Metabolic Engineering Communications 1 (2014) 1–11

Contents lists available at ScienceDirect

Metabolic Engineering Communications

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

Expanding metabolic engineering algorithms using feasible space and shadow price constraint modules

Christopher J. Tervo, Jennifer L. Reed *

Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA

ARTICLE INFO

Article history:
Received 2 June 2014
Accepted 25 June 2014
Available online 15 August 2014

Keywords:
Flux balance analysis
Constraint-based model
Strain design
OptKnock

ABSTRACT

While numerous computational methods have been developed that use genome-scale models to propose mutants for the purpose of metabolic engineering, they generally compare mutants based on a single criteria (e.g., production rate at a mutant's maximum growth rate). As such, these approaches remain limited in their ability to include multiple complex engineering constraints. To address this shortcoming, we have developed feasible space and shadow price constraint (FaceCon and ShadowCon) modules that can be added to existing mixed integer linear adaptive evolution metabolic engineering algorithms, such as OptKnock and OptORF. These modules allow strain designs to be identified amongst a set of multiple metabolic engineering algorithm solutions that are capable of high chemical production while also satisfying additional design criteria. We describe the various module implementations and their potential applications to the field of metabolic engineering. We then incorporated these modules into the OptORF metabolic engineering algorithm. Using an *Escherichia coli* genome-scale model (iJO1366), we generated different strain designs for the anaerobic production of ethanol from glucose, thus demonstrating the tractability and potential utility of these modules in metabolic engineering algorithms.

1. Background

Genome-scale models (GEMS) are powerful tools allowing for the prediction of cellular growth, flux profiles, and mutant strain phenotypes (Orth et al., 2010). Over the last decade, with the development of new computational algorithms, GEMS have been used to guide the design of strains for biochemical production, such as biofuels and commodity chemicals (reviewed in Curran and Alper (2012), Zomorrodi et al. (2012), and Lee et al. (2012)). While GEMs are valuable tools, new computational algorithms are still needed to evaluate them and apply them in new ways.

Many strain design algorithms exist that identify which network modifications are needed to improve chemical production. These modifications can involve reaction deletions (OptKnock), metabolic or regulatory gene deletions (OptGene and OptORF), reaction additions (OptStrain and SimOptStrain), or flux increases/decreases (OptReg, OptForce, CosMos, FSEOF) (Zomorrodi et al., 2012; Kim and Reed, 2010; Burgard et al., 2003; Pharkya et al., 2004; Ranganathan et al., 2010; Patil et al., 2005; Pharkya and Maranas, 2006; Cotten and Reed, 2013; Choi et al., 2010; Kim et al., 2011). The bi-level optimization approaches used to identify these modifications can be computationally expensive and recent efforts have improved their run-time performances (Patil et al., 2005; Kim et al., 2011; Ohno et al., 2013; Lun et al., 2009; Yang et al., 2011). Many of these metabolic engineering algorithms focus on improving the desired chemical production when the proposed mutant is operating at its maximal growth rate. By coupling chemical production to growth, selection for growth rate using a chemostat or sequential batch cultures can enrich for strains with increased chemical production (Fong et al., 2005). One such algorithm, OptORF, is used extensively in this work (Kim and Reed, 2010). The OptORF algorithm extends upon OptKnock by using gene rather than reaction deletions as potential modifications. By accounting for gene and transcriptional regulatory network information, OptORF proposes deleting or overexpressing metabolic or regulatory genes (as opposed to reaction level deletions proposed by OptKnock) to increase chemical production. By doing this, OptORF avoids designs that would be impossible to implement, due to genetic interactions between reactions or regulatory effects.

While metabolic engineering methods have been successful (Curran and Alper, 2012; Ranganathan et al., 2010; Fong et al., 2005; Yim et al., 2011), most of these approaches cannot consider the ramifications of undesirable suboptimal flux distributions

* Corresponding author. Tel.: +1 608 262 0188; fax: +1 608 262 5434.
E-mail address: reed@engr.wisc.edu (J.L. Reed).

http://dx.doi.org/10.1016/j.meteno.2014.06.001

(e.g. production with low productivity) (Patil et al., 2005; Feist et al., 2010; Lin et al., 2005; Sánchez et al., 2005; Vadali et al., 2005), or production phenotypes at or near stationary phase in batch cultures. Additionally, these algorithms are limited in their ability to tailor a strain's behavior to address more complex problems (e.g., the co-utilization of multiple substrates (Gawand et al., 2013; Lian et al., 2014; Trinh et al., 2008) or elimination of undesirable by-products (Aristidou et al., 1994; Eiteman and Altman, 2006; Jantama et al., 2008; Zha et al., 2009)). Consequently, while these approaches are valuable in designing adaptive evolutionary strains based on single criteria (e.g., high production at maximal growth rates), they often lack the ability to efficiently propose strains meeting multiple design criteria that are of interest to investigators. To address these problems in small networks, techniques such as constrained minimal cut sets (Hädicke and Klamt, 2011) can be used to allow researchers to meet additional design criteria (e.g., elimination of undesired by-products) without affecting the desired chemical production phenotype. Recent advances allow enumeration of the smallest minimal cut sets in genome-scale networks, from which constrained minimal cut sets can be identified (von Kamp and Klamt, 2014). However, all minimal cut sets can still not be enumerated for genome-scale networks, and the smallest minimal cut sets identified first might not correspond to constrained minimal cut sets meeting additional design criteria. Additionally, strategies for finding constrained minimal cut sets that consider transcriptional regulation, media selection or degree of coupling between biomass and chemical production have not been developed.

Previously, we developed the forced coupling algorithm (FOCAL) to identify conditions (e.g., gene deletions or media conditions) that ensure directional coupling between two fluxes (flux through v_x implies flux through v_y) (Tervo and Reed, 2012). By changing media conditions or deleting genes, FOCAL affects the shape of the resulting feasible solution space. We also showed how FOCAL can be modified to design a mutant strain that must co-utilize xylose and glucose simultaneously in order to grow. While these modifications were interesting, they did not work to increase the overall productivity of the organism since no metabolic engineering objective was included. Moreover, this approach could only enforce directional coupling between fluxes which is often an overly stringent condition for metabolic engineering strain designs.

Recently, Ohno et al. (2013) used shadow prices from flux balance analysis (FBA) solutions to guide a greedy algorithm for increasing chemical productivity as reaction deletions are added. Double deletion mutants with the top desired shadow prices (which indicate the rate of change in growth divided by the rate of change in chemical production) were used as "parent" strains to find triple deletion knockouts with the best shadow prices. This greedy search process, called FastPros, was repeated for up to 25 knockouts, and for each iterative screening step, any sets of deletions which resulted in a non-negative shadow prices (indicating coupling between growth and chemical production) were stored as candidates for further analysis and excluded from further screening. The authors then used OptKnock to maximize chemical production using only the stored reaction knockouts found by their FastPros process. Because they use a greedy algorithm, their method does not guarantee that the set of knockouts with the highest shadow prices are discovered. Additionally, since the authors use OptKnock to propose strain designs, their approach does not control or optimize the degree of coupling between chemical production and cellular growth when mutants are proposed.

Here, we have developed modules *F*easible Sp*ace Con*straint (FaceCon) and *Shadow Con*straint (ShadowCon) modules for controlling the shape an organism's feasible space. These modules allow many additional types of design criteria to be considered besides directional coupling. These modules can be easily added to mixed integer linear adaptive evolution metabolic engineering algorithms to incorporate additional design criteria, while retaining the original objective of the method (e.g., coupling growth and chemical production). Since there are often many possible solutions to these strain design algorithms, embedding these modules allows only the subset of those mutants to be found if the criteria associated with these modules is met. Such filtering is needed as models become larger and the computational cost (i.e., CPU time) of generating numerous strain designs increases, due to the combinatorial explosion associated with increasing numbers of integer variables and integer cuts needed to find alternate solutions. To date, the only type of filtering that can be done works to prevent finding solutions that have large ranges of chemical production at the maximum growth rate (Feist et al., 2010; Tepper and Shlomi, 2010).

FaceCon modules are included as additional inner optimization problems and ensure that any proposed mutant cannot operate within a user-defined region (i.e., no feasible flux distribution can exist within a user-defined region). By defining this excluded region, various feasible space characteristics can be enforced. Below we describe three FaceCon modules:

1. *Coupling module*: This module allows a researcher to enforce different types of coupling (directional or weak) between a flux of interest (v_y) and another flux (v_x) depending on the formulation and parameter selection. This module can be used to find mutants with directional coupling (i.e., flux through v_x implies flux through v_y for all values of v_x (Burgard et al., 2004)) or weak coupling (where flux through v_x implies flux for v_y only for some positive values of v_x). Depending on how the coupling module is implemented one can require mutants having directional coupling, weak coupling, or either directional or weak coupling. The result of any of these implementations is that a defined portion of the v_x axis is excluded from the solution space of a proposed mutant.
2. *Chemical level module*: The chemical level module ensures proposed mutants meet criteria associated with the production level of a chemical of interest, v_y (e.g., a desired product or undesired by-product). This module finds mutants whose solution space excludes solutions with v_y below (or above) a user-defined threshold (β) within a defined region (e.g., v_y must be greater than β when v_x is greater than v^{min}).
3. *Direct constraint module*: This module is the most comprehensive and with proper parameter selection can encompass the functions of the two previous FaceCon modules. This module allows the user to define a particular region that must be excluded from the solution space of any proposed mutant; thus, the researcher is able to directly influence the solution space of any mutant proposed by a metabolic engineering algorithm.

In the following sections, we detail the application, function and relevant parameters for each of these FaceCon modules. We then introduce the concept of *shadow con*straint (ShadowCon) modules, which can be used to control the degree of coupling once coupling between two fluxes occurs. To illustrate each module's functionality and potential use, we have included the FaceCon and ShadowCon modules as additional inner problems within the OptORF algorithm, to find metabolic gene deletions that couple growth and chemical production and that satisfy additional module criteria. Additionally, to demonstrate the methods are applicable on genome-scale networks we have applied them to identify mutants for ethanol production using the *Escherichia coli* model, iJO1366 (Orth et al., 2011). We demonstrate that when