Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Computational methods in metabolic engineering for strain design

Matthew R Long[1,3], Wai Kit Ong[1,2,3] and Jennifer L Reed[1,2]

Metabolic engineering uses genetic approaches to control microbial metabolism to produce desired compounds. Computational tools can identify new biological routes to chemicals and the changes needed in host metabolism to improve chemical production. Recent computational efforts have focused on exploring what compounds can be made biologically using native, heterologous, and/or enzymes with broad specificity. Additionally, computational methods have been developed to suggest different types of genetic modifications (e.g. gene deletion/addition or up/down regulation), as well as suggest strategies meeting different criteria (e.g. high yield, high productivity, or substrate co-utilization). Strategies to improve the runtime performances have also been developed, which allow for more complex metabolic engineering strategies to be identified. Future incorporation of kinetic considerations will further improve strain design algorithms.

Addresses
[1] Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, United States
[2] Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI, United States
[3] These authors contributed equally to this work.

Corresponding author: Reed, Jennifer L (reed@engr.wisc.edu)

Current Opinion in Biotechnology 2015, **34**:135–141

This review comes from a themed issue on **Systems biology**

Edited by **Sarah Maria Fendt** and **Costas D Maranas**

http://dx.doi.org/10.1016/j.copbio.2014.12.019

0958-1669/

Introduction

Cells can be used industrially to make a wide variety of products, including food and beverages, therapeutics, enzymes, commodity chemicals, and specialty chemicals. Many organisms can be manipulated to enhance chemical production or improve cellular properties by modifying genetics, media composition, and/or reactor operating conditions. Metabolic engineering primarily uses genetic approaches to adjust metabolic and regulatory networks in biocatalysts to enhance production of desired chemicals. These desired chemicals can be natural biological molecules (e.g. amino acids, antibiotics, and ethanol) or molecules that are not normally produced biologically (e.g. 1,4-butanediol [1••] and 3,4-dihydroxybutyric acid [2]). Regardless of whether the molecule is natural or non-natural, product biosynthesis pathways need to be identified and the metabolic network needs to be manipulated to increase flux towards the product of interest. Metabolic engineering can involve increasing precursor production, recycling cofactors (e.g. NAD(P)H), improving substrate utilization, removing bottlenecks in biosynthesis pathways, eliminating feedback regulation, and/or improving tolerance to products and their intermediates.

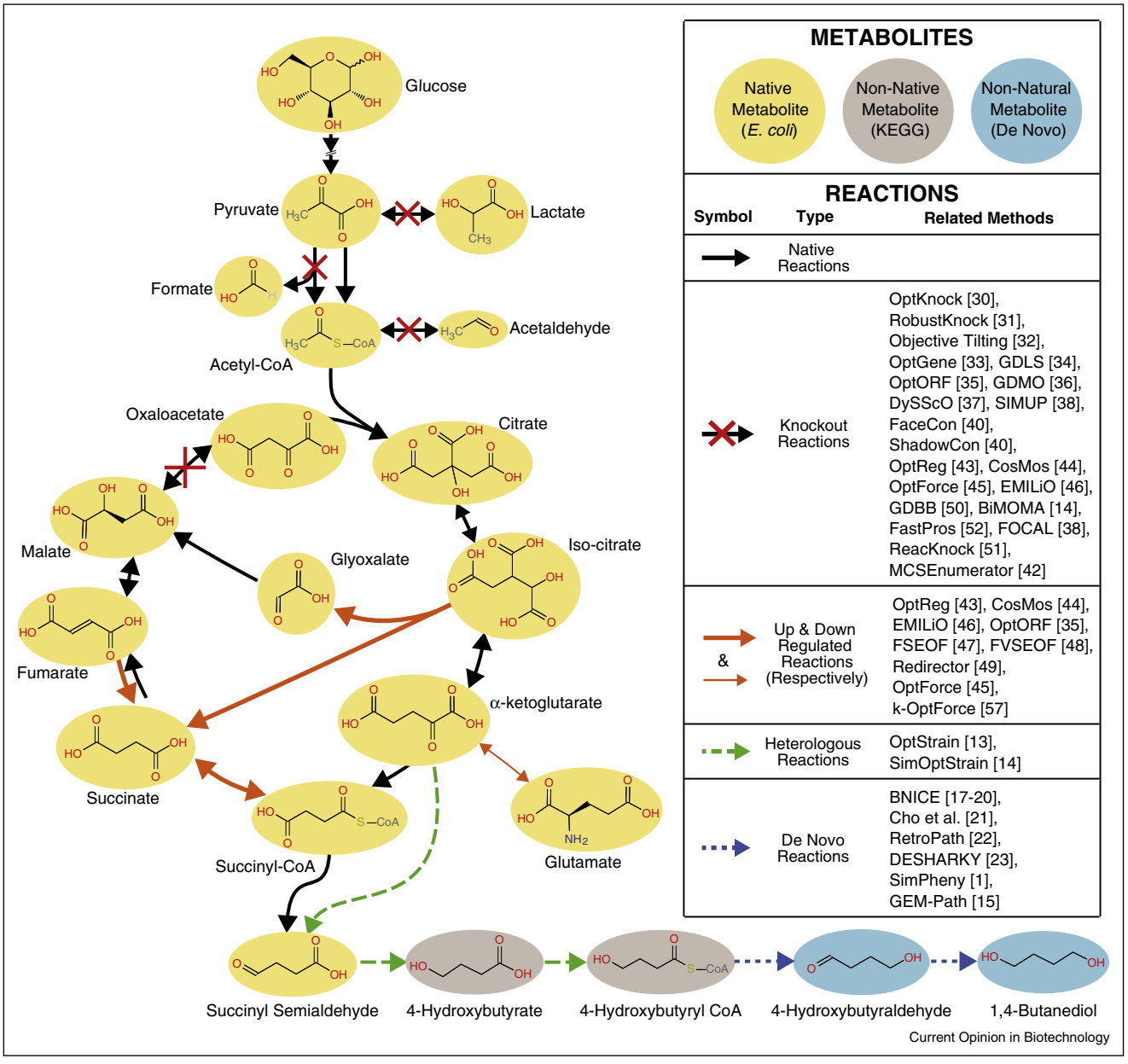
To facilitate metabolic engineering efforts, a variety of computational tools have been developed to identify genetic strategies (i.e. strain designs) for enhancing chemical production. Many of these tools model the integration of product biosynthesis pathways with host metabolism to predict how metabolic and regulatory changes will affect chemical production yields, titers, and productivity. Genome-scale constraint-based models have been useful in metabolic engineering since they can account for the entire metabolic (and regulatory) capabilities of a cell and do not require detailed kinetic parameters to generate them. Constraint-based models can be quickly generated by a variety of software packages (reviewed in [3]) using an organism's genomic, biochemical, and physiological data. All constraint-based models use two types of constraints to define a metabolic solution space: steady-state mass balances (which equate metabolite production and consumption rates), and flux bounds (which set maximum and minimum limits for metabolic fluxes based on an enzyme's capacity and directionality) [4]. Here, we discuss recent advances in biosynthesis pathway identification algorithms and constraint-based modeling approaches that can be used in metabolic engineering for strain design (Figure 1).

Identifying biosynthesis pathways

Pathways involving known reactions

A number of databases exist that describe the reactions that can be catalyzed by known enzymes, including KEGG, BioCyc, BRENDA, MetRxn, and SEED [5–9]. These databases can be used to develop organism-specific reaction networks (based on an organism's genome annotation), as well as comprehensive reaction networks (containing known biochemical reactions for all organisms). Native and non-native pathways that lead to product formation can be identified from these reaction networks using graphical and modeling approaches. While graphical approaches can identify shortest paths to a desired product, these paths might violate steady-state mass balances

Figure 1



Overview of strain design algorithms. Computational methods have been developed to identify pathways to enable production of desired compounds using known native (solid black and orange arrows) and known non-native (dashed green arrows) reactions, as well as *de novo* reactions (dotted blue arrows) that have not been known to occur biologically. Additional host metabolic changes can be suggested by a variety of constraint-based modeling algorithms, which can involve deletions (red crosses) and up/down regulation (orange arrows). The example *de novo* pathways and metabolic modifications shown are for illustrative purposes and were reported previously for 1,4-butanediol and succinate production [1••,44,45]. Chemical structures were generated using Open Babel [58].

(indicating that reactions used might consume or produce a metabolite that cannot be synthesized or degraded, respectively) [10]. Methods that include steady-state mass balances can be used to find mass-balanced paths [10], and constraint-based models can identify paths that are the shortest and/or have the highest maximum yield [11,12].

A variety of constraint-based modeling approaches can identify biosynthesis pathways involving native and/or non-native pathways by analyzing organism-specific reaction networks augmented with all known reactions. Flux balance analysis (FBA) can identify mass-balanced pathways with the highest yields [12]. OptStrain [13] and