Journal of Biotechnology 251 (2017) 1–13

Contents lists available at ScienceDirect

Journal of Biotechnology

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

Genome-scale metabolic modelling common cofactors metabolism in microorganisms

Nan Xu[a,b,d], Chao Ye[a,c,d], Xiulai Chen[a,c,d], Jia Liu[a,c,d], Liming Liu[a,c,d,*]

[a] State Key Laboratory of Food Science and Technology, Jiangnan University, 1800 Lihu Road, Wuxi, Jiangsu 214122, China
[b] College of Bioscience and Biotechnology, Yangzhou University, Yangzhou, Jiangsu 225009, China
[c] Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, 1800 Lihu Road, Wuxi, Jiangsu 214122, China
[d] The Laboratory of Food Microbial-Manufacturing Engineering, Jiangnan University, Wuxi 214122, China

ARTICLE INFO

Keywords:
Acetyl-CoA
ATP
Cofactor function
Genome-scale modeling
NADH
NADPH

ABSTRACT

The common cofactors ATP/ADP, NAD(P)(H), and acetyl-CoA/CoA are indispensable participants in biochemical reactions in industrial microbes. To systematically explore the effects of these cofactors on cell growth and metabolic phenotypes, the first genome-scale cofactor metabolic model, *i*cmNX6434, including 6434 genes, 1782 metabolites, and 6877 reactions, was constructed from 14 genome-scale metabolic models of 14 industrial strains. The origin, consumption, and interactions of these common cofactors in microbial cells were elucidated by the *i*cmNX6434 model, and they played important roles in cell growth. The essential cofactor modules contained 2480 genes and 2948 reactions; therefore, improving cofactor biosynthesis, directing these cofactors into essential metabolic pathways, as well as avoiding cofactor utilization during byproduct biosynthesis and futile cycles, are three ways to increase cell growth. The effects of these common cofactors on the distribution and rate of the carbon flux in four universal modes, as well as an optimized metabolic flux, could be obtained by manipulating cofactor availability and balance. Significant changes in the ATP, NAD(H), NADP(H), or acetyl-CoA concentrations triggered relevant metabolic responses to acidic, oxidative, heat, and osmotic stress. Globally, the model *i*cmNX6434 provides a comprehensive platform to elucidate the physiological effects of these cofactors on cell growth, metabolic flux, and industrial robustness. Moreover, the results of this study are a further example of using a consensus genome-scale metabolic model to increase our understanding of key biological processes.

1. Introduction

According to http://www.ncbi.nlm.nih.gov/genome/browse/, 2330 microbial genome sequences have been completed, and the sequencing of 7556 microbial genomes is in progress. The increasing number of genome sequences has increased the amount of high-throughput data and biological knowledge. By combining omics data and laboratory-derived data, genome-scale metabolic models (GSMMs) have become very useful platforms and tools for understanding microbial physiology (Liu et al., 2010; Oberhardt et al., 2009), including understanding the diversity of organism-specific knowledge (Herrgard et al., 2008; Thiele et al., 2013), comparing the similarities and differences of microbes in terms of their phylogenetic distance (Oberhardt et al., 2011), resolving the metabolic characteristics of microbes with a similar industrial usage (Papini et al., 2012), and exploring the relationship between microorganisms (Stolyar et al., 2007; Ye et al., 2014). One hundred and sixty-nine GSMMs involving 116 microorganisms have been published since 1999, 21.3% of which were eukaryotic metabolic models and 78.7% of which were prokaryotic metabolic models. Although much progress regarding GSMMs has been achieved by modeling highly characterized organisms, such as *Escherichia coli* (Carrera et al., 2014; McCloskey et al., 2013) and *Saccharomyces cerevisiae* (Kim et al., 2012), the unique characteristics of GSMMs from lesser studied organisms make them more suitable for specific applications. Therefore, how to use GSMMs to decipher key biological issues at the systems level remains a major challenge.

A key biological issue is the metabolism of cofactors and its role in biological systems (Broderick, 2001). Cofactors can act as substrates, products, and/or catalysts of biochemical reactions, and they serve as the carriers of redox, electron, energy, and/or functional groups in anabolic and catabolic reactions. Cofactor engineering, an important branch of metabolic engineering, is accomplished mainly by changing the intracellular cofactor form and levels to manipulate metabolic fluxes for a particular metabolite or metabolic network. Such strategies

* Corresponding author at: State Key Laboratory of Food Science and Technology, Jiangnan University, 1800 Lihu Road, Wuxi, 214122, China.
E-mail address: mingll@jiangnan.edu.cn (L. Liu).

http://dx.doi.org/10.1016/j.jbiotec.2017.04.004
Received 21 January 2017; Received in revised form 2 April 2017; Accepted 3 April 2017
Available online 04 April 2017

include: (i) adding substrates with different oxidation potentials, activators, inhibitors, and/or competitors of coenzymes, precursors for cofactor biosynthesis, and cofactor structural analogs (Ji et al., 2011; Liu et al., 2006a; Liu et al., 2006b); (ii) overexpressing, deleting, or reducing the cofactor-related reaction(s) in the biosynthetic and degradation pathways of target products (Dulermo and Nicaud, 2011); (iii) maintaining cofactor balance in the precursor supply and product formation pathways (Du et al., 2012; Feng and Zhao, 2013; Ghosh et al., 2011; van Rossum et al., 2016; Zhuang et al., 2013); (iv) constructing heterologous regeneration pathways of cofactors (Han et al., 2013); (v) controlling cofactor levels by converting between different cofactors (Gameiro et al., 2013); (vi) eliminating redundant cofactor reactions (Tang et al., 2013); and (vii) optimizing cofactor specificity in oxidation-reduction reactions (Hoelsch et al., 2013). However, those engineering strategies did not always achieve the expect results, and some of them decreased cellular fitness and were accompanied by metabolic stress. These detrimental effects occurred because the influence of cofactors on cellular metabolism tends to be complicated and comprehensive, and the metabolism and behavior of cofactors during fermentation were less well known. Therefore, it is necessary to understand the origin, consumption, interaction, and function of common cofactors in microorganisms. If we have a clear understanding of the metabolic pattern of common cofactors, it will be possible to develop more efficient strategies to enhance the production of target metabolites and to improve the economic efficiency of industrial biotechnology applications.

Fortunately, the number of GSMMs is increasing rapidly, and their quality is improving as well. If the GSMMs from 116 microorganisms were used to construct a consensus metabolic model, the resulting model should include 115,731 genes, 7570 metabolites, and 24,312 biochemical reactions. We believe that this will provide a systems level platform for understanding the metabolic characteristics of common cofactors in microorganisms. These highly-curated GSMMs are the basis for constructing the consensus platform. Besides, integrating approach required re-annotating cofactor-related gene-protein-reaction associations for the comprehensiveness, unifying and refining all the chemical entities for the consistency, and filling metabolic gaps for the connectivity and functionality. Based on these, we constructed the first genome-scale cofactor metabolic model, *i*cmNX6434, which includes the cofactors ATP/ADP, NADH/NAD, NADPH/NADP, and acetyl-CoA/CoA. Using the *i*cmNX6434 model, we determined the production, consumption, and interactions of these common cofactors. Furthermore, the physiological effects of these cofactors on cell growth, carbon flux, and industrial robustness were highlighted.

2. Materials and methods

2.1. Choice of organisms

The model microorganisms *E. coli*, *Bacillus subtilis*, *S. cerevisiae* and *Aspergillus niger* were considered first. Then, producers of some typical fermentation products, including various organic acids, alcohols, ketones, lipids, amino acids, vitamins, antibiotics, and other primary and secondary metabolites were included. The chosen seven prokaryotic and seven eukaryotic microorganisms were all heterotrophic and included obligate anaerobes, facultative anaerobes, and aerobes.

2.2. Model construction

Construction of metabolic cofactor models generally includes three steps. In the first step, a basic skeleton of metabolic cofactor models was formed from metabolic cofactor subnetworks in the 14 organism-specific GSMMs, and new gene-protein reaction relations were acquired by KEGG Automatic Annotation Server annotation (Moriya et al., 2007). Moreover, these newly annotated genes and original genes served as the local libraries for BLASTP searches with a filter identity $\geq 40\%$, match lengths $\geq 70\%$ of the length of both the subject and query sequences, and an *e*-value $\leq 1 \times 10^{-5}$ for prokaryotes and $\leq 1 \times 10^{-30}$ for eukaryotes. Transport reactions were annotated and classified from the TCDB. In the second step, a unified form of the model contents was made as follows. (i) Metabolites were marked with entries from the public database KEEG compounds, Chemical Entries of Biological Interest, the PubChem Compound Database, the SEED database, and ChemSpider. (ii) Reactions were normalized by our in-house script after deleting water, hydrogen, and coefficients of participants. In addition, reaction compartments were predicted by the Covariance Estimation and Learning through Likelihood Optimization (Yu et al., 2006) and WoLF PSORT (Horton et al., 2007) algorithms, and the reaction direction was determined, in turn, from the original GSMMs, MetaCyc (Caspi et al., 2014), and Biopath (Reitz et al., 2004). (iii) Metabolic subsystems that were not in accordance in different GSMMs were modified according to the KEGG pathway. In the third step, the functional connectivity of the cofactor models was considered from various perspectives: (i) their own biomass equations for eukaryotic, Gram-positive, and Gram-negative microorganisms should contain the relevant biomass components of as many of the 14 microorganisms as possible; (ii) non-cofactor reactions in the reduced models of the three major types of microbes were used to fill metabolic gaps; and (iii) the *in silico* growth of the three major types of microbial cofactor models was simulated as constraints of the minimal synthetic medium. Steps (ii) and (iii) formed an iterative debug until the three major types of microbial cofactor models could support their own biomass synthesis.

2.3. Simulation constraints

The effects of the cofactors on cell growth in minimal synthetic medium were simulated using three major types of microbial cofactor models, with glucose as the carbon source and without considering genetic characteristics. The effects of the cofactors on metabolic phenotypes were simulated using three major types of microbial cofactor models with genetic characteristics under reasonable constraints of the fermentation parameters in the organism-specific GSMMs. These fitted fermentation parameters included mainly: for *C. beijerinckii*, the specific growth rate (μ), the glucose uptake rate, acetate production (consumption) rate, and butyrate, butanol, and acetone production rates in the early growth stages (1 h), the acidogenesis phase (3 h), the logarithmic growth phase (7 h), and the solventogenesis phase (18 h) (Shi and Blaschek, 2008); for *Candida glabrata*, μ, the glucose uptake rate, ethanol, glycerol, and carbon dioxide production rates at the maximum cell growth under 40%–50% dissolved oxygen (Hua et al., 2001); for *Y. lipolytica*, μ, citrate and lipid production rates of 11 sampling points using glucose as the carbon source (Morin et al., 2011), and μ, citrate, and isocitrate production rates, as well as triacylglycerol consumption rates, during the growth phase (6 h), transition phase (13 h), and nitrogen limitation phase (23 h) using waste oil as the carbon source (Liu et al., 2015). These cell growth rates were changed to constraints when simulating metabolite production.

2.4. Modeling methods

Cofactor-related concepts in the *i*cmNX6434 model were introduced as follows. (i) The definition of the cofactor current was referred to as the metabolite turnover rate (Chung and Lee, 2009), and it was calculated by the reaction flux of the cofactor reaction multiplied by the stoichiometric coefficients in each reaction; the coefficients were negative if the cofactor was a reactant and vice versa. (ii) The accumulation of cofactors was equal to the sum of the cofactor current in each reaction. (iii) The production and consumption of cofactors were investigated in both the thermodynamic direction and the flow direction of the reactions using 14 organism-specific genetic constraints. The synthetic reaction shared by more than 10 organism-