Metabolic Engineering 32 (2015) 220–231

Contents lists available at ScienceDirect

Metabolic Engineering

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

Rational design of efficient modular cells

Cong T. Trinh [a,b,c,d,e,*], Yan Liu [a], David J. Conner [a]

[a] Department of Chemical and Biomolecular Engineering, United States
[b] UTK-ORNL Joint Institute of Biological Science, United States
[c] Bredesen Center for Interdisciplinary Research and Graduate Education, United States
[d] Institute of Biomedical Engineering, The University of Tennessee, Knoxville, TN, United States
[e] BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN, USA

ARTICLE INFO

Article history:
Received 27 January 2015
Received in revised form
7 May 2015
Accepted 14 October 2015
Available online 21 October 2015

Keywords:
Modular cell
MODCELL
Minimal metabolic functionality
Elementary mode analysis
Alcohols
Esters

ABSTRACT

The modular cell design principle is formulated to devise modular (chassis) cells. These cells can be assembled with exchangeable production modules in a plug-and-play fashion to build microbial cell factories for efficient combinatorial biosynthesis of novel molecules, requiring minimal iterative strain optimization steps. A modular cell is designed to be auxotrophic, containing core metabolic pathways that are necessary but insufficient to support cell growth and maintenance. To be functional, it must tightly couple with an exchangeable production module containing auxiliary metabolic pathways that not only complement cell growth but also enhance production of targeted molecules. We developed a MODCELL (modular cell) framework based on metabolic pathway analysis to implement the modular cell design principle. MODCELL identifies genetic modifications and requirements to construct modular cell candidates and their associated exchangeable production modules. By defining the degree of similarity and coupling metrics, MODCELL can evaluate which exchangeable production module(s) can be tightly coupled with a modular cell candidate. We first demonstrated how MODCELL works in a step-by-step manner for example metabolic networks, and then applied it to design modular *Escherichia coli* cells for efficient combinatorial biosynthesis of five alcohols (ethanol, propanol, isopropanol, butanol and isobutanol) and five butyrate esters (ethyl butyrate, propyl butyrate, isopropyl butyrate, butyl butyrate and isobutyl butyrate) from pentose sugars (arabinose and xylose) and hexose sugars (glucose, mannose, and galactose) under anaerobic conditions. We identified three modular cells, MODCELL1, MODCELL2 and MODCELL3, that can couple well with Group 1 of modules (ethanol, isobutanol, butanol, ethyl butyrate, isobutyl butyrate, butyl butyrate), Group 2 (isopropanol, isopropyl butyrate), and Group 3 (propanol, isopropanol), respectively. We validated the design of MODCELL1 for anaerobic production of ethanol, butanol, and ethyl butyrate using experimental data available in literature.

1. Introduction

Developing efficient and robust microbial biocatalysts plays a key role in producing targeted molecules at high yields, titers, and productivities, which determines the overall efficiency of a biochemical process (Schmid et al., 2001). The current challenge to the conventional strain engineering approach is that it requires many iterative trial-and-error strain optimization steps to engineer a desirable strain. It can take seven or more years to engineer an optimal strain to produce a targeted molecule such as 1,3-PDO (Nakamura and Whited, 2003) and artemisinin (Westfall et al., 2012) for commercialization. The entire process must be repeated to engineer new optimal strains for producing other targeted chemicals, which is laborious and expensive. Therefore, it is transformative to develop a novel method to build desirable microbial cell factories systematically and rapidly from exchangeable, modular, biological parts by requiring minimal strain optimization steps.

The design concept of a complex system based on compatible and modular parts has been widely applied in every aspect of modern society, from manufacturing cars to building computers and software packages. Such a design enables a complex system to be built in an efficient, cost-effective, and systematic manner. It has been suggested that cellular metabolisms have evolved to exhibit modular organization to maximize their fitness (Almaas et al., 2004, 2005; Wagner et al., 2007). Driven by the modular design, the field of synthetic biology has emerged rapidly in recent years, and seeks to develop reusable, compatible, robust, and programmable biological parts assembled to build devices and systems that can be readily plugged in a chassis cell (Endy, 2005). Currently, such an approach has not yet been fully investigated at

* Corresponding author at: Department of Chemical and Biomolecular Engineering, United States.
E-mail address: ctrinh@utk.edu (C.T. Trinh).

http://dx.doi.org/10.1016/j.ymben.2015.10.005
1096-7176/

the cell level (Andrianantoandro et al., 2006) where a wildtype strain is commonly used as a chassis cell and may not couple well with synthetic devices or systems. One can envision the development of a universal optimal modular cell that, when combined with optimized exchangeable production modules, create desirable microbial cell factories to synthesize products of interest in a systematic and rapid manner without requiring many iterative strain optimization steps. Different from the wildtype, the modular cell is designed to have a very distinctive and advantageous characteristic in that its metabolism is reprogrammed to tightly couple with a diverse set of exchangeable and tunable production modules for efficient production of targeted molecules.

Flux balance analysis (FBA) (Fell and Small, 1986; Varma and Palsson, 1993) and elementary mode analysis (EMA) (Schuster et al., 2000, 1994) are the two common metabolic pathway analysis tools in the constraints-based metabolic network modeling that have been proven useful in guiding rational strain design for overproducing targeted molecules (Lewis et al., 2012; Trinh et al., 2009). Predictive FBA-based techniques such as OptKnock (Burgard et al., 2003), MOMA (Segre et al., 2002), OptGene (Patil et al., 2005), OptStrain (Pharkya et al., 2004), OPTREG (Pharkya and Maranas, 2006), OPTORF (Ranganathan et al., 2010), BiMOMA (Kim et al., 2011), SimOptStrain (Kim et al., 2011), EMILiO (Yang et al., 2011), Robustknock (Tepper and Shlomi, 2010), and ReacKnock (Xu et al., 2013), and predictive EMA-based techniques such as MMF (Trinh et al., 2008), cMCS (Hädicke and Klamt, 2011), CASOP (Hädicke and Klamt, 2010), FluxDesign (Melzer et al., 2009), and SMET (Flowers et al., 2013) have been developed to identify genetic modifications for rational strain design with many successful reports from production of primary to secondary metabolites. However, these techniques have not yet been developed for designing modular cells.

In this study, we formulated the modular cell design principle to devise modular cells. These cells can be assembled with exchangeable production modules to build microbial cell factories for efficient combinatorial biosynthesis of novel molecules in plug-and-play fashion, requiring minimal iterative strain optimization steps. We developed a modular cell (MODCELL) framework based on metabolic pathway analysis to implement the modular cell design principle, and applied MODCELL to design modular cells for two case studies. The first case study used example networks to illustrate how MODCELL works in a step-by-step manner. The second investigated *E. coli* metabolic networks for alcohol and ester production from pentoses and hexoses. We validated the design of a modular cell for the latter case with experimental data available in literature.

2. Materials and methods

2.1. Mass action of a metabolic network

Given a metabolic network with m metabolites and n reactions, the principal law of mass conservation for intracellular metabolites can be written as follows:

$$\frac{dC}{dt} = S \bullet r - \mu C \quad (1)$$

where $\mathbf{C}$ $(m \times 1)$ is the concentration vector; $\mathbf{S}$ $(m \times n)$ is the stoichiometric matrix; $\mathbf{r}$ $(n \times 1)$ is the reaction rate (flux) vector; and μ is the dilution factor. At quasi steady state with the negligible dilution factor, Eq. (1) becomes:

$$S \bullet r = 0 \quad (2)$$

$$r_{irr} \geq 0 \quad (3)$$

Inequality 3 is the thermodynamic constraint requiring fluxes of irreversible reactions (r_{irr}) to be either equal or greater than zero. Solutions of Eq. (2) constrained by inequality 3 constitute a polyhedral flux cone (Schuster et al., 1994).

2.2. Elementary mode analysis (EMA)

EMA calculates all admissible solutions of Eq. (2) that can span the flux cone by imposing an additional non-decomposability constraint. Each solution is called an elementary mode (EM) containing a minimal and unique set of reactions. For any given EM_1 and EM_2, the non-decomposability constraint specifies that $S(EM_1)$ is not a subset of $S(EM_2)$ where S(EM) is a set of indices of non-zero fluxes in EM. A cell can use any non-negative combination of these EMs to function for a given operating condition. In this study, EMs were calculated by using Metatool (von Kamp and Schuster, 2006).

2.3. MODCELL design and implementation

2.3.1. Formulating modular cell design principle

Motivated by the modular organization of cellular metabolisms (Almaas et al., 2004, 2005; Barabási and Albert, 1999; Trinh, 2012; Wagner et al., 2007), we formulated the general modular cell design principle as follows: a modular cell is designed to be auxotrophic and contain core metabolic pathways that are necessary but insufficient to support cell growth and maintenance. To grow and efficiently produce targeted molecules, the modular cell must be coupled with exchangeable production modules, e.g., auxiliary metabolic pathways synthesizing targeted molecules. The modular cell is auxotrophic to growth due to cofactor imbalance and insufficient supply (or lack) of precursor metabolites required for biomass synthesis. The coupling design between the modular cell and exchangeable production modules can facilitate systematic and rapid construction of desirable microbial cell factories for efficient production of targeted molecules, and require minimal iterative strain optimization steps.

2.3.2. Implementing modular cell design principle

We developed a MODCELL framework to design modular cells based on the modular cell design principle defined above. The MODCELL workflow has four steps: step 1: input metabolic networks; step 2: calculate phenotypic spaces; step 3: constrain desirable phenotypic spaces by identifying all minimal deleted reaction sets (MDRSs) for each network; and step 4: identify common deleted reaction sets (CDRSs) to design modular cells (Fig. 1).

Step 1*: Input metabolic networks*. This step constructs metabolic networks used for determining feasible phenotypic spaces. These networks contain different auxiliary pathways required to synthesize targeted molecules (Fig. 1A). Metabolic networks investigated in this study include 3 example networks (Fig. 2, Supplementary Table S1) and 50 alcohol/ester metabolic networks (Fig. 3, Supplementary Table S2).

Step 2*: Calculate phenotypic spaces*. Elementary mode analysis (EMA) is used to calculate all possible elementary modes (EMs) for each metabolic network. From the complete set of EMs computed, yields of species *i* with respect to a substrate *S* (r_i/r_S) can be calculated for each EM and used to determine the phenotypic space for a network and efficient pathways having the highest yields of targeted molecules (Fig. 1B).

Step 3*: Constrain desirable phenotypic spaces*. Minimal metabolic functionality (MMF), a reaction intervention problem, is applied to identify all possible minimal deleted reaction sets (MDRSs) for each metabolic network that can destroy all EMs except a