



Research review paper

Modular design: Implementing proven engineering principles in biotechnology

Sergio Garcia^{a,b}, Cong T. Trinh^{a,b,*}^a Department of Chemical and Biomolecular Engineering, University of Tennessee, Knoxville, TN, United States of America^b Center for Bioenergy Innovation, Oak Ridge National Laboratory, Oak Ridge, TN, United States of America

ARTICLE INFO

Keywords:

Modular design
Modularity
Modular cell
Modular cell engineering
ModCell
Systems biology
Metabolic engineering
Synthetic biology
Robustness
Evolvability
Networks
Pareto optimality
Industrialization of biology
Microbial biocatalysis

ABSTRACT

Modular design is at the foundation of contemporary engineering, enabling rapid, efficient, and reproducible construction and maintenance of complex systems across applications. Remarkably, modularity has recently been discovered as a governing principle in natural biological systems from genes to proteins to complex networks within a cell and organism communities. The convergent knowledge of natural and engineered modular systems provides a key to drive modern biotechnology to address emergent challenges associated with health, food, energy, and the environment. Here, we first present the theory and application of modular design in traditional engineering fields. We then discuss the significance and impact of modular architectures on systems biology and biotechnology. Next, we focus on the very recent theoretical and experimental advances in modular cell engineering that seeks to enable rapid and systematic development of microbial catalysts capable of efficiently synthesizing a large space of useful chemicals. We conclude with an outlook towards theoretical and practical opportunities for a more systematic and effective application of modular cell engineering in biotechnology.

1. Introduction

Complex engineered systems such as computers, vehicles, or factories can be assembled from exchangeable units of self-contained functionality known as modules. Modular design enables efficient production, maintenance, and customization across modern engineering technologies. The inception of modular design has had a revolutionary impact on many industries. For instance, the first modular computer, named IBM System/360 and built in the 1960s, allowed to use the same software for different application-dependent hardware, shaping information technology as we know it today. Undoubtedly, modular design will continue to drive innovation in both established and emergent fields of engineering.

Among trending engineering disciplines, biotechnology is encompassing far-reaching applications driven by the recent development of enabling technologies in interdisciplinary areas of genome engineering (Barrangou and Doudna, 2016), systems and synthetic biology (Kahl and Endy, 2013), metabolic engineering (Nielsen and Keasling, 2016), and bioprocessing (Cramer and Holstein, 2011; Olson et al., 2012).

Amid many applications to address issues related to health, energy, and the environment, the chemical industry will benefit from metabolic reprogramming of microbes as cell factories to catalyze the synthesis of therapeutics, chemicals, and fuels, from renewable and sustainable feedstocks (e.g., lignocellulosic biomass, sugar cane) or waste products (e.g., waste gas from steel manufacturing, plastic waste). Even though there exists a naturally large space of molecules that can be synthesized by metabolically engineered microorganisms (Lee et al., 2019), fewer than a dozen molecules are industrially produced (Nielsen and Keasling, 2016). A major roadblock is attributed to the very laborious and costly strain engineering process partly arising from the lack of standardization and repetition of genetic manipulation tasks (King et al., 2017; Winkler et al., 2015). Recently, modular cell engineering has been proposed as an innovative approach to accelerate strain engineering, harnessing a large space of molecules derived from rich and diverse cellular metabolism and thus pushing whole-cell biocatalysis towards an industrially competitive technology (Trinh and Mendoza, 2016).

In this paper, we first examine the theory and application of

* Corresponding author at: Department of Chemical and Biomolecular Engineering, University of Tennessee, 1512 Middle Drive, Dougherty 432, Knoxville, TN 37996, United States of America.

E-mail address: ctrinh@utk.edu (C.T. Trinh).

<https://doi.org/10.1016/j.biotechadv.2019.06.002>

Received 11 February 2019; Received in revised form 23 April 2019; Accepted 4 June 2019

Available online 07 June 2019

0734-9750/ © 2019 Elsevier Inc. All rights reserved.

modular design in conventional engineering disciplines, such as mechanical, chemical, nuclear, and civil engineering, with the aim to provide perspectives and innovative methods that can be transferred to biotechnology. We next present the importance of modularity that exists in natural biological systems. Finally, we highlight the most recent theoretical and experimental developments in modular cell design for synthetic biology and metabolic engineering applications.

2. Modularity in engineered systems

2.1. Basic concepts

Based on the definition by Miller and Elgard (Miller and Elgard, 1998), a module is “an essential and self-contained functional unit relative to the product of which it is part. The module has, relative to a system definition, standardized interfaces and interactions that allow composition of products by combination”. We find this definition, out of the many available (Salvador, 2007), to be general yet descriptive enough to illustrate the topics in this paper. Based on the above definition, modules must have a standardized interface and exchangeability in order to enable rapid and systematic assembly of components into a system with various types of modular architectures (Ulrich, 1995). A fully modular or sectional architecture has all the components to be modules, for instance, fluid pipes or sectional couches, while an integral architecture lacks any type of modules (Fig. 1A). Chassis-based architectures are also common in modular design, including bus and slot. The bus modular architecture uses the same interface for all modules, e.g., universal serial bus (USB) and peripheral component interconnect (PCI) ports found in computers, while the slot modular architecture has specific interfaces for corresponding modules, e.g., tires in an automobile. The chassis-based architectures enable the use of alternative chassis that can be combined with the same modules to efficiently generate a variety of products or vice versa (Jose and Tollenaere, 2005).

2.2. Driving forces and potential tradeoffs of modular design

The driving forces for system modularization are to achieve increased efficiency and robustness, reduced complexity and cost, and better customization and maintenance options (Bonvoisin et al., 2016; Miller and Elgard, 1998). Modular design has been the core of many innovative technologies across engineering disciplines (Fig. 1B). For instance, modularized plants in chemical engineering allow faster and more cost-effective deployment, making small operations viable and hence providing economic and environmental advantages (Baldea et al., 2017; Kim et al., 2017). Likewise, modular buildings in civil engineering enable more rapid and economical construction (Kamali and Hewage, 2016). In nuclear engineering, the use of small modularized reactors overcomes potential hazards of traditional large-scale operations, allows plant customization to energetic demand, and reduces construction and manufacturing costs (Vujić et al., 2012). The emerging area of highly automated manufacturing also implements modular production systems (Weyer et al., 2015). In addition, modular design principles have been applied with great success in abstract engineering disciplines such as software (Abelson et al., 1996; Gancarz, 2003) and management engineering (Campagnolo and Camuffo, 2010). The decomposition of software elements into modules of defined functionality is essential to manage complexity, ensure robustness, and allow for concurrent development. Similarly, organizations and projects can be structured into modules to accomplish parallel task execution and avoid repetition.

Even though modular design is ubiquitous, it may not be desirable or feasible in every circumstance. The disadvantages and limitations of modularity tend to be field specific. When modularity is part of an innovative approach, such as modular chemical plants, a lack of experience and higher upfront costs are regarded as common drawbacks

(Baldea et al., 2017). Design constraints may also limit the applicability of modularity; for example, a quantitative study (Hölttä-Otto and De Weck, 2007) suggested that the portability requirements of cell phones and laptops make them less modular than their static counterparts. In the case of buildings and chemical plants, the size of components may prevent off-site modular manufacturing if transportation is unfeasible. Thus, when choosing modular design for engineering applications, it is important to ensure that advantages outweigh disadvantages.

2.3. Theoretical frameworks of modular design

Modular product design is complex and field-specific but can be generally formulated using the language of graph theory. The Design Structure Matrix (DSM) (Browning, 2016) is a commonly used technique to model a system as a graph, where nodes represent basic components and links between nodes describe their functional interactions. Component interactions can be represented in a binary manner (i.e., whether a relationship exists or not) or in a detailed complex fashion with multiple dimensions and values to increase model accuracy. For example, with a complex interaction, a metric can be used to quantitatively assign interaction desirability between two components, i.e., a high desirability score if they require each other to work or a negative desirability score otherwise. In some cases, detailed interaction types can also be classified and integrated in the modular design; for example, a cooling system with a radiator and a fan is required to have not only a *spatial* interaction (i.e., both elements need to be in close proximity) but also a *material* interaction (i.e., the fan provides airflow across the radiator) (Fig. 1 C) (Helmer et al., 2010).

DSM can reveal properties of the system through different analyses, including singular value decomposition to capture the modular architecture type (e.g., integral, chassis-based, bus, or slot) (Hölttä-Otto and De Weck, 2007), node centrality ranking to identify key interactions between modules and interfaces (Sosa et al., 2007), and most importantly, clustering to identify modules (Fig. 1 C). While many approaches exist to cluster a DSM, not all can successfully identify modules due to underlying design conflicts in product modularization; for these scenarios, integrated use of a highly descriptive DSM model to account for complex interactions and multi-objective optimization to identify Pareto optimal solutions is needed to accurately design modular systems (Helmer et al., 2010).

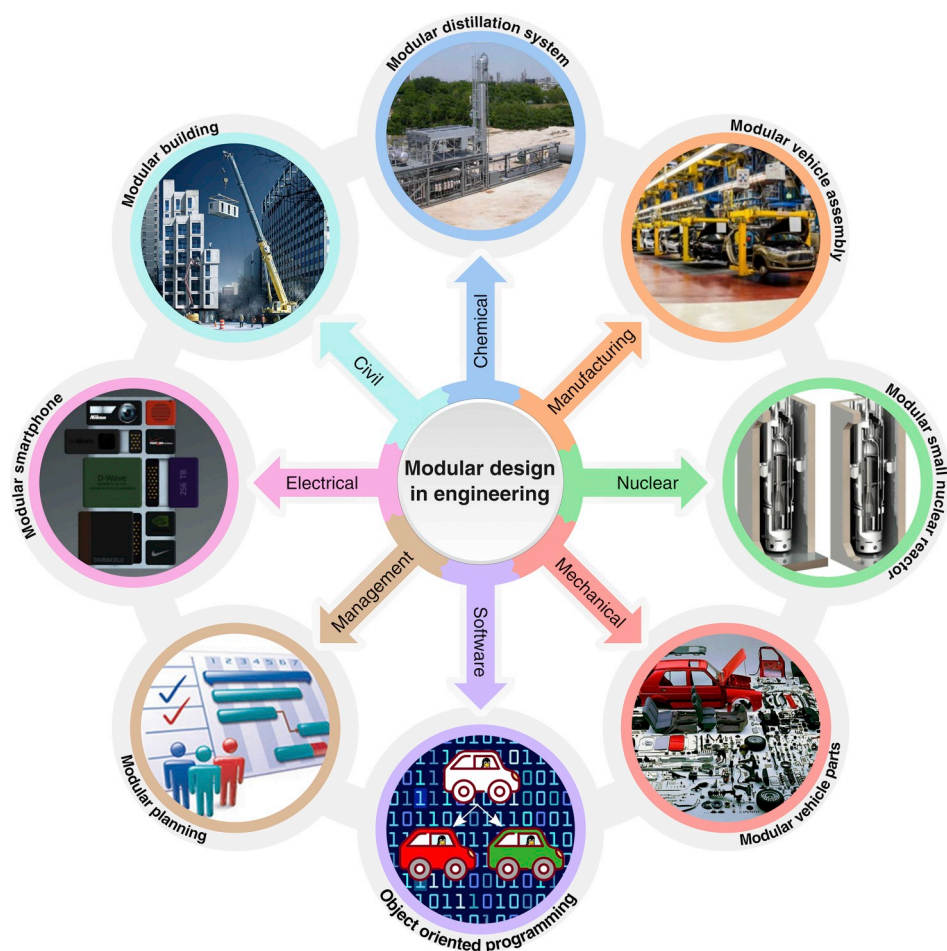
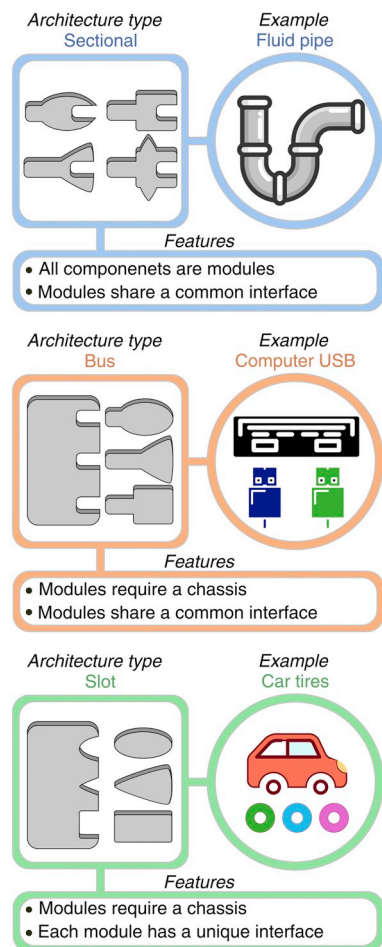
3. Modularity in biological systems

In the high-throughput and quantitative era of biology, modularity has become a fundamental abstraction in understanding and re-designing biological systems that have existed and evolved for billions of years (Hartwell et al., 1999; Wagner et al., 2007). A variety of definitions of biological modules co-exist, arising from the multi-scale and multi-interaction nature of biological systems. From a mathematical perspective, two general approaches are commonly used to define modules: (i) modules as clusters of highly interconnected nodes in a biological interaction network and (ii) modules as programmable circuits that can be described by laws of mass and energy conservation and control theory. These paradigms differ in that network modules provide a holistic and simplified description, while circuit modules provide a reductionistic and detailed description. Despite their differences, both approaches seek to understand the evolutionary origin of modules, their role in fundamental biological properties such as robustness and evolvability, and their biotechnological applications.

3.1. Modularity exists across all scales of biology

Modularity is a ubiquitous organizing principle across all scales of biology. Within a scale, modules often interact in a hierarchical manner (Fig. 2). At the molecular level, DNA transcription activation and rate can be controlled by a variety of modular promoter elements (Dyana,

A. Types of modular architecture **B. Applications of modular design in engineering**



C. General mathematical models for modular product design

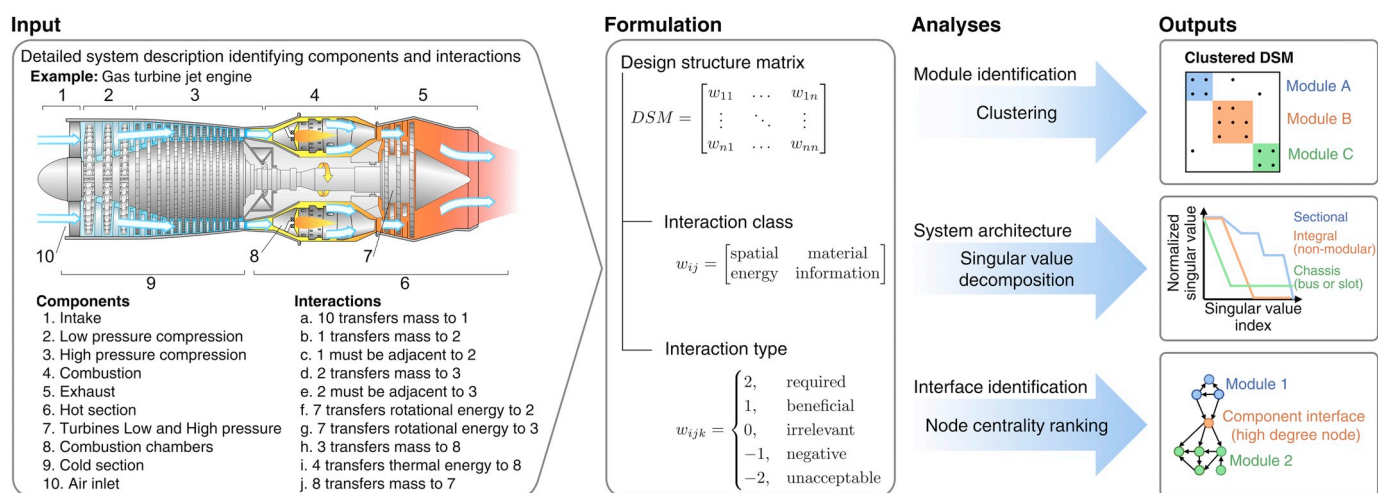


Fig. 1. Modular design in engineering. (A) Common types of modular architectures. (B) Current applications of innovative modular design. (C) General mathematical framework of modular design. The input is illustrated with a gas turbine jet engine, that intakes air through the front section for heating and compression followed by air expansion to generate thrust. The interaction between system components can be formalized in a DSM model and analyzed to identify the most effective modules and interfaces.

1989). Additionally, certain proteins are highly modular, such as enzymes (e.g., polyketide synthase and non-ribosomal peptide synthetase (Hutchinson, 2003)) with modular substrate identification elements that enable biosynthesis of a large space of secondary metabolites (Khosla and Harbury, 2001). At the cellular level, modularity is present

in all biomolecule interaction networks (Mitra et al., 2013), including DNAs (Tettelin et al., 2005), RNAs (Stuart et al., 2003), proteins (Spirin and Mirny, 2003), and metabolites (Ravasz et al., 2002). In all cases, modules are associated with specific cellular functions or pathways, which often interact in a hierarchical manner (Ravasz et al., 2002). At

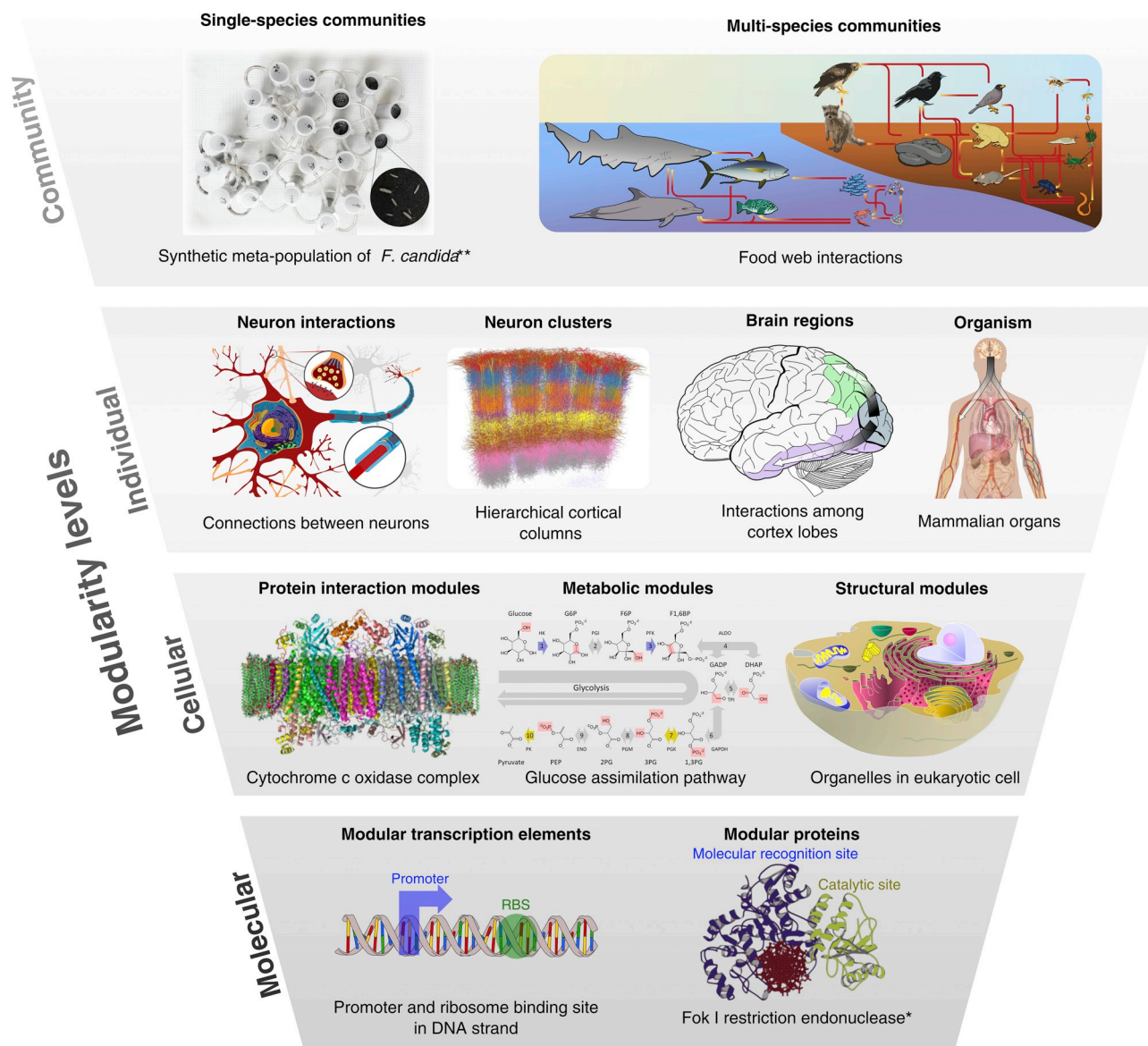


Fig. 2. Hierarchical modularity across all scales of biology. Images marked with * and ** are adapted from Khosla and Harbury (Khosla and Harbury, 2001) and Gilarranz et al. (Gilarranz et al., 2017), respectively.

the multi-cellular level, organs and tissues are also structured modularly. For example, in the human brain, neuron interaction networks contain modules associated with specific cognitive functions (Sporns and Betzel, 2016). These modules are hierarchically organized into submodules to integrate and contextualize specialized functions of the brain (Meunier et al., 2010); for instance, visual perception that requires the functional integration of multiple neuron clusters in the cortical columns (Park and Friston, 2013). At the ecological scale, organism communities can be represented by networks, where nodes are species or subpopulations and links are interactions such as consumption, pollination, or competition (Grilli et al., 2016). The capability of ecological networks to avoid global failure due to small perturbations has been attributed to their modular structure both theoretically (Grilli et al., 2016) and experimentally (Gilarranz et al., 2017).

3.2. Modularity explains functions of components and interactions of biological systems

Prior to the systems biology era, the view of modular biological systems can be traced back as early as the late 19th century with the

initial study of what later became known as glycolysis to investigate how yeast fermentation made wine have a good taste. Throughout the first half of the 20th century, the complete knowledge of many major metabolic pathways of well-defined functionality across organisms, including the EMP pathway, the Entner-Doudoroff pathway, Krebs cycle, pentose phosphate pathway, and so on, was established. These pathways were mapped to qualitatively describe the modular inter-connection of functional elements within a cell, i.e., cellular metabolism that governs cell physiology. Pioneering work in the 1980s, including the comprehensive description of measurable bacterial cellular components (Neidhardt et al., 1990) and constraint-based simulations of microbial metabolism (Fell and Small, 1986), helped establish a foundation for quantitative modular analysis of cell physiology. With the explosion of high-throughput 'omics technologies in the late 90's, complex biological systems with thousands of interacting elements can now be studied holistically and quantitatively at multiple levels from genes to proteins and metabolites within the cell and microbial communities (DeRisi et al., 1997; Lowe et al., 2017; Patti et al., 2012; Wilkins et al., 1996). Graphs (or networks) can represent these systems (Alon, 2003) and their complex interactions, e.g., metabolites-enzymes

(Ma and Zeng, 2003; Ravasz et al., 2002), genes-diseases (Goh et al., 2007), protein-protein (Szkarczyk et al., 2017), or a combination of all known protein and genetic interactions (Chatr-Aryamontri et al., 2017). Module analysis of biological networks can provide two insights: (i) identification of modules that represent transferable and self-contained functions (Mitra et al., 2013), for example, a cis-regulatory element and its associated genes (Ohta et al., 2008), the subunits of a protein complex (Guruharsha et al., 2011; Havugimana et al., 2012), and the genes associated with a disease phenotype (Goh et al., 2007) and (ii) interactions among modules of a system. For example, by analyzing the metabolic networks of 43 organisms, Ravasz et al. (Ravasz et al., 2002) identified a hierarchical modular architecture containing hubs of highly connected metabolites and modules with specific metabolic functions. Remarkably, the identified architecture overlaps with the known primary and secondary metabolic pathways. Likewise, by analyzing metabolic networks of 63 organisms, Ma and Zeng (Ma and Zeng, 2003) could identify a bow-tie architecture of the network containing a core of highly interconnected components linking input and output node clusters. Graph-based analysis of metabolic networks also revealed a small-world architecture (i.e., a small number of reactions between any two metabolites) that was hypothesized to confer cellular metabolism with the capability to quickly adapt to perturbations (Wagner and Fell, 2001).

3.3. Modularity is a foundational tool to control programmable cells

Genetic circuits are modules that define a universal programming language of cells, such as logical gates (Brophy and Voigt, 2014; Nielsen et al., 2016), and can be widely found in natural biological systems. Classical examples of genetic circuits are the lac operon that enables carbon catabolite repression in bacteria and the MAPK/ERK signaling pathway in mammalian cells that controls cell division among other cellular functions. Modularity of natural signaling pathways can be harnessed for novel functions (Lim, 2010; Pryciak, 2009), including the production of valuable metabolites, synthesis of nanomaterials, treatment of disease, and sensors to detect hazardous molecules (Brophy and Voigt, 2014). The laws of mass and energy conservation and principles of control theory that have been well developed in traditional engineering disciplines can be applied to enable modular design of synthetic biological systems. For instance, Del Vecchio et al. (Del Vecchio et al., 2008) developed a mathematical model of retroactivity that captures the impact of a downstream module on the function of an upstream module, and used this model to enhance module insulation. Even though synthetic genetic circuit modules operate correctly and reliably in an isolated environment, it remains challenging to integrate these modules into complex systems (Purnick and Weiss, 2009).

3.4. Modularity enables evolutionary advantage and robustness

Biological robustness is the ability of a system to maintain its function upon genetic and environmental perturbations. Both the graph- and circuit-based descriptions of modules (Whitacre, 2012) can help explain their contributions to system robustness. For example, the bow-tie architecture of biological networks (Friedlander et al., 2015) has been suggested (Kitano, 2004) to enhance the biological robustness of chassis-based modularity, where the essential core processes belong to the conserved chassis and adaptable modules allow for evolution to experiment safely. This view, however, raises the question of the evolutionary origin and function of modules. It has been hypothesized that modules may arise due to natural selection or biased mutational mechanisms. Computational studies on the topic are abundant but compatible with both hypotheses (Wagner et al., 2007). Several models have suggested that when a single fitness goal is present, a modular architecture does not emerge since it does not provide a fitness advantage. However, when multiple fitness goals are pursued, either sequentially (Kashtan et al., 2007) or simultaneously (Clune et al., 2013),

a modular architecture is favored. Additionally, it has been demonstrated that macroscopic (Shoval et al., 2012) and molecular (Schuetz et al., 2012) phenotypes are weighted combinations of optimal phenotypes specialized for single tasks. This view is important in the context of modular cell engineering that can be formulated as a multi-objective optimization problem (Garcia and Trinh, 2019), where each fitness goal corresponds to a module for making a desirable chemical. Thus, modular cell engineering can harness the existing features of biological modularity instead of creating entirely synthetic properties.

4. Modular cell engineering

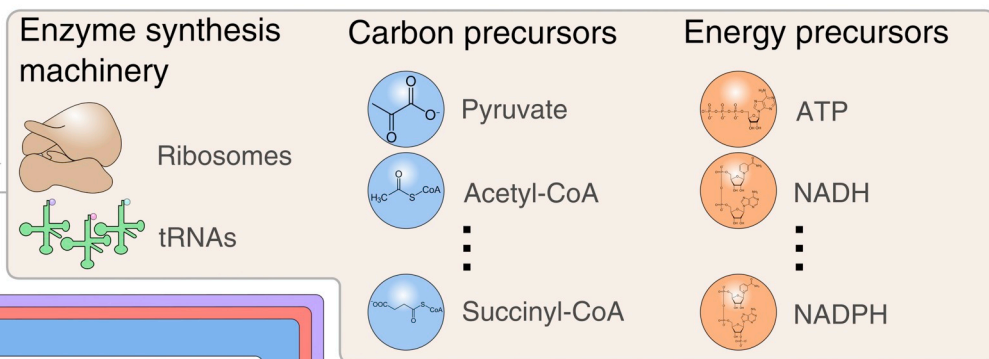
Diverse and complex cellular metabolism encompasses a large space of molecules, providing a potential path towards broader industrialization of biology (Connelly et al., 2015). To realize this potential, rapid development of novel microbial biocatalysts to efficiently synthesize these molecules is critical but faces multiple challenges due to laborious and costly requirement of extensive strain optimization cycles (Nielsen and Keasling, 2016; Trinh and Mendoza, 2016). Effective exploitation of modular design as seen in natural and engineered systems for biocatalyst development can offer innovative strategies to tackle these challenges.

To date, modular cell engineering has been mostly applied at the pathway level, where enzyme module expression is adjusted to increase target metabolite production. This topic has been extensively reviewed (Biggs et al., 2014; Jeschek et al., 2017; Lu et al., 2018; Yadav et al., 2012) and is not elaborated in detail here. At the cellular level, platform strains engineered to eliminate common byproducts and increase availability of important precursors for overproduction of target molecules have been reported with some success by implementing the conventional “push-and-pull” metabolic engineering strategy (Nielsen and Keasling, 2016). More recently, a system-level modular cell design method has been developed to systematically and simultaneously design both the chassis cell and production modules for the synthesis of various target chemicals (Garcia and Trinh, 2019; Trinh et al., 2015). Unlike conventional strain optimization methods that target one product, modular cell design seeks to enable rapid and predictable creation of multiple production strains to achieve superior performance with minimal strain optimization cycle where each synthesizes a different product. Each optimal production strain is obtained by assembling a reusable modular (chassis) cell with an exchangeable production module(s) in a plug-and-play fashion, resembling the advantages of modular design in traditional engineering disciplines. Specifically, a modular cell contains core metabolic phenotypes shared among production modules (Fig. 3A). The chassis interfaces with the modules through enzyme synthesis machinery and precursor metabolites (Fig. 3B). Modules contain auxiliary regulatory and metabolic pathways (Fig. 3C) that enable a desired phenotype for optimal biosynthesis of a target molecule, such as growth-coupled-to-product formation (GCP design) or stationary-phase product synthesis (NGP design) (Fig. 3D). These design principles are formulated to integrate state-of-the-art techniques developed in the fields of synthetic biology and metabolic engineering, including, (i) computational models that identify genetic interventions towards desirable phenotypes, (ii) rapid discovery and optimization of metabolic pathways for target product synthesis (i.e., production modules), and (iii) design of minimal cells and orthogonal pathways to generate a toolkit of parts that can be assembled into various functional systems.

In the following sections, we highlight the recent advancements in modular cell engineering that capture: (1) design principles and computational tools to enable construction of a modular cell and associated production modules, (2) discovery and optimization of metabolic pathways as production modules both theoretically and experimentally, and (3) experimental implementation of modular cell design principles.

A. Chassis

B. Interfaces



C. Modules

D. Production strains

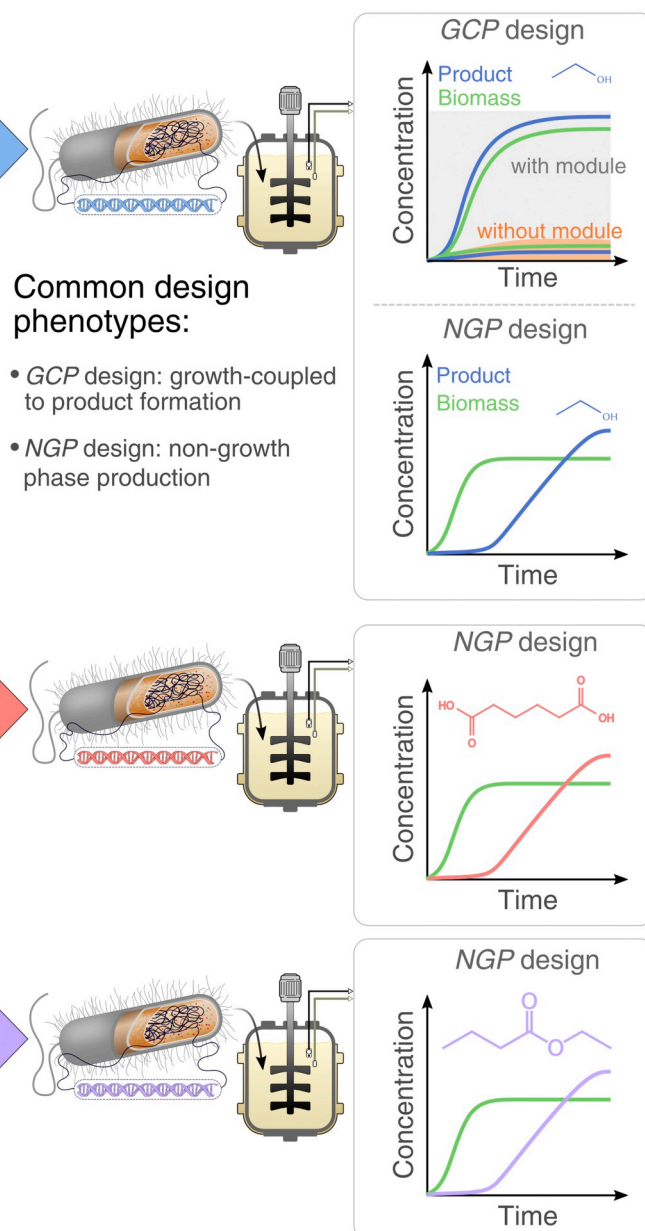
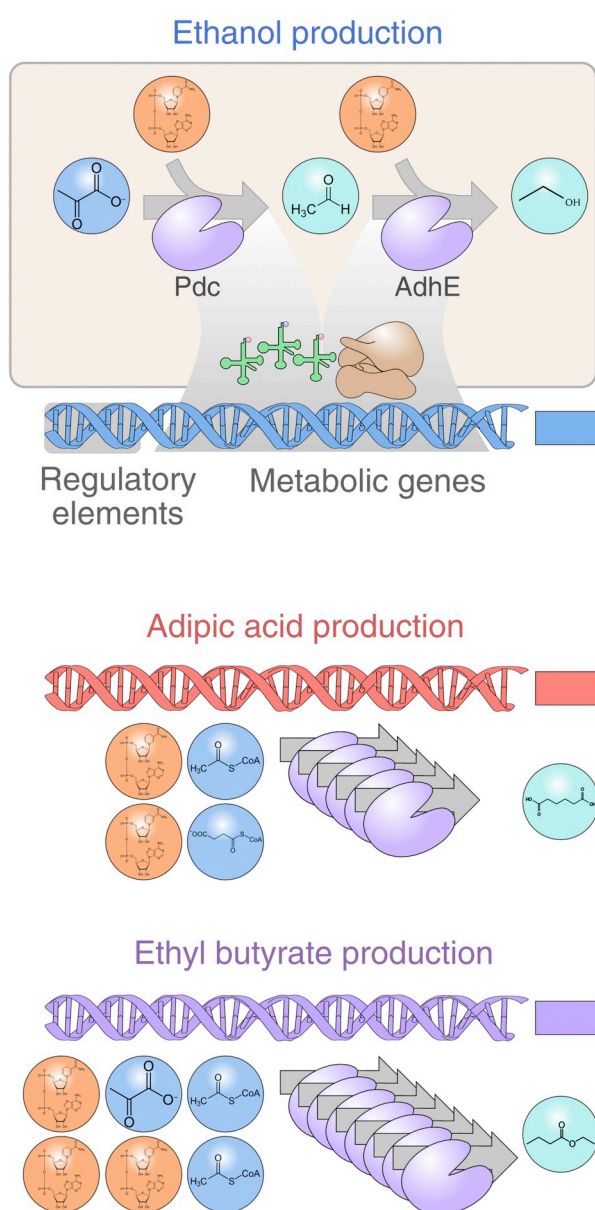


Fig. 3. Generalized concept of modular cell design. (A) Modular (chassis) cell. (B) Interfaces. (C) Production modules. (D) Production strains. A modular cell is designed to provide the necessary precursors for biosynthesis pathway modules that are independently assembled with the modular cell to generate production strains exhibiting desirable phenotypes.

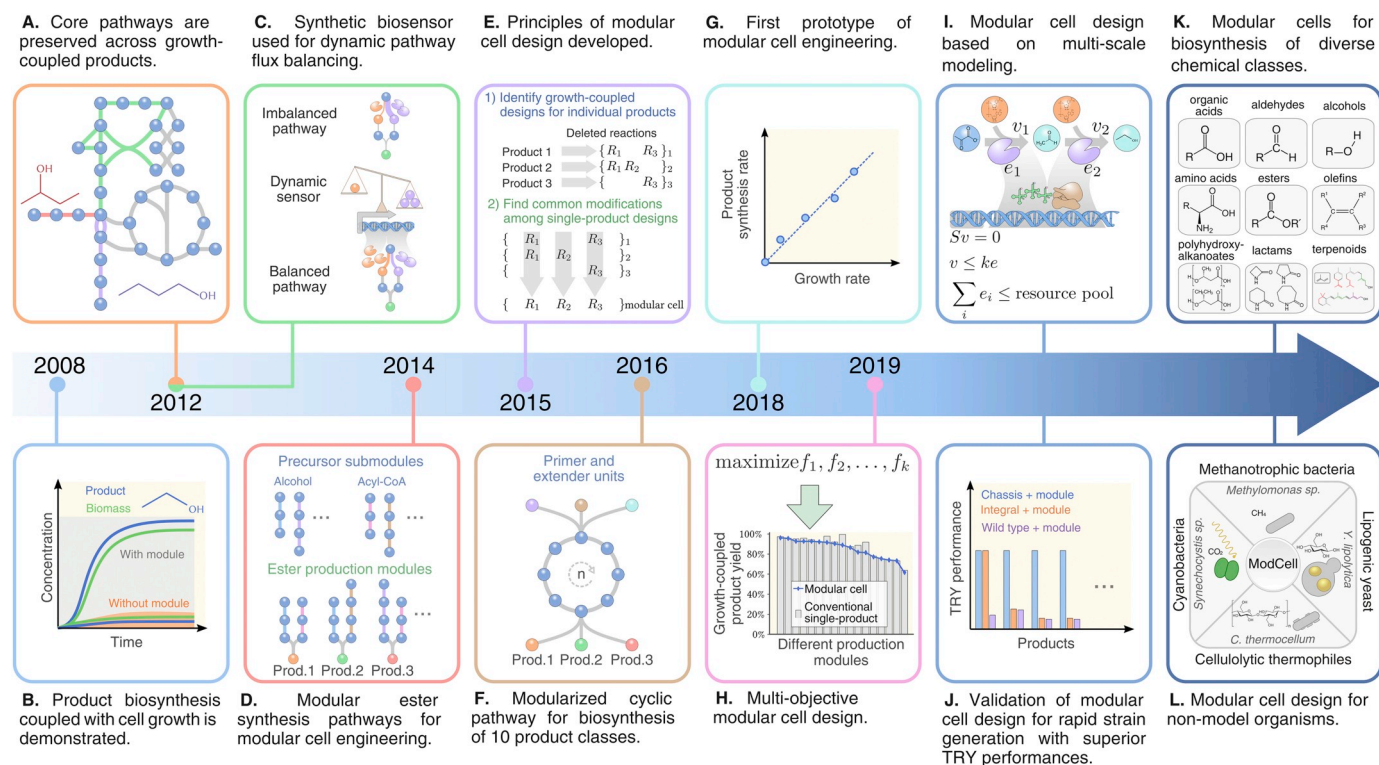


Fig. 4. Key advances and opportunities in the design and implementation of modular cells and exchangeable production modules. (A) The same core metabolic pathways were revealed for n-butanol and isobutanol growth-coupled production based on elementary mode analysis (Trinh, 2012). (B) An *E. coli* cell with minimal metabolic functionality designed to require product (ethanol) synthesis for cell growth was experimentally validated (Trinh et al., 2008b). (C) A dynamic sensor-regulator system was designed and implemented to balance metabolic fluxes for enhanced biosynthesis of fatty acid-derived molecules (Zhang et al., 2012a). (D) A modular ester fermentative pathway platform derived from alcohol and acyl-CoA pathway submodules was designed and experimentally demonstrated in a chassis strain (Layton and Trinh, 2014). (E) First modular cell design method, named MODCELL, was proposed and used to design 3 modular cells for the production of a group of alcohols and derived esters (Trinh et al., 2015). (F) Carbon and energy efficient cyclic pathway was developed to synthesize 10 product classes from a variety of primer and extender precursors (Cheong et al., 2016). (G) Prototypes of modular cell engineering were demonstrated for growth-coupled to product synthesis predicted by MODCELL (Trinh et al., 2015) and pathway optimization by adaptive laboratory evolution (Wilbanks et al., 2017). (H) Multi-objective optimization-based modular cell design method, named ModCell2, was developed and demonstrated to reveal negligible trade-offs between modular and integral designs (Garcia and Trinh, 2019). (I) Next-generation of modular cell design framework is proposed to account for enzyme biosynthesis cost using metabolism and expression (ME) or kinetic models with enzyme constraints. (J) Experimental demonstration of rapid and systematic generation of modular production strains with many different types of production modules to achieve superior performances over the wildtype and single-product (integral) strains in terms of titer, rate, and yield (TRY). (K) Design of a universal modular cell compatible with a large and diverse space of production modules. (L) Future demonstration of modular cell design is proposed for industrially-relevant, non-model organisms with difficult-to-transfer phenotypes, such as efficient assimilation of CO₂, CH₄, or cellulose.

4.1. Recent developments in mathematical formulation of modular cell design

The primer for the modular cell design principles started with the observation (Trinh, 2012) that the optimal design of the n-butanol- and isobutanol-producing *E. coli* cells, based on constraint-based modeling (Palsson, 2015) and conventional strain design methods (Trinh et al., 2009), exhibited the same core metabolism (Fig. 4A-B). To systematically explore this property, the first modular cell design method, called MODCELL, was proposed and used to design an *E. coli* modular cell for alcohol and ester production (Trinh et al., 2015) (Fig. 4E). A new formulation of the modular cell design problem, ModCell2 (Garcia and Trinh, 2019), was proposed based on the framework of multi-objective optimization to capture the tradeoffs that may exist among highly diverse biochemical pathways. ModCell2 enables analysis of genome-scale metabolic networks with many production modules and design of optimal modules without extensive prior knowledge of target pathways. ModCell2 was used to demonstrate that a modular cell can be designed to exhibit (i) desirable production phenotypes for many molecules under various culturing conditions beyond the growth-coupled to product synthesis phase (Klamt and Mahadevan, 2015), such as the stationary-phase phase (Klamt et al., 2018), and (ii) minimal trade-off between compatibility, performance, and robustness (Fig. 4H).

These results highlight that the systems-level modularization of cellular metabolism can accelerate strain engineering without sacrificing performance requirements. While ModCell2 is to our knowledge the only tool that involves simultaneous design of chassis, modules, and interfaces, other recently developed computational tools can be applied to design these elements individually. For example, MinGenome (Wang and Maranas, 2018) is used to design genomes of minimal cells that can serve as chassis whereas ValveFind (Pandit et al., 2017) enables design of orthogonal pathways to build production modules. Topics on model-guided strain design techniques for “integral” strains, i.e., strains optimized for production of only a single product, can be found in recent excellent reviews. (Long et al., 2015; Machado and Herrgård, 2015; Ng et al., 2015).

Future development of modular cell design tools will incorporate enzyme kinetics of cellular metabolism (King et al., 2015; Noor et al., 2016) to account for potential metabolic burdens (Wu et al., 2016) in the design of modular cells and exchangeable production modules (Fig. 4I). Enzyme cost analysis is also particularly useful for identifying robust strategies for adaptive laboratory evolution that prevent an unintended pathway(s) to be optimized instead of a targeted pathway(s) (Dinh et al., 2018). The lack of experimental data on enzyme catalytic efficiencies needed for these modeling approaches can be addressed through random parameter sampling (Dinh et al., 2018), *omics*

integration (Ebrahim et al., 2016; Khodayari and Maranas, 2016), and machine learning (Heckmann et al., 2018). Recently developed models that predict *in vivo* enzyme concentrations from the genetic sequence help bridge the gap between metabolic model predictions and experimental implementation (Farasat et al., 2014; Meng et al., 2013; Salis et al., 2009).

4.2. Recent advances in discovery and optimization of metabolic pathways as production modules

With the arrival of quantitative 'omics era in mid 1990s, we started to gain a deeper understanding of complex biological systems and categorize them into the functional biological parts, i.e., regulatory and functional genes for retrieval through development of biological databases (e.g., Registry of Standard Biological Parts (<https://parts.igem.org>), KEGG (Kanehisa and Goto, 2000), Biocyc (Caspi et al., 2013), Brenda (Scheer et al., 2010), among many others). Synergistically, the interdisciplinary areas of bioinformatics, synthetic biology, metabolic and protein engineering have also emerged, advanced, and now enabled rapid and systematic identification and assembly of these biological parts into production modules to probe a large space of molecules that are only limited by one's imagination. Nowadays, development of production modules can start by using computational tools with increasing scope and accuracy (Kumar et al., 2018; Wang et al., 2017), that help identify metabolic steps, associated enzymes, and genetic parts (i.e., promoters, terminators, ribosome binding sites, regulatory/sensory elements, and so on) using a combination of elementary reaction rules, yield, thermodynamic, and biophysical analyses (Dugar and Stephanopoulos, 2011; Kumar et al., 2018). Next, the genetic parts can be synthesized, assembled, and characterized for accuracy in a rapid manner to build production modules to make desirable molecules in a target host (Annaluru et al., 2014; Bitinaite et al., 2007; Blake et al., 2010; Chen et al., 2013; Colloms et al., 2014; Gibson et al., 2009; Kok et al., 2014; Li and Elledge, 2007; Li and Elledge, 2005; Shao et al., 2009; Trubitsyna et al., 2014; Tsuge et al., 2003; Zhang et al., 2012b). Some recent achievements, highlighting innovations in retrofitting cellular metabolism for novel biocatalysis from sustainable feedstocks, include: (i) redirection of central metabolism to make environmentally friendly, non-natural bioplastics (Rehm, 2010), (ii) redesign of fermentative pathways to produce a large space of designer bioesters used as flavors, fragrances, biofuels, and solvents (Layton and Trinh, 2014; Rodriguez et al., 2014) (Fig. 4D), (iii) repurposing of the beta-oxidation pathway for combinatorial biosynthesis of alcohols, dicarboxylic acids, hydroxyl acids, and lactones as industrial platform chemicals (Cheong et al., 2016) (Fig. 4F), and (iv) refactoring of polyketide and isoprenoid pathways to explore a large space of secondary metabolites as drugs (Ajikumar et al., 2010; Galanie et al., 2015; Martin et al., 2003).

While the design, construction, and characterization of production modules can be streamlined, compatibility between the modules and a target host has always posed a significant challenge mainly due to the intricate flux imbalance resulting in low product titers, rates, and yields (Nielsen and Keasling, 2016). In the case of heterologous enzyme expression, undesirable regulatory interactions at the transcriptional and metabolic levels might hinder the pathway operation in a new host. This problem can be addressed by refactoring pathways into isolated orthogonal modules that are independent of the source organism regulation and do not interfere with heterologous host regulation (Galanie et al., 2015; Sánchez-Pascuala et al., 2017; Tan and Liu, 2017; Temme et al., 2012). In most design scenarios, even after regulatory issues are addressed, metabolic flux imbalance is likely to occur due to the differences in expression and catalytic efficiencies of pathway enzymes. Metabolic flux balancing of engineered pathways is a combinatorial optimization problem that requires modification of gene expression elements (e.g., promoters, ribosome binding sites, etc.) and enzyme engineering to achieve the desired fluxes. Currently, this combinatorial problem is often tackled in two ways: (i) identification of key design

variables and screening and (ii) pathway selection. The search space in variable screening approaches can be reduced by pathway-level modularization, known as Multivariate Modular Metabolic Engineering (Biggs et al., 2014; Jeschek et al., 2017; Yadav et al., 2012). Screening approaches can be effort-intensive and impractical for certain scenarios due to the extensive strain characterizations required to effectively sample the design space. Additionally, these approaches may not be able to solve poor pathway-host interactions where precursor metabolite(s) of the pathway of interest becomes the bottleneck. Alternatively, simultaneous host and pathway optimization can be accomplished by adaptive laboratory evolution, provided that a simple selectable phenotype such as growth is tightly coupled with the desirable product synthesis phenotype. For example, Wilbanks et al. (Wilbanks et al., 2017) recently demonstrated a linear correlation between growth rate and product synthesis rate for computationally designed growth-coupled strains (Trinh et al., 2015) (Fig. 4G). Pathway optimization through growth-coupled design has been applied successfully (Fong et al., 2005), and a recent computational study suggest its applicability to many different products and organisms (von Kamp and Klamt, 2017). Modular cell design is compatible with the two described module optimization tools; particularly, the design of a chassis growth-coupled to its production modules can enable rapid optimization of diverse pathways.

4.3. Recent advances in the experimental implementation of modular cell design

Experimental implementation of modular cell design is still at infancy of validation. Construction of modular cells can be implemented by two methods: (i) top-down approach that aims to remove undesired phenotypes from a naturally existing organism under defined environmental conditions (Trinh et al., 2006; Trinh et al., 2008a) and (ii) bottom-up approach that seeks to create a new organism derived from a synthetic minimal cell. (Hutchison et al., 2016) These two methods draw the same analogy as those in the engineering modular design, classified as "product modularization" and "design with modules" (Bonvoisin et al., 2016). In the "product modularization" approach like the "top-down" approach, the chassis, modules, and interfaces are simultaneously designed either by defining functional carriers and interfaces as part of the design (*ex ante* design) or by clustering existing functions into modules (*ex post* design). Alternatively, in the "design with modules" approach like the "bottom-up" approach, a product is designed out of a collection of predefined compatible parts; for example, the design of a personal computer that is assembled from the existing modules, including a motherboard, a graphic card, a monitor, etc., with standard interfaces. In the modular cell design context, a minimal cell can function as chassis whereas orthogonal refactored pathways that function as modules are independently designed and combined with the chassis to build modular production strains with desirable phenotypes.

To date, all recent computational (Garcia and Trinh, 2019; Trinh et al., 2015) and experimental (Layton and Trinh, 2014; Wilbanks et al., 2017) efforts in implementing modular cell design have been focused on the top-down approach, since it is more feasible and accessible with the current knowledge and available genetic tools. Even though the bottom-up approach is much more challenging (Hutchison et al., 2016), the design principles developed for top-down construction can be applied to create bottom-up minimal modular cells, which would be less prone to failure due to their simpler architecture. Regardless of which approach is chosen, the underlying genotypes essential to target product synthesis appear to be conserved, according to the recent surveys of over two decades of metabolic engineering reports (King et al., 2017; Winkler et al., 2015). This suggests that modular cell design can serve as a unifying platform for rapid strain engineering (Fig. 4J).

An anticipated challenge of modular cell design is that the chassis must provide enough precursor metabolites and enzyme synthesis

machinery to support the target flux through each module. This issue may become increasingly difficult as the biochemical diversity and number of products supported by a single chassis expands (Fig. 4K). Recent development of biosensors coupled with gene expression regulation tools (Liu et al., 2015; Meyer et al., 2018; Moser et al., 2018; Yang et al., 2018; Zhang et al., 2012a) can achieve tunable control over the host metabolism to meet the requirements of each production module (Fig. 4C). In practice, such regulatory elements may be implemented in the host or as a part of specific modules.

5. Conclusions

Inspired by natural and conventional engineering modularity, bioengineers have started applying modular design principles to engineer biological systems at genetic, enzymatic, and cellular levels. Modular cell design aims to integrate all three levels for rapidly creating novel microbial biocatalysts in a plug-and-play fashion with minimal strain optimization cycles. Advancements in genome reading (Goodwin et al., 2016), writing (Casini et al., 2015; Kosuri and Church, 2014), and editing (Barrangou and Doudna, 2016) will provide a unique opportunity to streamline modular cell engineering that effectively harnesses a large space of molecules from cellular metabolism using single organisms or microbial consortia, especially from non-model organisms with industrially-relevant but not-easy-to-transfer traits (Fig. 4L) (Abdel-Mawgoud et al., 2018; Carroll et al., 2018; Kalyuzhnaya et al., 2015; Lynd et al., 2016; Thompson et al., 2016). Particularly, these advancements help streamline the construction of modular cells and exchangeable production modules from both top-down and bottom-up approaches. To further advance modular cell engineering, it is important not only to optimize the interfaces between production modules and modular cell but also to account for robustness and evolvability towards desirable engineered phenotypes. By combining proven engineering methods rooted in the physical and chemical laws with system modeling frameworks (e.g., Pareto optimality theory, graph theory), we can elucidate the modular design principles in biology from natural systems to engineered ones, leading towards fundamental understanding of essential rules of life and broader industrialization of biology.

Acknowledgements

This research was financially supported in part by a NSF CAREER award (NSF#1553250) and The Center for Bioenergy Innovation (CBI), the U.S. Department of Energy (DOE) Bioenergy Research Centers funded by the Office of Biological and Environmental Research in the DOE Office of Science (DE-SC0019412). Due to space limitation, we sincerely apologize to those whose work was not cited.

References

Abdel-Mawgoud, A.M., Markham, K.A., Palmer, C.M., Liu, N., Stephanopoulos, G., Alper, H.S., 2018. Metabolic engineering in the host *Yarrowia lipolytica*. *Metab. Eng.* 50, 192–208.

Abelson, H., Sussman, G.J., Sussman, J., 1996. *Structure and Interpretation of Computer Programs*. MIT Press, Cambridge Massachusetts.

Ajikumar, P.K., Xiao, W.-H., Tyo, K.E., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T.H., Pfeifer, B., Stephanopoulos, G., 2010. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* 330 (6000), 70–74.

Alon, U., 2003. Biological networks: the tinkerer as an engineer. *Science* 301 (5641), 1866–1867.

Annaluru, N., Muller, H., Mitchell, L.A., Ramalingam, S., Stracquadanio, G., Richardson, S.M., Dymond, J.S., Kuang, Z., Scheifele, L.Z., Cooper, E.M., 2014. Total synthesis of a functional designer eukaryotic chromosome. *Science* 344 (6179), 55–58.

Baldea, M., Edgar, T.F., Stanley, B.L., Kiss, A.A., 2017. Modular manufacturing processes: status, challenges and opportunities. *AIChE J.* 63 (10), 4262–4272.

Barrangou, R., Doudna, J.A., 2016. Applications of CRISPR technologies in research and beyond. *Nat. Biotechnol.* 34 (9), 933.

Biggs, B.W., De Paep, B., Santos, C.N.S., De Mey, M., Ajikumar, P.K., 2014. Multivariate modular metabolic engineering for pathway and strain optimization. *Curr. Opin. Biotechnol.* 29, 156–162.

Bitinaite, J., Rubino, M., Varma, K.H., Schildkraut, I., Vaisvila, R., Vaiskunaite, R., 2007. USER™ friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Res.* 35 (6), 1992–2002.

Blake, W.J., Chapman, B.A., Zindal, A., Lee, M.E., Lippow, S.M., Baynes, B.M., 2010. Pairwise selection assembly for sequence-independent construction of long-length DNA. *Nucleic Acids Res.* 38 (8), 2594–2602.

Bonvoisin, J., Halstenberg, F., Buchert, T., Stark, R., 2016. A systematic literature review on modular product design. *J. Eng. Des.* 27 (7), 488–514.

Brophy, J.A., Voigt, C.A., 2014. Principles of genetic circuit design. *Nat. Methods* 11 (5), 508.

Browning, T.R., 2016. Design structure matrix extensions and innovations: a survey and new opportunities. *IEEE Trans. Eng. Manag.* 63 (1), 27–52.

Campagnolo, D., Camuffo, A., 2010. The concept of modularity in management studies: a literature review. *Int. J. Manag. Rev.* 12 (3), 259–283.

Carroll, A.L., Case, A.E., Zhang, A., Atsumi, S., 2018. Metabolic engineering tools in model cyanobacteria. *Metab. Eng.* 50, 47–56.

Casini, A., Storch, M., Baldwin, G.S., Ellis, T., 2015. Bricks and blueprints: methods and standards for DNA assembly. *Nat. Rev. Mol. Cell Biol.* 16 (9), 568.

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C.A., Holland, T.A., Keseler, I.M., Kothari, A., Kubo, A., 2013. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 42 (D1), D459–D471.

Chatr-Aryamontri, A., Oughtred, R., Boucher, L., Rust, J., Chang, C., Kolas, N.K., O'Donnell, L., Oster, S., Theesfeld, C., Sellam, A., 2017. The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* 45 (D1), D369–D379.

Chen, W.-H., Qin, Z.-J., Wang, J., Zhao, G.-P., 2013. The MASTER (methylation-assisted tailorable ends rational) ligation method for seamless DNA assembly. *Nucleic Acids Res.* 41 (8), e93.

Cheong, S., Clomburg, J.M., Gonzalez, R., 2016. Energy- and carbon-efficient synthesis of functionalized small molecules in bacteria using non-decarboxylative Claisen condensation reactions. *Nat. Biotechnol.* 34, 556–561.

Clune, J., Mouret, J.-B., Lipson, H., 2013. The evolutionary origins of modularity. *Proc. R. Soc. B* 280 (1755), 20122863.

Colloms, S.D., Merrick, C.A., Olorunniji, F.J., Stark, W.M., Smith, M.C., Osbourn, A., Keasling, J.D., Rosser, S.J., 2014. Rapid metabolic pathway assembly and modification using serine integrase site-specific recombination. *Nucleic Acids Res.* 42 (4), e23.

Connelly, T., Chang, C., Clarke, L., Ellington, A., Hillson, N., Johnson, R., Keasling, J., Laderman, S., Ossorio, P., Prather, K., 2015. *Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals*. National Academies Press, Washington, DC.

Cramer, S.M., Holstein, M.A., 2011. Downstream bioprocessing: recent advances and future promise. *Curr. Opin. Chem. Eng.* 1 (1), 27–37.

Del Vecchio, D., Ninfa, A.J., Sontag, E.D., 2008. Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.* 4 (1), 161.

DeRisi, J.L., Iyer, V.R., Brown, P.O., 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278 (4), 680–686.

Dinh, H.V., King, Z.A., Palsson, B.O., Feist, A.M., 2018. Identification of growth-coupled production strains considering protein costs and kinetic variability. *Metab. Eng. Commun.* 7, e00080.

Dugar, D., Stephanopoulos, G., 2011. Relative potential of biosynthetic pathways for biofuels and bio-based products. *Nat. Biotechnol.* 29 (12), 1074–1078.

Dynan, W.S., 1989. Modularity in promoters and enhancers. *Cell* 58 (1), 1–4.

Ebrahim, A., Brunk, E., Tan, J., O'Brien, E.J., Kim, D., Szubin, R., Lerman, J.A., Lechner, A., Sastry, A., Bordbar, A., 2016. Multi-omic data integration enables discovery of hidden biological regularities. *Nat. Commun.* 7, 13091.

Farasat, I., Kushwaha, M., Collens, J., Easterbrook, M., Guido, M., Salis, H.M., 2014. Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. *Mol. Syst. Biol.* 10 (6), 731.

Fell, D.A., Small, J.R., 1986. Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochem. J.* 238 (3), 781–786.

Fong, S.S., Burgard, A.P., Herring, C.D., Knight, E.M., Blattner, F.R., Maranas, C.D., Palsson, B.O., 2005. In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol. Bioeng.* 91 (5), 643–648.

Friedlander, T., Mayo, A.E., Thusty, T., Alon, U., 2015. Evolution of bow-tie architectures in biology. *PLoS Comput. Biol.* 11 (3), e1004055.

Galanie, S., Thodey, K., Trenchard, I.J., Interrante, M.F., Smolke, C.D., 2015. Complete biosynthesis of opioids in yeast. *Science* 349 (6252), 1095–1100.

Gancarz, M., 2003. *Linux and the Unix Philosophy*. Digital Press.

Garcia, S., Trinh, C.T., 2019. Multiobjective strain design: a framework for modular cell engineering. *Metab. Eng.* 51, 110–120.

Gibson, D., Young, L., Chuang, R., Venter, J., Hutchison, C., Smith, H., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345.

Gilarranz, L.J., Rayfield, B., Liñán-Cembrano, G., Bascompte, J., Gonzalez, A., 2017. Effects of network modularity on the spread of perturbation impact in experimental metapopulations. *Science* 357 (6347), 199–201.

Goh, K.-I., Cusick, M.E., Valle, D., Childs, B., Vidal, M., Barabási, A.-L., 2007. The human disease network. *Proc. Natl. Acad. Sci.* 104 (21), 8685–8690.

Goodwin, S., McPherson, J.D., McCombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* 17 (6), 333.

Grilli, J., Rogers, T., Allesina, S., 2016. Modularity and stability in ecological communities. *Nat. Commun.* 7, 12031.

Guruharsha, K., Rual, J.-F., Zhai, B., Mintseris, J., Vaidya, P., Vaidya, N., Beekman, C., Wong, C., Rhee, D.Y., Cenaj, O., 2011. A protein complex network of *Drosophila melanogaster*. *Cell* 147 (3), 690–703.

Hartwell, L.H., Hopfield, J.J., Leibler, S., Murray, A.W., 1999. From molecular to modular

- cell biology. *Nature* 402 (6761supp), C47.
- Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S., 2012. A census of human soluble protein complexes. *Cell* 150 (5), 1068–1081.
- Heckmann, D., Lloyd, C.J., Mih, N., Ha, Y., Zielinski, D.C., Haiman, Z.B., Desouki, A.A., Lercher, M.J., Palsson, B.O., 2018. Machine learning applied to enzyme turnover numbers reveals protein structural correlates and improves metabolic models. *Nat. Commun.* 9 (1), 5252.
- Helmer, R., Yassine, A., Meier, C., 2010. Systematic module and interface definition using component design structure matrix. *J. Eng. Des.* 21 (6), 647–675.
- Hölttä-Otto, K., De Weck, O., 2007. Degree of modularity in engineering systems and products with technical and business constraints. *Concurr. Eng.* 15 (2), 113–126.
- Hutchinson, C.R., 2003. Polyketide and non-ribosomal peptide synthases: falling together by coming apart. *Proc. Natl. Acad. Sci.* 100 (6), 3010–3012.
- Hutchison, C.A., Chuang, R.-Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L., 2016. Design and synthesis of a minimal bacterial genome. *Science* 351 (6280), aad6253.
- Jeschek, M., Gerngross, D., Panke, S., 2017. Combinatorial pathway optimization for streamlined metabolic engineering. *Curr. Opin. Biotechnol.* 47, 142–151.
- Jose, A., Tollenaere, M., 2005. Modular and platform methods for product family design: literature analysis. *J. Intell. Manuf.* 16 (3), 371–390.
- Kahl, L.J., Endy, D., 2013. A survey of enabling technologies in synthetic biology. *J. Biol. Eng.* 7 (1), 13.
- Kalyuzhnaya, M.G., Puri, A.W., Lidstrom, M.E., 2015. Metabolic engineering in methanotrophic bacteria. *Metab. Eng.* 29, 142–152.
- Kamali, M., Hewage, K., 2016. Life cycle performance of modular buildings: a critical review. *Renew. Sust. Energ. Rev.* 62, 1171–1183.
- Kanehisa, M., Goto, S., 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30.
- Kashtan, N., Noor, E., Alon, U., 2007. Varying environments can speed up evolution. *Proc. Natl. Acad. Sci.* 104 (34), 13711–13716.
- Khodayari, A., Maranas, C.D., 2016. A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nat. Commun.* 7, Khosla, C., Harbury, P.B., 2001. Modular enzymes. *Nature* 409 (6817), 247.
- Kim, Y.-h., Park, L.K., Yiacoymi, S., Tsouris, C., 2017. Modular chemical process intensification: a review. *Ann. Rev. Chem. Biomol. Eng.* 8, 359–380.
- King, Z.A., Lloyd, C.J., Feist, A.M., Palsson, B.O., 2015. Next-generation genome-scale models for metabolic engineering. *Curr. Opin. Biotechnol.* 35, 23–29.
- King, Z.A., O'Brien, E.J., Feist, A.M., Palsson, B.O., 2017. Literature mining supports a next-generation modeling approach to predict cellular byproduct secretion. *Metab. Eng.* 39, 220–227.
- Kitano, H., 2004. Biological robustness. *Nat. Rev. Genet.* 5 (11), 826.
- Klamt, S., Mahadevan, R., 2015. On the feasibility of growth-coupled product synthesis in microbial strains. *Metab. Eng.* 30, 166–178.
- Klamt, S., Mahadevan, R., Hädicke, O., 2018. When do two-stage processes outperform one-stage processes? *Biotechnol. J.* 13 (2), 1700539.
- Kok, S.D., Stanton, L.H., Slaby, T., Durot, M., Holmes, V.F., Patel, K.G., Platt, D., Shapland, E.B., Serber, Z., Dean, J., 2014. Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synth. Biol.* 3 (2), 97–106.
- Kosuri, S., Church, G.M., 2014. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* 11 (5), 499.
- Kumar, A., Wang, L., Ng, C.Y., Maranas, C.D., 2018. Pathway design using de novo steps through uncharted biochemical spaces. *Nat. Commun.* 9 (1), 184.
- Layton, D.S., Trinh, C.T., 2014. Engineering modular ester fermentative pathways in *Escherichia coli*. *Metab. Eng.* 26, 77–88.
- Lee, S.Y., Kim, H.U., Chae, T.U., Cho, J.S., Kim, J.W., Shin, J.H., Kim, D.I., Ko, Y.-S., Jang, W.D., Jang, Y.-S., 2019. A comprehensive metabolic map for production of bio-based chemicals. *Nat. Catal.* 2 (1), 18.
- Li, M.Z., Elledge, S.J., 2005. MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules. *Nat. Genet.* 37 (3), 311–319.
- Li, M., Elledge, S., 2007. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods* 4 (3), 251–256.
- Lim, W.A., 2010. Designing customized cell signalling circuits. *Nat. Rev. Mol. Cell Biol.* 11 (6), 393.
- Liu, D., Evans, T., Zhang, F., 2015. Applications and advances of metabolite biosensors for metabolic engineering. *Metab. Eng.* 31, 35–43.
- Long, M.R., Ong, W.K., Reed, J.L., 2015. Computational methods in metabolic engineering for strain design. *Curr. Opin. Biotechnol.* 34, 135–141.
- Lowe, R., Shirley, N., Bleakley, M., Dolan, S., Shafee, T., 2017. Transcriptomics technologies. *PLoS Comput. Biol.* 13 (5), e1005457.
- Lu, H., Villada, J.C., Lee, P.K., 2018. Modular metabolic engineering for biobased chemical production. *Trends Biotechnol.* 37 (2), 152–166.
- Lynd, L.R., Guss, A.M., Himmel, M.E., Beri, D., Herring, C., Holwerda, E.K., Murphy, S.J., Olson, D.G., Payne, J., Rydzak, T., 2016. Advances in Consolidated Bioprocessing Using *Clostridium Thermocellum* and *Thermoanaerobacter Saccharolyticum*. *Industrial Biotechnology: Microorganisms*. pp. 365–394.
- Ma, H.-W., Zeng, A.-P., 2003. The connectivity structure, giant strong component and centrality of metabolic networks. *Bioinformatics* 19 (11), 1423–1430.
- Machado, D., Herrgård, M., 2015. Co-evolution of strain design methods based on flux balance and elementary mode analysis. *Metab. Eng. Commun.* 2, 85–92.
- Martin, V.J., Pitera, D.J., Withers, S.T., Newman, J.D., Keasling, J.D., 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* 21 (7), 796–802.
- Meng, H., Wang, J., Xiong, Z., Xu, F., Zhao, G., Wang, Y., 2013. Quantitative design of regulatory elements based on high-precision strength prediction using artificial neural network. *PLoS One* 8 (4), e60288.
- Meunier, D., Lambiotte, R., Bullmore, E.T., 2010. Modular and hierarchically modular organization of brain networks. *Front. Neurosci.* 4, 200.
- Meyer, A.J., Segall-Shapiro, T.H., Glassey, E., Zhang, J., Voigt, C.A., 2018. *Escherichia coli* “marionette” strains with 12 highly optimized small-molecule sensors. *Nat. Chem. Biol.* 1.
- Miller, T.D., Elgard, P., 1998. Defining modules, modularity and modularization. In: *Proceedings of the 13th IPS Research Seminar, Fuglsoe. Aalborg Universiy, Fuglsoe.*
- Mitra, K., Carvunis, A.-R., Ramesh, S.K., Ideker, T., 2013. Integrative approaches for finding modular structure in biological networks. *Nat. Rev. Genet.* 14 (10), 719.
- Moser, F., Borujeni, A.E., Ghodasara, A.N., Cameron, E., Park, Y., Voigt, C.A., 2018. Dynamic control of endogenous metabolism with combinatorial logic circuits. *Mol. Syst. Biol.* 14 (11), e8605.
- Neidhardt, F.C., Ingraham, J.L., Schaechter, M., 1990. *Physiology of the Bacterial Cell: A Molecular Approach*. Sinauer Associates Sunderland, MA.
- Ng, C.Y., Khodayari, A., Chowdhury, A., Maranas, C.D., 2015. Advances in de novo strain design using integrated systems and synthetic biology tools. *Curr. Opin. Chem. Biol.* 28, 105–114.
- Nielsen, J., Keasling, Jay D., 2016. Engineering cellular metabolism. *Cell* 164, 1185–1197.
- Nielsen, A.A., Der, B.S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E.A., Ross, D., Densmore, D., Voigt, C.A., 2016. Genetic circuit design automation. *Science* 352 (6281), aac7341.
- Noor, E., Flamholz, A., Bar-Even, A., Davidi, D., Milo, R., Liebermeister, W., 2016. The protein cost of metabolic fluxes: prediction from enzymatic rate laws and cost minimization. *PLoS Comput. Biol.* 12 (11), e1005167.
- Ohta, H., Kinoshita, K., Saeki, M., Hayashi, S., Obayashi, T., 2008. ATTED-II provides coexpressed gene networks for Arabidopsis. *Nucleic Acids Res.* 37 (suppl_1), D987–D991.
- Olson, D.G., McBride, J.E., Shaw, A.J., Lynd, L.R., 2012. Recent progress in consolidated bioprocessing. *Curr. Opin. Biotechnol.* 23 (3), 396–405.
- Palsson, B.O., 2015. *Systems Biology: Constraint-Based Reconstruction and Analysis*, 2nd ed. Cambridge University Press, Cambridge, United Kingdom.
- Pandit, A.V., Srinivasan, S., Mahadevan, R., 2017. Redesigning metabolism based on orthogonality principles. *Nat. Commun.* 8, 15188.
- Park, H.-J., Friston, K., 2013. Structural and functional brain networks: from connections to cognition. *Science* 342 (6158), 1238411.
- Patti, G.J., Yanes, O., Siuzdak, G., 2012. Innovation: metabolomics: the omics trilogy. *Nat. Rev. Mol. Cell Biol.* 13 (4), 263.
- Pryciak, P.M., 2009. Designing new cellular signaling pathways. *Chem. Biol.* 16 (3), 249–254.
- Purnick, P.E.M., Weiss, R., 2009. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* 10, 410–422.
- Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N., Barabási, A.-L., 2002. Hierarchical organization of modularity in metabolic networks. *Science* 297, 1551–1555.
- Rehm, B.H., 2010. Bacterial polymers: biosynthesis, modifications and applications. *Nat. Rev. Microbiol.* 8 (8), 578.
- Rodriguez, G.M., Tashiro, Y., Atsumi, S., 2014. Expanding ester biosynthesis in *Escherichia coli*. *Nat. Chem. Biol.* 10, 259–265.
- Salis, H.M., Mirsky, E.A., Voigt, C.A., 2009. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* 27 (10), 946.
- Salvador, F., 2007. Toward a product system modularity construct: literature review and reconceptualization. *IEEE Trans. Eng. Manag.* 54 (2), 219–240.
- Sánchez-Pascuala, A., de Lorenzo, V., Nikel, P., 2017. Refactoring the Embden–Meyerhof–Parnas pathway as a whole of portable GlucoBricks for implantation of glycolytic modules in gram-negative bacteria. *ACS Synth. Biol.* 6 (5), 793–805.
- Scheer, M., Grote, A., Chang, A., Schomburg, I., Munaretto, C., Rother, M., Söhngen, C., Stelzer, M., Thiele, J., Schomburg, D., 2010. BRENDA, the enzyme information system in 2011. *Nucleic Acids Res.* 39 (suppl_1), D670–D676.
- Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M., Sauer, U., 2012. Multidimensional optimality of microbial metabolism. *Science* 336 (6081), 601–604.
- Shao, Z., Zhao, H., Zhao, H., 2009. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* 37 (2), e16.
- Shoval, O., Sheftel, H., Shinar, G., Hart, Y., Ramote, O., Mayo, A., Dekel, E., Kavanagh, K., Alon, U., 2012. Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science* 336 (6085), 1157–1160.
- Sosa, M.E., Eppinger, S.D., Rowles, C.M., 2007. A network approach to define modularity of components in complex products. *J. Mech. Des.* 129 (11), 1118–1129.
- Spirin, V., Mirny, L.A., 2003. Protein complexes and functional modules in molecular networks. *Proc. Natl. Acad. Sci.* 100 (21), 12123–12128.
- Sporns, O., Betzel, R.F., 2016. Modular brain networks. *Annu. Rev. Psychol.* 67, 613–640.
- Stuart, J.M., Segal, E., Koller, D., Kim, S.K., 2003. A gene-coexpression network for global discovery of conserved genetic modules. *Science* 302 (5643), 249–255.
- Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., 2017. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* 45 (D1), D362–D368.
- Tan, G.-Y., Liu, T., 2017. Rational synthetic pathway refactoring of natural products biosynthesis in actinobacteria. *Metab. Eng.* 39, 228–236.
- Temme, K., Zhao, D.H., Voigt, C.A., 2012. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc. Natl. Acad. Sci. U. S. A.* 109 (18), 7085–7090.
- Tettelin, H., Massignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome”. *Proc. Natl. Acad. Sci.* 102 (39), 13950–13955.
- Thompson, R.A., Dahal, S., Garcia, S., Nookaew, I., Trinh, C.T., 2016. Exploring complex

- cellular phenotypes and model-guided strain design with a novel genome-scale metabolic model of *Clostridium thermocellum* DSM 1313 implementing an adjustable cellulosome. *Biotechnol. Biofuels* 9 (1), 194.
- Trinh, C.T., 2012. Elucidating and reprogramming *Escherichia coli* metabolisms for obligate anaerobic n-butanol and isobutanol production. *Appl. Microbiol. Biotechnol.* 95 (4), 1083–1094.
- Trinh, C.T., Mendoza, B., 2016. Modular cell design for rapid, efficient strain engineering toward industrialization of biology. *Curr. Opin. Chem. Eng.* 14, 18–25.
- Trinh, C.T., Carlson, R., Wlaschin, A., Sreenc, F., 2006. Design, construction and performance of the most efficient biomass producing *E. coli* bacterium. *Metab. Eng.* 8 (6), 628–638.
- Trinh, C.T., Unrean, P., Sreenc, F., 2008a. Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. *Appl. Environ. Microbiol.* 74 (12), 3634–3643.
- Trinh, C.T., Unrean, P., Sreenc, F., 2008b. Minimal *Escherichia coli* cell for the Most efficient production of ethanol from hexoses and pentoses. *Appl. Environ. Microbiol.* 74, 3634–3643.
- Trinh, C.T., Wlaschin, A., Sreenc, F., 2009. Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl. Microbiol. Biotechnol.* 81, 813–826.
- Trinh, C.T., Liu, Y., Conner, D.J., 2015. Rational design of efficient modular cells. *Metab. Eng.* 32, 220–231.
- Trubitsyna, M., Michlewski, G., Cai, Y., Elfick, A., French, C.E., 2014. PaperClip: rapid multi-part DNA assembly from existing libraries. *Nucleic Acids Res.* gku829.
- Tsuge, K., Matsui, K., Itaya, M., 2003. One step assembly of multiple DNA fragments with a designed order and orientation in *Bacillus subtilis* plasmid. *Nucleic Acids Res.* 31 (21), e133.
- Ulrich, K., 1995. The role of product architecture in the manufacturing firm. *Res. Policy* 24 (3), 419–440.
- von Kamp, A., Klamt, S., 2017. Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms. *Nat. Commun.* 8, 15956.
- Vujić, J., Bergmann, R.M., Škoda, R., Miletić, M., 2012. Small modular reactors: simpler, safer, cheaper? *Energy* 45 (1), 288–295.
- Wagner, A., Fell, D.A., 2001. The small world inside large metabolic networks. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 268 (1478), 1803–1810.
- Wagner, G.P., Pavlicev, M., Cheverud, J.M., 2007. The road to modularity. *Nat. Rev. Genet.* 8 (12), 921–931.
- Wang, L., Maranas, C.D., 2018. MinGenome: an in silico top-down approach for the synthesis of minimized genomes. *ACS Synth. Biol.* 7 (2), 462–473.
- Wang, L., Dash, S., Ng, C.Y., Maranas, C.D., 2017. A review of computational tools for design and reconstruction of metabolic pathways. *Synth. Syst. Biotechnol.* 2 (4), 243–252.
- Weyer, S., Schmitt, M., Ohmer, M., Gorecky, D., 2015. Towards industry 4.0-standardization as the crucial challenge for highly modular, multi-vendor production systems. *Ifac-Papersonline* 48 (3), 579–584.
- Whitacre, J.M., 2012. Biological robustness: paradigms, mechanisms, and systems principles. *Front. Genet.* 3, 67.
- Wilbanks, B., Layton, D., Garcia, S., Trinh, C., 2017. A prototype for modular cell engineering. *ACS Synth. Biol.* 7 (1), 187–199.
- Wilkins, M.R., Sanchez, J.-C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D.F., Williams, K.L., 1996. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev.* 13 (1), 19–50.
- Winkler, J.D., Halweg-Edwards, A.L., Gill, R.T., 2015. The LASER database: formalizing design rules for metabolic engineering. *Metab. Eng. Commun.* 2, 30–38.
- Wu, G., Yan, Q., Jones, J.A., Tang, Y.J., Fong, S.S., Koffas, M.A., 2016. Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. *Trends Biotechnol.* 34 (8), 652–664.
- Yadav, V.G., De Mey, M., Giaw Lim, C., Kumaran Ajikumar, P., Stephanopoulos, G., 2012. The future of metabolic engineering and synthetic biology: towards a systematic practice. *Metab. Eng.* 14, 233–241.
- Yang, Y., Lin, Y., Wang, J., Wu, Y., Zhang, R., Cheng, M., Shen, X., Wang, J., Chen, Z., Li, C., 2018. Sensor-regulator and RNAi based bifunctional dynamic control network for engineered microbial synthesis. *Nat. Commun.* 9 (1), 3043.
- Zhang, F., Carothers, J.M., Keasling, J.D., 2012a. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* 30 (4), 354.
- Zhang, Y., Werling, U., Edelmann, W., 2012b. SLICE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res.* 40 (8), e55.