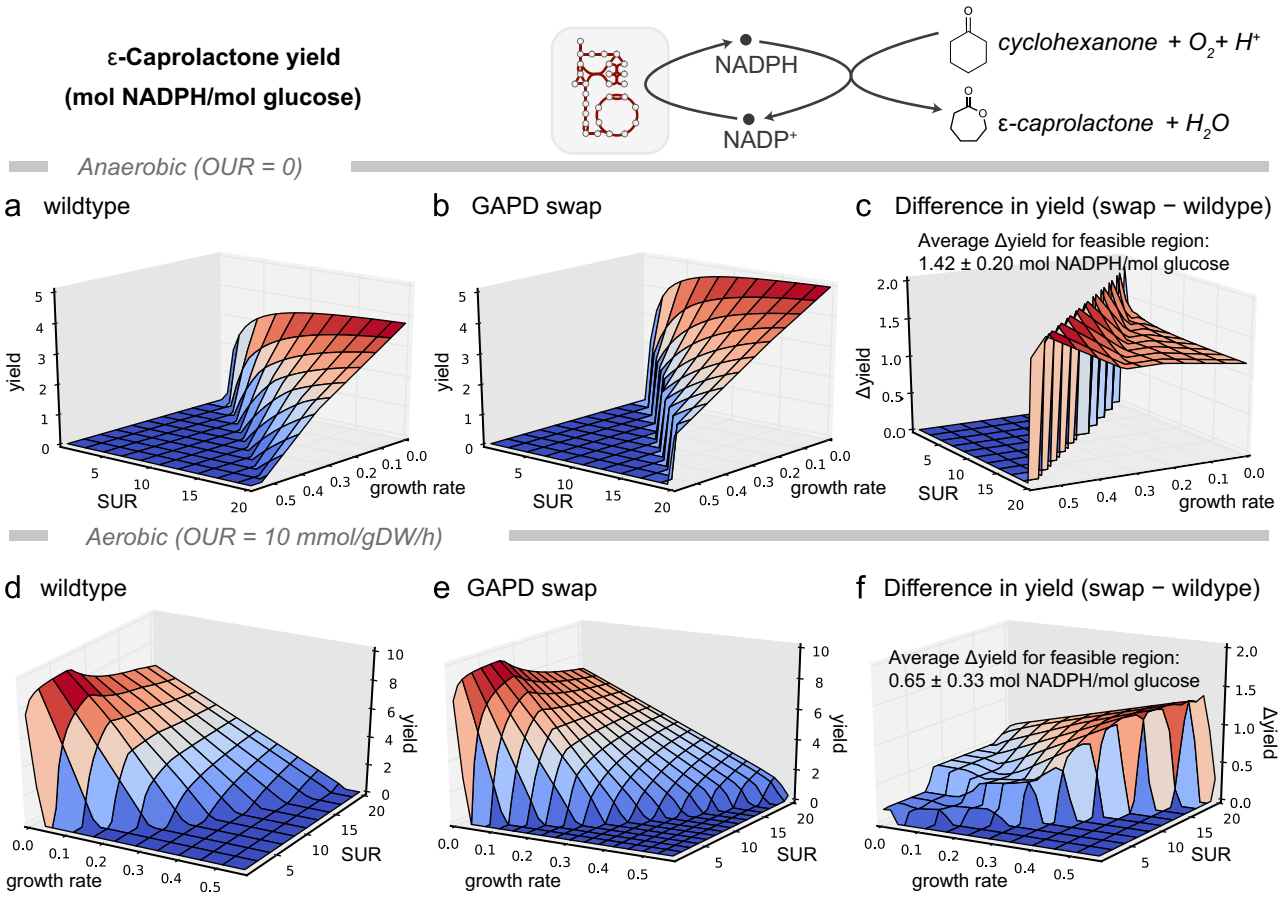

Fig. 5. A sensitivity analysis on the impact of cofactor swapping when varying modeling parameters. The conversion of cyclohexanone to ε-caprolactone requires reducing equivalents which can be provided by *E. coli* growing in glucose minimal media. For this conversion, a sensitivity analysis was performed for (a–c) anaerobic and (d–f) aerobic conditions. Substrate uptake rate (SUR) and the minimum growth rate (μ_{min}) were varied across a wide range of values, and it was observed that swapping GAPD increased carbon yield across all parameters. For infeasible μ_{min} values, no growth is possible, so yield is zero (dark blue regions). Across all feasible datapoints, the average and standard deviation of the Δyield are given and can be compared to the increase in yield observed by Martínez et al. (2008) of 1.25 ± 0.19 mol NADPH/mol glucose.

theoretical yields. The theoretical yield improvements were found for a number of native products that are produced on an industrial scale and have been considered for bioproduction using *E. coli* or *S. cerevisiae*. These include the native compounds L-lysine, L-isoleucine, L-proline, L-serine, L-threonine, L-aspartate, L-lactate, 1,5-diaminopentane (cadaverine), and putrescine, and the non-native compounds 3-propanediol, 3-hydroxybutyrate, 3-hydroxypropanoate, 3-hydroxyvalerate, styrene, and lycopene. (For review articles describing the bioproduction of these compounds in *E. coli*, see Becker and Wittmann, 2012; Jang et al., 2012).

Nearly all single swaps selected by the optimization for native and non-native products in *E. coli* were for the GAPD enzyme. This is a central enzyme in glycolysis, and the change in electron carrier specificity causes a shift in central metabolism toward NADPH production. Because the GAPD cofactor swap has such a global impact on biosynthesis in *E. coli* metabolism, a stable strain with NADPH-dependent GAPD could be a useful starting point for engineering *E. coli* for production of any of the products reported here (Figs. 2 and 4). Published results have shown that experimentally replacing the GAPD in *E. coli* with an NADPH-dependent GAPD increases yield and productivity of lycopene and e-caprolactone (Martínez et al., 2008). The optimizations presented here demonstrate that the GAPD cofactor swap is an ideal choice. However, for some products, swapping the cofactor specificity of other important oxidoreductase enzymes can have the same effects on yield (e.g., pyruvate dehydrogenase [PDH, EC 1.2.4.1], malate dehydrogenase [MDH, EC 1.1.1.37], and phosphoglycerate dehydrogenase [PGCD, EC 1.1.1.95]).

The environmental conditions (aerobicity, substrate) were a major determinant of the impact of cofactor swapping. Simulations with D-xylose as a substrate showed greater increases in yield after swapping than those with glucose substrate (Figs. 2–4), and, with glycerol as a substrate, very little improvement was possible with cofactor swapping (Fig. 4). Aerobicity also had a major effect. In anaerobic simulations, cofactor swapping had a much greater impact than in aerobic simulations. SUR and $\mu_{\min}$ also affected the results of cofactor swapping (as described below). Some cases (e.g., aerobic production of glycine in yeast) are exceptions to these trends, so the relationship between environment conditions, cofactor swapping, and theoretical yield is complex, and the cofactor balance depends on the exact metabolic state of a cell.

Simulations utilizing the *S. cerevisiae* metabolic network differed from *E. coli* in that more cofactor swaps could theoretically produce the same improvement in yield, as seen in the many non-unique solutions (Fig. 3B). Thus, one has more flexibility to choose interventions as various swaps can produce the same effect. The exact enzyme to modify can be determined by other factors, such as the strength of regulation of the native enzyme or the availability of an appropriate enzyme with alternate cofactor specificity. As in *E. coli*, the native GAPD in *S. cerevisiae* has been experimentally replaced with a NADPH-dependent GAPD, resulting in higher D-xylose fermentation (Verho et al., 2003) and the same design has been patented for the production of many products (ethanol, lactic acid, polyhydroxyalkanoates, amino acids, fats, vitamins, and nucleotides) (Londesborough et al., 2003).

However, the results of cofactor swapping optimizations demonstrate that modifying ALCD2x can have the same impact on theoretical yield, and, in many cases, any one of a group of enzymes can increase theoretical yield to optimal levels.

4.2. Simulated theoretical yield improvement matches experimental observations for a GAPD swap

To determine how cofactor swapping modifies the capabilities of the metabolic network, a simple scenario was explored where *E. coli* generates reducing equivalents of NADPH that drive a bioprocessing reaction. In such a scenario, the effect of cofactor balance can be isolated from biomass production and carbon metabolism. In an experimental study, Martínez et al. (2008) compared the ability of wildtype and GAPD-swap *E. coli* strains to produce reducing equivalents of NADPH for a reaction that converts cyclohexanone to ϵ-caprolactone (utilizing one mole of NADPH per mole of ϵ-caprolactone produced). The authors reported an increase in NADPH yield from 1.72 ± 0.19 to 2.97 ± 0.05 mol NADPH/mol glucose. For *in silico* optimizations of ϵ-caprolactone production, GAPD is the optimal cofactor swap across the range of SUR, OUR, and $\mu_{\min}$ parameters, and swapping GAPD increases theoretical yield for all parameter values (Fig. 5). The increase in theoretical yield after swapping GAPD is somewhat consistent across parameters: on average 1.42 ± 0.20 mol NADPH/mol glucose under anaerobic conditions and 0.65 ± 0.33 mol NADPH/mol glucose under the examined aerobic conditions (Fig. 5C,F). Thus, the simulated increase in theoretical yield is a plausible explanation for the 1.25 ± 0.19 mol NADPH/mol glucose increase in yield observed by Martínez et al. (2008).

4.3. Optimal cofactor swaps increase ATP availability

It has been reported that transhydrogenase enzymes in *E. coli* are responsible for producing 35–45% of the NADPH necessary for biosynthesis (Sauer et al., 2004). However, the transfer of reducing equivalents from NAD(H) to NADP(H) requires energy from proton translocation to proceed. Thus, *in vivo* and *in silico*, the amount of NADPH that can be produced by the transhydrogenase enzyme is limited by this energy requirement. Some yield increases for products of biosynthetic reactions can be achieved by overexpressing the membrane-bound transhydrogenase encoded by *sthA* or down-regulating the soluble transhydrogenase encoded by *pntAB* (Sanchez et al., 2006; Jan et al., 2013). However, the metabolic model predicts that modifying transhydrogenases will not lead to the maximum theoretical yield. Producing NADPH directly during a metabolic transformation is inherently more efficient than utilizing the transhydrogenase enzyme and the electrochemical gradient to generate NADPH from NADH. At the theoretical limit of NADPH production under anaerobic conditions (Supplementary Fig. 4), the GAPD cofactor swap allows for a decrease in NADPH production by the energy-coupled transhydrogenase enzyme. Thus, demand for transmembrane electrochemical potential (i.e., ATP equivalents) from the transhydrogenase enzyme decreases, and more flux can be directed away from glycolysis and ATP production and towards NADPH producing enzymes (e.g., in the pentose phosphate pathway). These trends were also investigated for native products of *E. coli* iJO1366 for which the GAPD swap increased theoretical yield. For these products, it was observed that, at the theoretical maximum production state, the flux through glycolysis, the pentose phosphate pathway, and TCA cycle were not generally shifted after swapping GAPD (Supplementary Fig. 5). Thus, glycolytic flux remains high in the GAPD-swap simulation, and any ATP that was no longer necessary for transhydrogenase activity can be directed to product and biomass generation.

In contrast to *E. coli*, no transhydrogenase enzymes have been identified in yeast (Nissen et al., 2001). Therefore, balancing cofactor production and consumption in *S. cerevisiae* is even more critical. In yeast, under anaerobic conditions, the theoretical yield of NADPH increases after the ALCD2x swap because NADH production decreases, so less flux is directed to ethanol fermentation (which would act to oxidize NADH), and more flux can be directed to acetate fermentation (which has a higher ATP yield than ethanol fermentation). Thus, the decrease in NADH production can theoretically increase the availability of ATP at steady state—even in the absence of transhydrogenase enzymes.

4.4. Cofactor swapping in yeast has a greater effect with D-xylose as a substrate

When D-xylose is used as a substrate, swapping the specificity of a single oxidoreductase can have an even greater effect on the production of native products in *S. cerevisiae*. There has been much interest in the use of 5-carbon sugars as a feedstock for *S. cerevisiae* production strains, which can be accomplished with heterologous expression of the *XYL1* and *XYL2* genes, encoding XR and XDH, respectively, from *P. stipitis* (Bengtsson et al., 2009). This XR is NADPH-dependent, and modifications to the cofactor specificity of XR have been shown to increase xylose fermentation, experimentally (Bengtsson et al., 2009) and computationally (Ghosh et al., 2011). The simulations presented here show that swapping the cofactor specificity of high-flux, central metabolic enzymes (i.e., GAPD, ALCD2x) can increase theoretical yields by 2–3 fold across native metabolism in *S. cerevisiae* by generating more NADPH to drive D-xylose uptake by XR (Fig. 3, see Methods for how the native XR and XDH in the model are accounted for). Our results also show that swapping the cofactor specificity of these central metabolic enzymes has a greater impact on theoretical yield in yeast than cofactor swaps for XR or XDH. Utilizing the cofactor-balanced XI enzyme (Bettiga et al., 2008) is another approach to cofactor balancing D-xylose uptake, and the XI uptake pathway also maximizes the theoretical yield *in silico*.

While theoretical yield improves for many products with D-xylose as a substrate, the theoretical yield of ethanol fermentation was not much improved by cofactor swapping. Under aerobic conditions, simulations showed an increase in theoretical yield from 0.606 to 0.619 Cmol/Cmol (a 2% increase) after swapping the cofactor specificity of ALCD2x. Under anaerobic conditions, no increase was observed. In a previous study, Verho et al. (2003) compared ethanol fermentation from D-xylose with the native GAPD enzyme to ethanol fermentation with an NADPH-GAPD, and the authors reported a significant increase in ethanol yield (from 0.24 to 0.41 Cmol/Cmol). The results of this work suggest that the theoretical yield was not significantly increased in the GAPD-swap strain, and so other effects (e.g., regulation and kinetics) could be considered to explain the improved yield.

4.5. Theoretical yield of non-native products increases with swaps

Optimizations for the non-native pathways reconstructed in *E. coli* demonstrate that cofactor swapping can also improve maximum theoretical yields for non-native industrial products. In particular, the theoretical yields of 1,3-propanediol, 3-hydroxybutyrate, 3-hydroxypropanoate, 3-hydroxyvalerate, styrene, and lycopene were increased after one cofactor swap. For biomonomer production strains, yield is an extremely important consideration. Final titer, productivity, and yield are the three most important design parameters in bioprocessing, and yield plays a central role in determining the economic viability of a production process for high volume/low cost products like biomonomers (Villadsen et al., 2011). The same type of modifications of cofactor specificity that

have been used to increase yield of native products in *E. coli* (Martínez et al., 2008) and *S. cerevisiae* (Bengtsson et al., 2009; Verho et al., 2003) are predicted to increase the maximum yield of these non-native products.

In the global analysis, theoretical yields of 1,4-butanediol and 2,3-butanediol were not affected by cofactor swapping. For these products, a detailed analysis of swapping across the parameter space of SUR, OUR, and μ_{min} (Supplementary Fig. 2) showed that some improvement in theoretical yield *is* possible for 1,4-butanediol at certain parameter values. Specifically, improvements in theoretical yield are prevalent when μ_{min} is near its maximum value for a given SUR. Only at higher growth rates does cofactor balance become a limitation for production of 1,4-butanediol. However, across the parameter range, cofactor swaps did not increase the thoeretical yield for either of the 2,3-butanediol isomers. The case of 1,4-butanediol demonstrates that one should pay close attention to SUR, OUR, and μ_{min} and their effects when simulating theoretical yield and optimizing cofactor balance.

4.6. Theoretical yields are sensitive to knowledge of cofactor preference and enzyme promiscuity

A limitation of this analysis is the possibility of alternative cofactor usage by promiscuous oxidoreductase enzymes (Olavarría et al., 2012). This limitation can be addressed by improving the model to include additional reactions for those oxidoreductases that are known to catalyze flux with multiple cofactors, and, in fact, some reactions do have this kind of annotation (e.g., *frdABCD* in *i*JO1366, Orth et al., 2011). To engineer a strain that generates a product near the theoretical yield, it is still necessary to optimize the native regulation and the kinetics and thermodynamics of the optimal pathways. Also, the theoretical yields presented here are purely stoichiometric, so any theoretical kinetic or thermodynamic limitations beyond reaction reversibilities (Feist et al., 2007) are not included in the analysis. Furthermore, the kinetic and thermodynamic impacts of cofactor swapping a particular enzyme could lead to unintended consequences in the network that are not reflected in the thoeretical yield. Published cases of successfully swapping the cofactor specificity of oxidoreductase enzymes have not identified this as a major limitation (Verho et al., 2003; Martínez et al., 2008), but it is an important consideration, especially as one approaches the theoretical limits of bioproduction.

5. Conclusion

Constraint-based modeling is uniquely suited for modeling optimal metabolic states, because optimizations like cofactor swapping can be performed for large sets of products and environmental conditions. The optimizations reported here demonstrate the importance of cofactor swapping in *E. coli* and *S. cerevisiae* for native and non-native products, and show that microbial byproducts can be organized according to the need for a synthetic increase in NADPH production. These results also highlight the centrality of certain enzymes and that swapping the cofactor specificity of these enzymes (GAPD, ALCD2x) has a global effect on cofactor balance in the metabolic network. These methods are especially applicable for highly engineered strains that can generate metabolic products with yields approaching the maximum theoretical yield (Murarka et al., 2008; Trinh, 2012; Lin et al., 2005). Cofactor swapping could be used to tune high-yield industrial bioprocessing strains, and experimental validation of the optimal swaps predicted for *E. coli* and *S. cerevisiae* could have an immediate application in industry.

Acknowledgments

We would like to thank Bernhard Palsson and Daniel Zielinski for their guidance and suggestions during the course of this project. Funding for this work was provided by the Novo Nordisk Foundation and by the National Science Foundation Graduate Research Fellowship under Grant no. DGE-1144086.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2014.05.009.

References

Adkins, J., Pugh, S., McKenna, R., Nielsen, D.R., 2012. Engineering microbial chemical factories to produce renewable "biomonomers". Front. Microbiol. 3, 313, http://dx.doi.org/10.3389/fmicb.2012.00313.

Becker, J., Wittmann, C., 2012. Systems and synthetic metabolic engineering for amino acid production – the heartbeat of industrial strain development. Curr. Opin. Biotechnol. 23, 718–726, http://dx.doi.org/10.1016/j.copbio.2011.12.025.

Becker, S.A., Feist, A.M., Mo, M.L., Hannum, G., Palsson, B.Ø., Herrgård, M.J., 2007. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. Nat. Protoc. 2, 727–738, http://dx.doi.org/10.1038/nprot.2007.99.

Bengtsson, O., Hahn-Hägerdal, B., Gorwa-Grauslund, M.F., 2009. Xylose reductase from *Pichia stipitis* with altered coenzyme preference improves ethanolic xylose fermentation by recombinant *Saccharomyces cerevisiae*. Biotechnol. Biofuels 2, 9, http://dx.doi.org/10.1186/1754-6834-2-9.

Berríos-Rivera, S.J., Bennett, G.N., San, K.Y., 2002a. Metabolic engineering of *Escherichia coli*: increase of NADH availability by overexpressing an NAD^+-dependent formate dehydrogenase. Metab. Eng. 4, 217–229, http://dx.doi.org/10.1006/mben.2002.0227.

Berríos-Rivera, S.J., Bennett, G.N., San, K.Y., 2002b. The effect of increasing NADH availability on the redistribution of metabolic fluxes in *Escherichia coli* chemostat cultures. Metab. Eng. 4, 230–237, http://dx.doi.org/10.1006/mben.2002.0228.

Berríos-Rivera, S.J., San, K.Y., Bennett, G.N., 2002. The effect of NAPRTase overexpression on the total levels of NAD, The NADH/NAD^+ ratio, and the distribution of metabolites in *Escherichia coli*. Metab. Eng. 4, 238–247, http://dx.doi.org/10.1006/mben.2002.0229.

Berríos-Rivera, S.J., Sánchez, a.M., Bennett, G.N., San, K.Y., 2004. Effect of different levels of NADH availability on metabolite distribution in *Escherichia coli* fermentation in minimal and complex media. Appl. Microbiol. Biotechnol. 65, 426–432, http://dx.doi.org/10.1007/s00253-004-1609-3.

Bettiga, M., Hahn-Hägerdal, B., Gorwa-Grauslund, M.F., 2008. Comparing the xylose reductase/xylitol dehydrogenase and xylose isomerase pathways in arabinose and xylose fermenting *Saccharomyces cerevisiae* strains. Biotechnol. Biofuels 1, 16, http://dx.doi.org/10.1186/1754-6834-1-16.

Blaschkowski, H., Neuer, G., Ludwig-Festl, M., Knappe, J., 1982. Routes of flavodoxin and ferredoxin reduction in *Escherichia coli*. Eur. J. Biochem. 123, 563–569, http://dx.doi.org/10.1111/j.1432-1033.1982.tb06569.x.

Chin, J.W., Khankal, R., Monroe, C.A., Maranas, C.D., Cirino, P.C., 2009. Analysis of NADPH supply during xylitol production by engineered *Escherichia coli*. Biotechnol. Bioeng. 102, 209–220, http://dx.doi.org/10.1002/bit.22060.

Chung, B.K.S., Lakshmanan, M., Klement, M., Mohanty, B., Lee, D.Y., 2013. Genome-scale *in silico* modeling and analysis for designing synthetic terpenoid-producing microbial cell factories. Chem. Eng. Sci. 103, 100–108, http://dx.doi.org/10.1016/j.ces.2012.09.006.

Covert, M.W., Knight, E.M., Reed, J.L., Herrgård, M.J., Palsson, B.Ø., 2004. Integrating high-throughput and computational data elucidates bacterial networks. Nature 429, 92–96, http://dx.doi.org/10.1038/nature02456.

Feist, A.M., Henry, C.S., Reed, J.L., Krummenacker, M., Joyce, A.R., Karp, P.D., Broadbelt, L.J., Hatzimanikatis, V., Palsson, B.Ø., 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, http://dx.doi.org/10.1038/msb4100155.

Feist, A.M., Zielinski, D.C., Orth, J.D., Schellenberger, J., Herrgård, M.J., Palsson, B.Ø., 2010. Model-driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. Metab. Eng. 12, 173–186, http://dx.doi.org/10.1016/j.ymben.2009.10.003.

Ghosh, A., Zhao, H., Price, N.D., 2011. Genome-scale consequences of cofactor balancing in engineered pentose utilization pathways in *Saccharomyces cerevisiae*. PLoS One 6, e27316, http://dx.doi.org/10.1371/journal.pone.0027316.

Gottschalk, G., 1986. Bacterial Metabolism, 2nd ed. Springer-Verlag, New York, NY.

Hyung Lim, J., Woo Seo, S., Yeon Kim, S., Yeol Jung, G., 2013. Model-driven rebalancing of the intracellular redox state for optimization of a heterologous n-butanol pathway in *Escherichia coli*. Metab. Eng. 20, 56–62, http://dx.doi.org/10.1016/j.ymben.2013.09.003.

Jan, J., Martinez, I., Wang, Y., Bennett, G.N., San, K.Y., 2013. Metabolic engineering and transhydrogenase effects on NADPH availability in *Escherichia coli*. Biotechnol. Prog. 29, 1124–1130, http://dx.doi.org/10.1002/btpr.1765.

Jang, Y.S., Kim, B., Shin, J.H., Choi, Y.J., Choi, S., Song, C.W., Lee, J., Park, H.G., Lee, S.Y., 2012. Bio-based production of C2–C6 platform chemicals. Biotechnol. Bioeng. 109, 2437–2459, http://dx.doi.org/10.1002/bit.24599.

Kauffman, K.J., Prakash, P., Edwards, J.S., 2003. Advances in flux balance analysis. Curr. Opin. Biotechnol. 14, 491–496, http://dx.doi.org/10.1016/j.copbio.2003.08.001.

King, Z.A., Feist, A.M., 2013. Optimizing cofactor specificity of oxidoreductase enzymes for the generation of microbial production strains—OptSwap. Ind. Biotechnol. 9, 236–246, http://dx.doi.org/10.1089/ind.2013.0005.

Lakshmanan, M., Chung, B.K.S., Liu, C., Kim, S.W., Lee, D.Y., 2013. Cofactor modificiation analysis: a computational framework to identify cofactor specificity engineering targets for strain improvement. J. Bioinform. Biotechnol. 11, 1343006.

Lee, W.H., Kim, M.D., Jin, Y.S., Seo, J.H., 2013. Engineering of NADPH regenerators in Escherichia coli for enhanced biotransformation. Appl. Microbiol. Biotechnol. 97, 2761–2772, http://dx.doi.org/10.1007/s00253-013-4750-z.

Lewis, N.E., Hixson, K.K., Conrad, T.M., Lerman, J.A., Charusanti, P., Polpitiya, A.D., Adkins, J.N., Schramm, G., Purvine, S.O., Lopez-Ferrer, D., Weitz, K.K., Eils, R., Konig, R., Smith, R.D., Palsson, B.Ø., 2010. Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. Mol. Syst. Biol. 6, 390, http://dx.doi.org/10.1038/msb.2010.47.

Lin, H., Bennett, G.N., San, K.Y., 2005. Metabolic engineering of aerobic succinate production systems in *Escherichia coli* to improve process productivity and achieve the maximum theoretical succinate yield. Metab. Eng. 7, 116–127, http://dx.doi.org/10.1016/j.ymben.2004.10.003.

Londesborough, J., Penttilae, M., Richard, P., Verho, R., 2003. Fungal micro-organism having an increased ability to carry out biotechnological process(es). WO 2003/038067 A1.

Lunzer, M., Miller, S.P., Felsheim, R., Dean, A.M., 2005. The biochemical architecture of an ancient adaptive landscape. Science 310, 499–501, http://dx.doi.org/10.1126/science.1115649.

Martínez, I., Zhu, J., Lin, H., Bennett, G.N., San, K.Y., 2008. Replacing *Escherichia coli* NAD-dependent glyceraldehyde 3-phosphate dehydrogenase (GAPDH) with a NADP-dependent enzyme from *Clostridium acetobutylicum* facilitates NADPH dependent pathways. Metab. Eng. 10, 352–359, http://dx.doi.org/10.1016/j.ymben.2008.09.001.

McKenna, R., Nielsen, D.R., 2011. Styrene biosynthesis from glucose by engineered *E. coli*. Metab. Eng. 13, 544–554, http://dx.doi.org/10.1016/j.ymben.2011.06.005.

Mo, M.L., Palsson, B.Ø., Herrgård, M.J., 2009. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. BMC Syst. Biol. 3, 37, http://dx.doi.org/10.1186/1752-0509-3-37.

Murarka, A., Dharmadi, Y., Yazdani, S.S., Gonzalez, R., 2008. Fermentative utilization of glycerol by *Escherichia coli* and its implications for the production of fuels and chemicals. Appl. Environ. Microbiol. 74, 1124–1135, http://dx.doi.org/10.1128/AEM.02192-07.

Nissen, T.L., Anderlund, M., Nielsen, J., Villadsen, J., Kielland-Brandt, M.C., 2001. Expression of a cytoplasmic transhydrogenase in *Saccharomyces cerevisiae* results in formation of 2-oxoglutarate due to depletion of the NADPH pool. Yeast 18, 19–32 http://dx.doi.org/10.1002/1097-0061(200101)18:1⟨19::AID-YEA650⟩3.0.CO;2-5.

Olavarría, K., Valdés, D., Cabrera, R., 2012. The cofactor preference of glucose-6-phosphate dehydrogenase from *Escherichia coli*–modeling the physiological production of reduced cofactors. FEBS J. 279, 2296–2309, http://dx.doi.org/10.1111/j.1742-4658.2012.08610.x.

Orth, J.D., Conrad, T.M., Na, J., Lerman, J.A., Nam, H., Feist, A.M., Palsson, B.Ø., 2011. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. Mol. Syst. Biol. 7, 535, http://dx.doi.org/10.1038/msb.2011.65.

Price, N.D., Reed, J.L., Palsson, B.Ø., 2004. Genome-scale models of microbial cells: evaluating the consequences of constraints. Nat. Rev. Microbiol. 2, 886–897, http://dx.doi.org/10.1038/nrmicro1023.

Qi, W.W., Vannelli, T., Breinig, S., Ben-Bassat, A., Gatenby, A.A., Haynie, S.L., Sariaslani, F.S., 2007. Functional expression of prokaryotic and eukaryotic genes in *Escherichia coli* for conversion of glucose to p-hydroxystyrene. Metab. Eng. 9, 268–276, http://dx.doi.org/10.1016/j.ymben.2007.01.002.

Qian, Z.G., Xia, X.X., Lee, S.Y., 2009. Metabolic engineering of *Escherichia coli* for the production of putrescine: a four carbon diamine. Biotechnol. Bioeng. 104, 651–662, http://dx.doi.org/10.1002/bit.22502.

Rathnasingh, C., Raj, S.M., 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, 729–739, http://dx.doi.org/10.1002/bit.22429.

Rathnasingh, C., Raj, S.M., Lee, Y., Catherine, C., Ashok, S., Park, S., 2012. Production of 3-hydroxypropionic acid via malonyl-CoA pathway using recombinant *Escherichia coli* strains. J. Biotechnol. 157, 633–640, http://dx.doi.org/10.1016/j.jbiotec.2011.06.008.

Russell, J., Cook, G., 1995. Energetics of bacterial growth: balance of anabolic and catabolic reactions. Microbiol. Mol. Biol. Rev. 59, 48–62.

Sanchez, A.M., Andrews, J., Hussein, I., Bennett, G.N., San, K.Y., 2006. Effect of overexpression of a soluble pyridine nucleotide transhydrogenase (UdhA) on the production of poly(3-hydroxybutyrate) in Escherichia coli. Biotechnol. Prog. 22, 420–425, http://dx.doi.org/10.1021/bp050375u.

Sauer, U., Canonaco, F., Heri, S., Perrenoud, A., Fischer, E., 2004. The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*. J. Biol. Chem. 279, 6613–6619, http://dx.doi.org/10.1074/jbc.M311657200.

Shen, C.R., Lan, E.I., Dekishima, Y., Baez, A., Cho, K.M., Liao, J.C., 2011. Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. Appl. Environ. Microbiol. 77, 2905–2915, http://dx.doi.org/10.1128/AEM.03034-10.

Stols, L., Donnelly, M., 1997. Production of succinic acid through overexpression of NAD^+-dependent malic enzyme in an *Escherichia coli* mutant. Appl. Environ. Microbiol. 63, 2695–2701.

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, 1628–1634, http://dx.doi.org/10.1128/AEM.02376-08.

Trinh, C.T., 2012. Elucidating and reprogramming *Escherichia coli* metabolisms for obligate anaerobic n-butanol and isobutanol production. Appl. Microbiol. Biotechnol. 95, 1083–1094, http://dx.doi.org/10.1007/s00253-012-4197-7.

Tseng, H.C., Harwell, C.L., Martin, C.H., Prather, K.L.J., 2010. Biosynthesis of chiral 3-hydroxyvalerate from single propionate-unrelated carbon sources in metabolically engineered *E. coli*. Microb. Cell Fact. 9, 96, http://dx.doi.org/10.1186/1475-2859-9-96.

Tseng, H.C., Martin, C.H., Nielsen, D.R., Prather, K.L.J., 2009. Metabolic engineering of *Escherichia coli* for enhanced production of (R)- and (S)-3-hydroxybutyrate. Appl. Environ. Microbiol. 75, 3137–3145, http://dx.doi.org/10.1128/AEM.02667-08.

Ui, S., Takusagawa, Y., Sato, T., Ohtsuki, T., Mimura, A., Ohkuma, M., Kudo, T., 2004. Production of L-2,3-butanediol by a new pathway constructed in *Escherichia coli*. Lett. Appl. Microbiol. 39, 533–537, http://dx.doi.org/10.1111/j.1472-765X.2004.01622.x.

Verho, R., Londesborough, J., Penttilä, M., Richard, P., 2003. Engineering redox cofactor regeneration for improved pentose fermentation in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 69, 5892, http://dx.doi.org/10.1128/AEM.69.10.5892.

Villadsen, J., Nielsen, J., Lidén, G., 2011. Chemicals from metabolic pathways. In: Bioreaction Engineering Principles. Springer, Boston, MA, USA, pp. 7–62 (Chapter 2). http://dx.doi.org/10.1007/978-1-4419-9688-6.

Wang, B., Wang, P., Zheng, E., Chen, X., Zhao, H., Song, P., Su, R., Li, X., Zhu, G., 2011. Biochemical properties and physiological roles of NADP-dependent malic enzyme in *Escherichia coli*. J. Microbiol. 49, 797–802, http://dx.doi.org/10.1007/s12275-011-0487-5.

Yan, Y., Lee, C.C., Liao, J.C., 2009. Enantioselective synthesis of pure (R,R)-2,3-butanediol in *Escherichia coli* with stereospecific secondary alcohol dehydrogenases. Org. Biomol. Chem. 7, 3914–3917, http://dx.doi.org/10.1039/b913501d.

Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., Khandurina, J., Trawick, J.D., Osterhout, R.E., Stephen, R., Estadilla, J., Teisan, S., Schreyer, H.B., Andrae, S., Yang, T.H., Lee, S.Y., Burk, M.J., Van Dien, S., 2011. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. Nat. Chem. Biol. 7, 445–452, http://dx.doi.org/10.1038/nchembio.580.

Zhang, X., Jantama, K., Moore, J.C., Shanmugam, K.T., Ingram, L.O., 2007. Production of L -alanine by metabolically engineered *Escherichia coli*. Appl. Microbiol. Biotechnol. 77, 355–366, http://dx.doi.org/10.1007/s00253-007-1170-y.

Zhu, G., Golding, G.B., Dean, A.M., 2005. The selective cause of an ancient adaptation. Science 307, 1279–1282, http://dx.doi.org/10.1126/science.1106974.