Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Engineering

Modular cell design for rapid, efficient strain engineering toward industrialization of biology

Cong T Trinh[1,2,3] and Brian Mendoza[1,3]

Transforming biology into an engineering practice has great potential to shape the industrialization of biology that will drive rapid development of novel microbial manufacturing platforms. These platforms will be capable of producing a vast number of sustainable industrial chemicals at scale from alternative renewable feedstocks or wastes (e.g., biomass residues, biogas methane, syngas, CO_2) without harming the environment. The challenge is to develop microbial platforms to produce targeted chemicals with high efficiency in a rapid, predictable, and reproducible fashion. This paper highlights recent progress in rational design of heterologous pathways for combinatorial biosynthesis of a large space of chemicals and modular cell design for rapid strain engineering.

Addresses
[1] Department of Chemical and Biomolecular Engineering, The University of Tennessee, Knoxville, United States
[2] Bredesen Center for Interdisciplinary Research and Graduate Education, The University of Tennessee, Knoxville, United States
[3] Bioenergy Science Center (BESC), Oak Ridge National Laboratory, Oak Ridge, United States

Corresponding author: Trinh, Cong T (ctrinh@utk.edu)

Current Opinion in Chemical Engineering 2016, **14**:18–25

This review comes from a themed issue on **Biotechnology and bioprocess engineering**

Edited by **Terry Papoutsakis** and **Nigel Titchener-Hooker**

http://dx.doi.org/10.1016/j.coche.2016.07.005

Introduction

Underlying the rapid advancement of technological innovation during the 20th century was the harnessing of petroleum for its diverse library of potential products. Recent research and socioeconomic developments have revealed the dangers of relying on fossil fuels as a singular source of specialty fuels and chemicals [1]. Chemicals derived from engineered microorganisms have been lauded as a promising sustainable alternative to the thousands of chemicals derived from petroleum [2]. While some success stories of biosynthesized molecules achieved commercialization in recent years [3], the complexity of constructing these microbial cell factories is surpassed by the vastness of the biochemical space open to exploration [4]. For biosynthesized products to compete with current solutions economically and in a variety of applications, research efforts must endeavor to achieve an understanding of the combinatorial space available, and devise an effective means for rapid strain engineering to produce these products based on nature's inherent modularity and synergies.

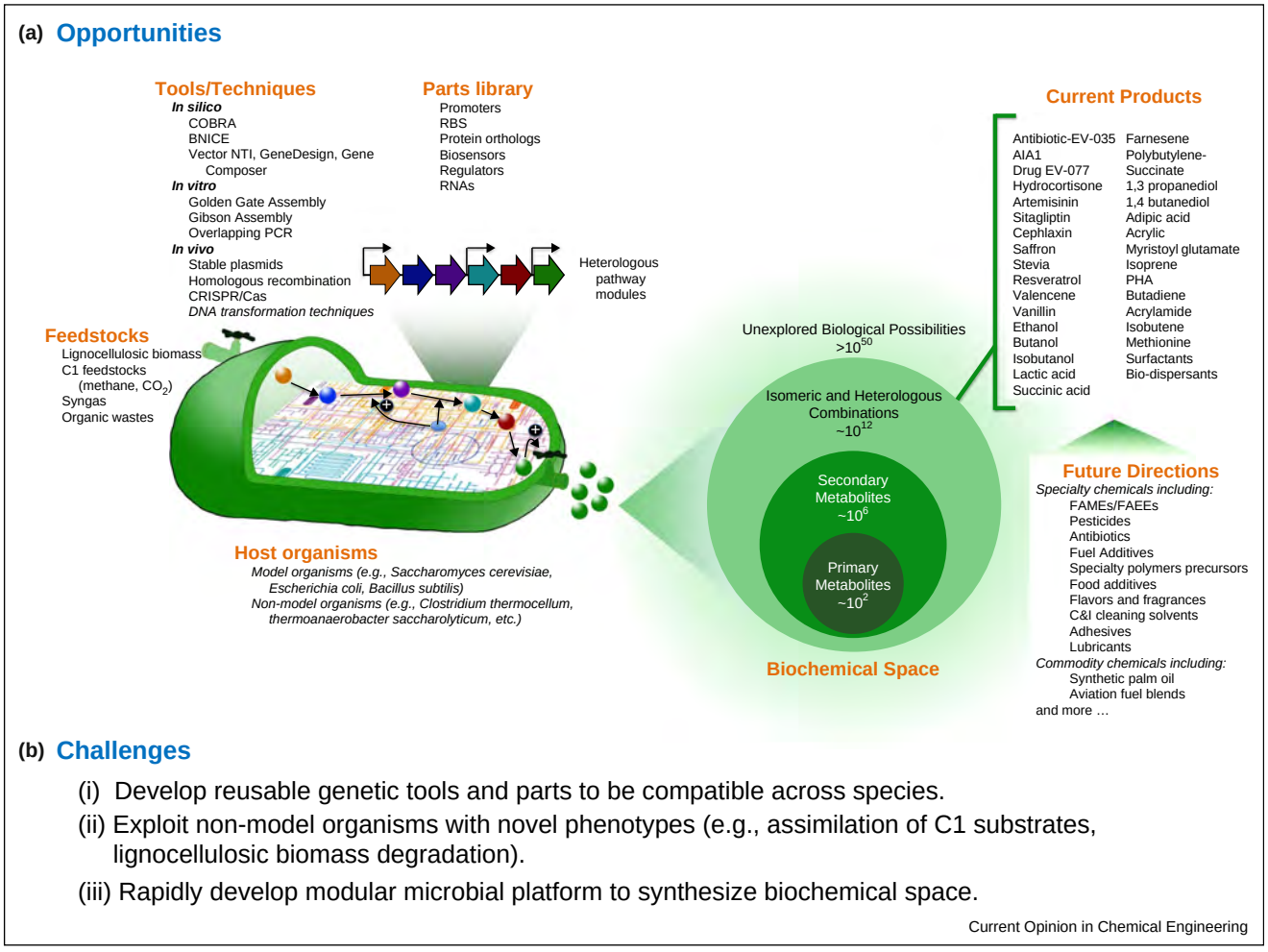
In this paper, we explore the progress made in rational design of heterologous pathways and highlight the importance of engineering dynamic control of heterologous pathways coupled with the host cell metabolism to achieve improved pathway efficiency. We envision the development of modular cell design principles will enable rapid strain engineering for combinatorial biosynthesis of a large, sustainable chemical space in a plug-and-play fashion requiring minimum strain optimization cycles.

Probing the combinatorial space of biobased chemicals

Cellular metabolisms are diverse and complex, generating thousands of unique chemicals. Advancements in comparative genomics, systems and synthetic biology, and metabolic engineering have enabled researchers to access a multitude of biological parts to assemble heterologous pathways and start probing a combinatorial space of biobased chemicals (Figure 1). Characterized biological parts have been compiled into databases such as the Synthetic Biology Parts Registry (http://parts.igem.org/Main_Page), KEGG [5], and Biocyc [6] among others that continue to expand as new genomes are discovered. While these databases continue to add to the breadth of biological parts information, the minimal depth of quantitative knowledge (e.g., transcription rates, enzyme kinetics, etc.) is limited, and will be the next challenge to address as the databases evolve to encompass a greater understanding of parts and their interactions with cellular systems.

Recent achievements in harnessing biological parts for chemical biosynthesis include: (i) rewiring central metabolism for making non-natural bioplastics from sustainable feedstocks [7], (ii) redesigning fermentative pathways for combinatorial biosynthesis of unique esters with tunable carbon backbones [8,9•], (iii) harnessing synthetic pathways of reverse beta-oxidation and non-decarboxylative Claisen condensation coupled with subsequent beta-reduction reactions for combinatorial biosynthesis of alcohols, dicarboxylic acids, hydroxyl acids, and lactones [10,11••], (iv) manipulating polyketide and isoprenoid

Figure 1



The power **(a)** and the daunting challenge **(b)** for microbial synthesis of the potential combinatorial space of specialty fuels and chemicals.

biosynthesis pathways to produce alcohol fuels [12], (v) engineering fatty acid biosynthesis for producing hydrocarbons [13–16], and (vi) rerouting amino acid biosynthesis pathways for making alcohols, drug precursors, and industrial chemicals [17,18]. These chemicals have broad applications related to health, energy, and the environment, but are sometimes difficult to synthesize by a conventional chemical method.

Rational design of heterologous pathways

Engineering heterologous pathways in a recombinant host can become very challenging especially as the pathway complexity increases as seen in the production of opioids in *Saccharomyces cerevisiae* [19••]. For assembling multiple parts into a heterologous pathway, it is critical to balance and optimize fluxes through not only the heterologous pathway to produce a desirable chemical but also the host's native pathways to maintain good cell viability [20–22]. The challenge is how one can identify an efficient heterologous pathway and choose a proper combination and assembly of parts for the pathway to create optimal phenotypes (e.g., efficient production of desirable chemicals at high yields, titers, and productivities) without going through extensive screening. For instance, designing an optimal multi-gene pathway in a bacterial host can generate a vast space of solutions that depend on finding the best pairing of appropriate promoters, ribosome binding sites, terminators, gene orders, tunable intergenic regions, and orthologous genes. In recent years, a collection of computation-based techniques has been developed to assist heterologous pathway design and can be classified into three groups: pathway prediction, yield analysis, and parts identification.

Pathway prediction is an essential tool to identify all thermodynamically feasible routes and associated enzymes to produce desirable chemicals from the existing databases. Hatzimanikatis and coworkers first developed

the BNICE (Biochemical Network Integrated Computational Explorer) framework to identify novel heterologous pathways [23]. Currently, there are a variety of available tools for pathway prediction [24•] with improved search algorithms to address the computational challenge in identifying heterologous pathways from large, putative biochemical reaction databases [25]. For instance, by employing both computational metabolic pathway search and analysis from KEGG, Zhang *et al.* interrogated the metabolic potential of *Escherichia coli* as a microbial platform capable of producing 1777 non-native chemicals, 279 of which have commercial use. Interestingly, more than 50% of these commercial products require a minimum of three heterologous reaction steps [26•].

Yield analysis. Once a desirable heterologous pathway is formulated, the next important step is to evaluate the pathway efficiency. Yield analysis can be employed to evaluate a pathway's potential [27] by considering thermodynamics, electron and carbon constraints for both the engineered heterologous pathway, and native metabolism of the host. Constraint-based metabolic network modeling such as flux balance or elementary mode analysis offers powerful tools for performing yield analysis with a large collection of useful software [28]. Recently, there has been a significant interest in engineering carbon-conserving pathways in the native host as they not only can potentially improve theoretical pathway yields but also reduce the CO_2 carbon footprint harming the environment. For instance, a native carbon-conserving pathway is the succinic acid-producing pathway that could yield 1 Cmol succinate/Cmol glucose equivalent from sugar fermentation. In contrast, the synthesis of alcohol fuels (e.g., ethanol and butanol) can only yield up to 0.67 Cmol product/Cmol glucose equivalent where 33% of carbon is lost to CO_2 as waste. Recent discovery of the non-oxidative glycolysis (NOG) pathway opens new opportunities to potentially engineer heterologous pathways to produce chemicals with product yields greater than the theoretical limits [29,30]. Other unique pathways like NOG likely exist, waiting to be discovered to build alternative novel carbon-conserving pathways.

Parts identification. The next critical step in heterologous pathway design is to identify parts that can be assembled to achieve optimal phenotypes. Since these pathway components are often synthetic and heterologous, it becomes very challenging to select compatible parts that display desirable phenotypes without going through iterative optimization. Currently, parts selection is mostly trial-and-error and relies on high-throughput screening to generate feedback for the design-build-test-learn cycles of pathway engineering. Toward addressing this challenge, Farasat *et al.* has recently developed the SEAMAP framework to design heterologous pathways with balanced and enhanced metabolic fluxes by manipulating ribosome binding sites to fine-tune translation rates and narrow the experimental strain engineering space [31•].

Once the heterologous pathway is designed with appropriate parts, a variety of tools and techniques exist for streamlined synthesis and assembly. One class of parts assembly techniques is based on homologous recombination including a large collection of in vitro [32–39] and in vivo [40–42] assembly methods. In parallel, the other popular class of parts assembly techniques is based on non-homologous recombination such as Golden Gate assembly [43] among many other options [44–49]. Combined, these techniques enable the creation of combinatorial libraries and large DNA fragments (>1 Mb) for constructing minimal cells [50••]. Computational tools (Vector NTI [51], j5 [52], Gene Designer [53], GeneDesign [54], and Gene Composer [55]) make the assembly process very seamless nowadays.

Pathway evaluation for design-build-test-learn cycles

Ideally, the engineered heterologous pathways work consistently with design to achieve desired products at high yields, titers, and productivities. However, the performance is often suboptimal in practice due to metabolic flux imbalances caused by incompatibility of the heterologous pathways and the host (e.g., inefficient protein expression, enzyme stability, metabolic burden, and redox imbalances). Metabolic flux analysis coupled with OMICS data is a state-of-the-art technique to identify metabolic bottlenecks [56–58]. Once the bottlenecks are identified, the control of these pathways together with their interaction with the regulatory machinery of the host is essential to obtain viable yields through the design-build-test-learn cycles. Next, we highlight significant progress in balancing and optimizing metabolic fluxes through static and dynamic controls at both pathway and cell population levels.

Balancing Act: Steady-state control. To overcome reaction bottlenecks, various techniques can be employed to modulate reaction fluxes via environmental (temperature, pH, substrates) and genetic (transcription, translation) manipulation to adjust metabolite concentrations as well as enzyme stability, concentrations, activities, and localization [20]. Similar to modularity implemented in the realm of industrial design, the complexity of pathways can be modularized for efficient flux control leading to modular production of related chemicals [8,9•,59–62]. Metabolic control theory has been developed, and should be of great value for guiding modular pathway engineering in the future [63,64].

Juggling Act: Dynamic pathway control. Fluctuations in the cellular environment, whether planned or inadvertent, can create metabolic imbalances in biological systems optimized for very particular conditions. Natural biological

systems have evolved with efficient sensor and regulator systems (e.g., the well-known *lac* operon of *E. coli*) so that they can effectively respond to environmental cues to maximize their fitness. Since heterologous pathways are often incompatible to the host, dynamic controllers engineered in these pathways would be highly advantageous to maximize product formation while maintaining healthy cells. Dynamic pathway control is the forefront of tool development, going beyond steady-state balance. Next, we highlight some recent advancement in harnessing biosensors and genetic circuit tools for dynamic pathway control.

To engineer dynamic control of any heterologous pathway, the critical element is to have controllable biosensors to construct the sensor-regulator system. Biosensors can be classified into many types including metabolic response transcription factors, two-component systems, regulatory RNAs, and protein allostery [65•]. At pathway levels, these sensor-regulator systems are hard-wired into heterologous pathways for sensing intracellular metabolites and/or environmental cues (e.g., temperature, pH, light) to control metabolic fluxes via feedback and/or feedforward mechanisms [66,67]. At cell levels, they can be used to decouple growth and production phases by implementing transcriptional toggle switches [68,69]. A CRISPRi/a (interference and activation) system with RNA scaffolding, or a genetic switchboard using riboswitches, prove to be powerful tools to dynamically toggle between heterologous pathways [70•,71]. Population quality control (PopQC) is also a useful genetic circuit to enforce the optimal performance of the heterologous pathway [72•]. Because of growth competition fitness, heterogeneity in a cell population carrying heterologous pathways often causes low product production. To address this problem, Xiao *et al.* developed the PopQC strategy by dynamically coupling heterologous pathway flux with antibiotic selection to select for the most hyperproducing strain in the population during the chemical production phase. The team demonstrated a 3-fold improvement in production of fatty acids and tyrosine. Since specificity toward target metabolites is critical for engineering controllable biosensors of heterologous pathways, Taylor *et al.* took a de novo protein design approach to reengineer an allosteric transcription factor (aTF) LacI of *E. coli* to sense alternative ligands [73]. This approach is powerful to design new properties of abundant, native aTFs for dynamic pathway control.

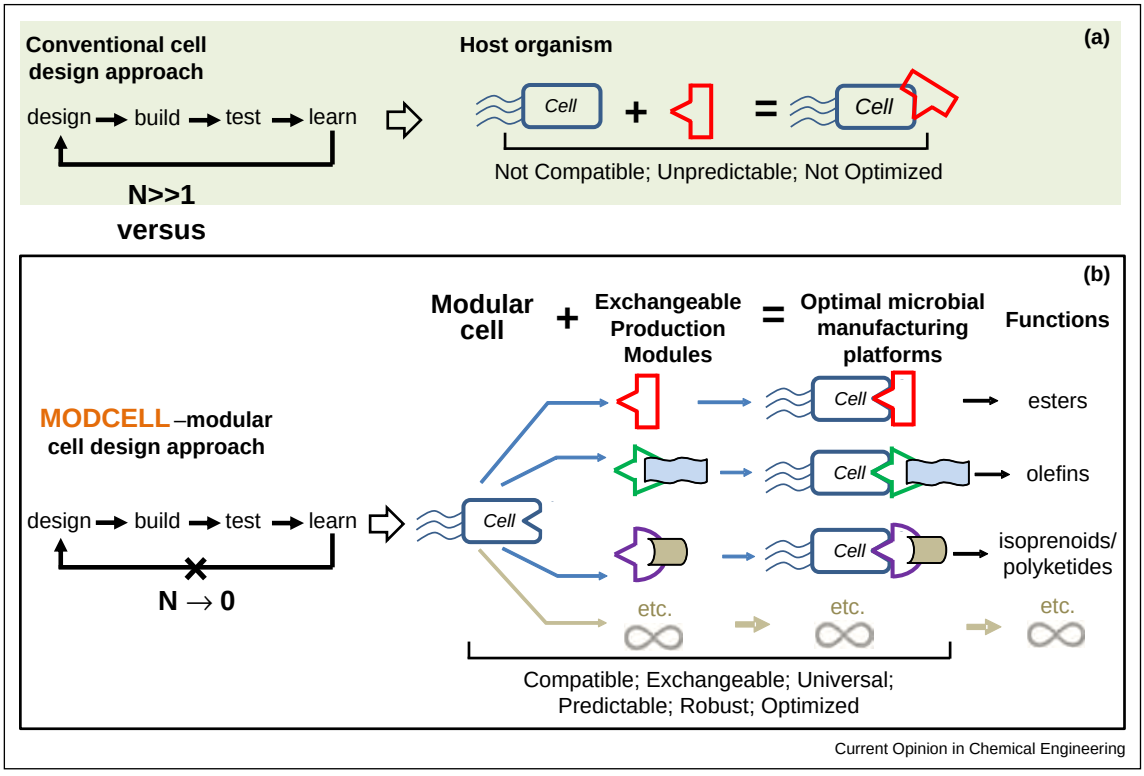
Genome and combinatorial engineering. With the ability to create a large number of variants no longer a limiting step, focus is necessarily shifting toward improved methods of genome and combinatorial engineering. Instead of characterizing individual cells carrying pathway variants one by one, the entire population can be probed simultaneously. The approach is very powerful if one can screen or select for the desirable phenotype, that is, high-yield production of target chemicals. For instance, if a heterologous pathway produces a chemical that can emit light, chemical-hyperproducing strains in the cell population can be isolated via fluorescence-activated cell sorting (FACS) or solid plate screening [74,75]. If production of a chemical is coupled with growth, chemical-hyperproducing strains can be isolated based on simple growth selection. Deep sequencing can also be employed for rapid strain isolation [76]. Both TRMR [77] and TRACE [78] are powerful genome engineering tools to identify potential targets for manipulation; when coupled with MAGE [74] outperforming strains can be generated. While TRMR and MAGE rely solely on the homologous recombination machinery of the host for effective genome editing, recent advances in the CRISPR technology streamlines the genome editing process for rapid strain engineering by utilizing double strand breaks (DSB) and single strand nicks to improve homologous recombination efficiency and perform multiplexing [79]. Even though these methods are powerful to improve engineered phenotypes, they do not often generate the most optimal phenotypes due to incomplete sampling spaces and unforeseen native regulation.

Challenges in heterologous pathway engineering

One significant challenge in engineering heterologous pathways is to deal with non-model organisms because availability of genetic tools and parts libraries does not exist. Parts incompatibility is very common because cellular machinery and its regulation can vary greatly between different microbes. For instance, *Clostridium thermocellum* is one of the potential consolidated bioprocessing thermophiles that is very efficient and robust in degrading complex plant biomass [80]. While many attempts have been explored, no engineered *E. coli* or *S. cerevisae* mutants reported to date could degrade biomass as efficiently as native *C. thermocellum*. It is of great interest to introduce heterologous pathways into *C. thermocellum* and rewire its metabolism for production of specialty biofuels and chemicals while exploiting its biomass-degrading machinery; however, reliable genetic tools and parts availability remain formidable bottlenecks. This challenge presents itself for each new organism and hence breaking these barriers will help industrialization of biology to exploit nature's best. Advances in the CRISPR technology shows potential for manipulating non-model organisms for metabolic engineering applications [81,82].

One other significant challenge in engineering heterologous pathways is the host cell must be re-engineered in an iterative manner to produce different chemicals. This process is laborious and expensive (Figure 2a). Ideally, it is advantageous to develop a blueprint of the universal modular cell that, when combined with optimized exchangeable production modules, creates microbial

Figure 2



(a) Conventional cell design. This approach is laborious and expensive with multiple iterative trial-and-error optimization cycles to engineer a desirable cell ($N \gg 1$). **(b)** MODCELL design. This approach minimizes optimization cycles ($N \rightarrow 0$) and can create desirable microbial platforms in a rapid and systematic manner from the modular cell and exchangeable production modules.

manufacturing platforms in a plug-and-play fashion for optimal production of desirable chemicals (Figure 2b).

Modular cell design toward industrialization of biology

Complexity has kept metabolic engineering confined to time-consuming validation and optimization via multiple design-build-test-learn cycles, despite best efforts to standardize and characterize parts [83]. The grand challenge is how to streamline pathway and strain engineering to rapidly explore the combinatorial space of chemicals (Figure 1). While the high-throughput screening approach is powerful, it is not advantageous and proportionally scaled to deal with the large chemical space; a need for rational modular chassis cell design is required.

Trinh *et al.* has laid out the computational framework named MODCELL (modular cell) for designing modular cells that couple with a diverse class of production modules (i.e., heterologous pathways) [84••]. The MODCELL design principles are formulated such that the modular cell must be auxotrophic and contain the core metabolic pathways that are necessary but insufficient to support cell growth and maintenance under controllable physiological conditions (e.g. anaerobic conditions). The modular cell is designed to be auxotrophic by imposing cofactor imbalance and/or insufficient supply of precursor metabolites required for biosynthesis of biomass and targeted chemicals. To efficiently produce targeted chemicals, the modular cell must be tightly coupled with exchangeable production modules, auxiliary metabolic pathways designed to synthesize target chemicals. The tighter the coupling between the modular cell and production module, the faster the cell growth, substrate consumption, and desirable chemical production rate become. Here, the modularity of the design will enable rapid development of microbial platforms from the modular cell and exchangeable production modules in a plug-and-play fashion, whereas the metabolic coupling will provide powerful selection for production of targeted chemicals at high efficiency during both growth and non-growth associated phases. The MODCELL has shown promise through the demonstrated production of alcohol fuels [85,86,87] and combinatorial biosynthesis of esters [9•].

Modular cells can be constructed from a bottom-up approach where a minimal cell can be synthetically designed and constructed as a host [88–95]. Alternatively, native metabolism of the existing strains must be restricted based on the MODCELL design [84••]. The metabolic coupling

design in modular cells is an ideal chassis for dynamic pathway engineering as well as genome and combinatorial engineering. One could envision the modular cell being constructed with the production module(s) integrated into the host chromosome and transcriptionally controlled by an environmental signal (e.g., light, temperature, IPTG, nutrients, and/or a desirable chemical) in a controllable genetic circuit. This design can impose the auxotrophic characteristic of the modular cell and trigger cell growth and chemical overproduction only if environmental signals are received to activate the production module(s). An additional advantage is the MODCELL design can provide an alternative, secure strategy for microbial containment due to the auxotrophic characteristic of modular cells [96].

Conclusions

Developments in synthetic biology, as well as rapid discovery of novel organisms and their genes, enzymes, and regulatory mechanisms continually grow, enabling metabolic engineers to explore the boundless chemical space. Computational models and curated databases of biological parts play vital roles in driving systematic pathway design, and continued effort into standardizing and improving these sources of information is imperative. Research into building dynamic sensor-regulator devices is still in its infancy, and must be developed if we are to succeed in producing targeted chemicals at high yields, titers, and productivities. The need to develop modular chassis cells for tight coupling with optimized heterologous pathways can potentially minimize the iterative design-build-test-learn cycles. Further, high-throughput technologies will enable rapid selection of the best phenotype amongst the set of the rationally designed pathways. It is of our opinion that achieving the goals of these research areas will lead to the successful engineering of economical microbial factories for efficient production of specialty fuels and chemicals, replacing traditional petroleum-based products.

Acknowledgements

This research was financially supported in part by the NSF awards (NSF #1553250, NSF #1360867, NSF #1553250) as well as the DOE subcontract grant (DE-AC05-000R22725) by the BioEnergy Science Center (BESC), the U.S. Department of Energy Bioenergy Research Center funded by the Office of Biological and Environmental Research in the DOE Office of Science.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick WJ Jr, Hallett JP, Leak DJ, Liotta CL *et al.*: **The path forward for biofuels and biomaterials**. *Science* 2006, **311**:484-489.

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

3. Choi S, Song CW, Shin JH, Lee SY: **Biorefineries for the production of top building block chemicals and their derivatives**. *Metab Eng* 2015, **28**:223-239.

4. Ruddigkeit L, Van Deursen R, Blum LC, Reymond J-L: **Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17**. *J Chem Inform Model* 2012, **52**:2864-2875.

5. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M: **KEGG as a reference resource for gene and protein annotation**. *Nucleic Acids Res* 2016, **44**:D457-D462.

6. Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, Holland TA, Keseler IM, Kothari A, Kubo A *et al.*: **The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases**. *Nucleic Acids Res* 2014, **42**:D459-D471.

7. Choi SY, Park SJ, Kim WJ, Yang JE, Lee H, Shin J, Lee SY: **One-step fermentative production of poly(lactate-co-glycolate) from carbohydrates in *Escherichia coli***. *Nat Biotechnol* 2016, **34**:435-440.

8. Layton DS, Trinh CT: **Expanding the modular ester fermentative pathways for combinatorial biosynthesis of esters from volatile organic acids**. *Biotechnol Bioeng* 2016, **113**:1764-1776.

9. • Layton DS, Trinh CT: **Engineering modular ester fermentative pathways in *Escherichia coli***. *Metab Eng* 2014, **26**:77-88.
This study established a general platform for a modular fermentative ester biosynthesis.

10. Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R: **Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals**. *Nature* 2011, **476**:355-359.

11. •• Cheong S, Clomburg JM, Gonzalez R: **Energy- and carbon-efficient synthesis of functionalized small molecules in bacteria using non-decarboxylative Claisen condensation reactions**. *Nat Biotechnol* 2016.
This study engineered an innovative modular pathway based on coupling of non-decarboxylative Claisen condensation and subsequent beta-reduction reactions to produce a family of ω-diacid and ω1-diacid, hydroxyacids, alcohols, and lactones.

12. Peralta-Yahya PP, Ouellet M, Chan R, Mukhopadhyay A, Keasling JD, Lee TS: **Identification and microbial production of a terpene-based advanced biofuel**. *Nat Commun* 2011, **2**:483.

13. Rui Z, Li X, Zhu X, Liu J, Domigan B, Barr I, Cate JHD, Zhang W: **Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase**. *Proc Natl Acad Sci* 2014, **111**:18237-18242.

14. Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, del Cardayre SB, Keasling JD: **Microbial production of fatty-acid-derived fuels and chemicals from plant biomass**. *Nature* 2010, **463**:559-562.

15. Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB: **Microbial biosynthesis of alkanes**. *Science* 2010, **329**:559-562.

16. Choi YJ, Lee SY: **Microbial production of short-chain alkanes**. *Nature* 2013, **502**:571-574.

17. Jones JA, Vernacchio VR, Lachance DM, Lebovich M, Fu L, Shirke AN, Schultz VL, Cress B, Linhardt RJ, Koffas MA: **ePathOptimize: a combinatorial approach for transcriptional balancing of metabolic pathways**. *Sci Rep* 2015:5.

18. Koma D, Yamanaka H, Moriyoshi K, Ohmoto T, Sakai K: **Production of aromatic compounds by metabolically engineered *Escherichia coli* with shikimate pathway expansion**. *Appl Environ Microbiol* 2012. AEM.01148-01112.

19. •• Galanie S, Thodey K, Trenchard IJ, Interrante MF, Smolke CD: **Complete biosynthesis of opioids in yeast**. *Science* 2015, **349**:1095-1100.
This study successfully engineered the most complex heterologous pathway known to date for the opioid biosynthesis in *S. cerevisae*. This pathway employed an expression of a set of 21 thebaine and 23 hydrocodone enzymes derived from plants, mammals, bacteria, and yeast itself.

20. Jones JA, Toparlak ÖD, Koffas MA: **Metabolic pathway balancing and its role in the production of biofuels and chemicals**. *Curr Opin Biotechnol* 2015, **33**:52-59.

21. Nielsen J, Keasling JD: **Engineering cellular metabolism**. *Cell* 2016, **164**:1185-1197.

22. Wu G, Yan Q, Jones JA, Tang YJ, Fong SS, Koffas MAG: **Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications**. *Trends Biotechnol* 2016.

23. Hatzimanikatis V, Li C, Ionita J, Henry C, Jankowski M: **Exploring the diversity of complex metabolic networks**. *Bioinformatics* 2005, **21**:1603-1609.

24. • Ng CY, Khodayari A, Chowdhury A, Maranas CD: **Advances in de novo strain design using integrated systems and synthetic biology tools**. *Curr Opin Chem Biol* 2015, **28**:105-114.
This work presents an excellent review on pathway and strain engineering design tools.

25. Pertusi DA, Stine AE, Broadbelt LJ, Tyo KE: **Efficient searching and annotation of metabolic networks using chemical similarity**. *Bioinformatics* 2014:btu760.

26. • Zhang X, Tervo CJ, Reed JL: **Metabolic assessment of *E. coli* as a biofactory for commercial products**. *Metab Eng* 2016, **35**: 64-74.
This study employed a computational biology approach to systematically analyze the metabolic feasibility of *E. coli* as a host for production of a large space of non-native chemicals.

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

28. Machado D, Herrgård M: **Co-evolution of strain design methods based on flux balance and elementary mode analysis**. *Metab Eng Commun* 2015.

29. Bogorad IW, Lin T-S, Liao JC: **Synthetic non-oxidative glycolysis enables complete carbon conservation**. *Nature* 2013, **502**: 693-697.

30. Chowdhury A, Maranas CD: **Designing overall stoichiometric conversions and intervening metabolic reactions**. *Sci Rep* 2015:5.

31. • Farasat I, Kushwaha M, Collens J, Easterbrook M, Guido M, Salis HM: **Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria**. *Mol Syst Biol* 2014, **10**:731.
This work presents the SEAMAP framework for designing heterologous pathways with optimized fluxes by manipulating ribosome binding sites.

32. Li M, Elledge S: **Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC**. *Nat Methods* 2007, **4**:251-256.

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

34. Guye P, Li Y, Wroblewska L, Duportet X, Weiss R: **Rapid, modular and reliable construction of complex mammalian gene circuits**. *Nucleic Acids Res* 2013.

35. Zhang Y, Werling U, Edelmann W: **SLiCE: a novel bacterial cell extract-based DNA cloning method**. *Nucleic Acids Res* 2012, **40**:e55.

36. Bitinaite J, Rubino M, Varma KH, Schildkraut I, Vaisvila R, Vaiskunaite R: **USER™ friendly DNA engineering and cloning method by uracil excision**. *Nucleic Acids Res* 2007, **35**:1992-2002.

37. Kok Sd, Stanton LH, Slaby T, Durot M, Holmes VF, Patel KG, Platt D, Shapland EB, Serber Z, Dean J: **Rapid and reliable, DNA assembly via ligase cycling reaction**. *ACS Synth Biol* 2014, **3**: 97-106.

38. Quan J, Tian J: **Circular polymerase extension cloning of complex gene libraries and pathways**. *PLoS ONE* 2009, **4**:e6441.

39. Trubitsyna M, Michlewski G, Cai Y, Elfick A, French CE: **PaperClip: rapid multi-part DNA assembly from existing libraries**. *Nucleic Acids Res* 2014. gku829.

40. Shao Z, Zhao H, Zhao H: **DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways**. *Nucleic Acids Res* 2009, **37**:e16.

41. Annaluru N, Muller H, Mitchell LA, Ramalingam S, Stracquadanio G, Richardson SM, Dymond JS, Kuang Z, Scheifele LZ, Cooper EM: **Total synthesis of a functional designer eukaryotic chromosome**. *Science* 2014, **344**:55-58.

42. Li MZ, Elledge SJ: **MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules**. *Nat Genet* 2005, **37**:311-319.

43. Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S: **A modular cloning system for standardized assembly of multigene constructs**. *PLoS ONE* 2011, **6**:e16765.

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

45. Sarrion-Perdigones A, Falconi EE, Zandalinas SI, Juárez P, Fernández-del-Carmen A, Granell A, Orzaez D: **GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules**. *PLOS ONE* 2011, **6**:e21622.

46. Blake WJ, Chapman BA, Zindal A, Lee ME, Lippow SM, Baynes BM: **Pairwise selection assembly for sequence-independent construction of long-length DNA**. *Nucleic Acids Res* 2010, **38**:2594-2602.

47. Tsuge K, Matsui K, Itaya M: **One step assembly of multiple DNA fragments with a designed order and orientation in Bacillus subtilis plasmid**. *Nucleic Acids Res* 2003, **31**:e133.

48. Colloms SD, Merrick CA, Olorunniji FJ, Stark WM, Smith MC, Osbourn A, Keasling JD, Rosser SJ: **Rapid metabolic pathway assembly and modification using serine integrase site-specific recombination**. *Nucleic Acids Res* 2014, **42**:e23.

49. Zhang L, Zhao G, Ding X: **Tandem assembly of the epothilone biosynthetic gene cluster by in vitro site-specific recombination**. *Sci Rep* 2011:1.

50. •• Hutchison CA, Chuang R-Y, Noskov VN, Assad-Garcia N, Deerinck TJ, Ellisman MH, Gill J, Kannan K, Karas BJ, Ma L: **Design and synthesis of a minimal bacterial genome**. *Science* 2016, **351** aad6253.
This work presents the construction of a synthetic minimal cell with large DNA synthesis and assembly.

51. Lu G, Moriyama EN: **Vector NTI, a balanced all-in-one sequence analysis suite**. *Brief Bioinform* 2004, **5**:378-388.

52. Hillson NJ, Rosengarten RD, Keasling JD: **j5 DNA assembly design automation software**. *ACS Synth Biol* 2011, **1**:14-21.

53. Villalobos A, Ness JE, Gustafsson C, Minshull J, Govindarajan S: **Gene Designer: a synthetic biology tool for constructing artificial DNA segments**. *BMC Bioinform* 2006, **7**:285.

54. Richardson SM, Wheelan SJ, Yarrington RM, Boeke JD: **GeneDesign: rapid, automated design of multikilobase synthetic genes**. *Genome Res* 2006, **16**:550-556.

55. Lorimer D, Raymond A, Walchli J, Mixon M, Barrow A, Wallace E, Grice R, Burgin A, Stewart L: **Gene composer: database software for protein construct design, codon engineering, and gene synthesis**. *BMC Biotechnol* 2009, **9**:1.

56. Flowers D, Thompson RA, Birdwell D, Wang T, Trinh CT: **SMET: systematic multiple enzyme targeting–a method to rationally design optimal strains for target chemical overproduction**. *Biotechnol J* 2013, **8**:605-618.

57. Chowdhury A, Zomorrodi AR, Maranas CD: **k-OptForce: integrating kinetics with flux balance analysis for strain design**. *PLoS Comput Biol* 2014, **10**:e1003487.

58. Thompson RA, Layton DS, Guss AM, Olson DG, Lynd LR, Trinh CT: **Elucidating central metabolic redox obstacles hindering ethanol production in *Clostridium thermocellum***. *Metab Eng* 2015, **32**:207-219.

59. Xu P, Gu Q, Wang W, Wong L, Bower AGW, Collins CH, Koffas MAG: **Modular optimization of multi-gene pathways for fatty acids production in *E. coli***. *Nat Commun* 2013, **4**:1409.

60. Tseng H-C, Prather KLJ: **Controlled biosynthesis of odd-chain fuels and chemicals via engineered modular metabolic pathways**. *Proc Natl Acad Sci* 2012, **109**:17925-17930.

61. Sheppard MJ, Kunjapur AM, Wenck SJ, Prather KL: **Retro-biosynthetic screening of a modular pathway design achieves selective route for microbial synthesis of 4-methyl-pentanol**. *Nat Commun* 2014:5.

62. Ajikumar PK, Xiao W-H, Tyo KEJ, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G: **Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli***. *Science* 2010, **330**:70-74.

63. Schuster S: **Use and limitations of modular metabolic control analysis in medicine and biotechnology**. *Metab Eng* 1999, **1**:232-242.

64. Schuster S, Kahn D, Westerhoff HV: **Modular analysis of the control of complex metabolic pathways**. *Biophys Chem* 1993, **48**:1-17.

65. • Liu D, Evans T, Zhang F: **Applications and advances of metabolite biosensors for metabolic engineering**. *Metab Eng* 2015, **31**:35-43.
This study presents a comprehensive overview of different classes of biosensors that can be harnessed to build synthetic sensor-regular systems for engineering heterologous pathways.

66. Zhang F, Carothers JM, Keasling JD: **Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids**. *Nat Biotechnol* 2012, **30**:354-359.

67. Xu P, Li L, Zhang F, Stephanopoulos G, Koffas M: **Improving fatty acids production by engineering dynamic pathway regulation and metabolic control**. *Proc Natl Acad Sci* 2014, **111**:11299-11304.

68. Soma Y, Tsuruno K, Wada M, Yokota A, Hanai T: **Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch**. *Metab Eng* 2014, **23**:175-184.

69. Tsuruno K, Honjo H, Hanai T: **Enhancement of 3-hydroxypropionic acid production from glycerol by using a metabolic toggle switch**. *Microb Cell Fact* 2015, **14**:1.

70. • Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, Tsai JC, Weissman JS, Dueber JE, Qi LS: **Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds**. *Cell* 2015, **160**:339-350.
This paper established a multilayered activation/inhibition system in *S. cerevisiae* using a Cas9 master switch for controlled production of a family of violaceins.

71. Callura JM, Cantor CR, Collins JJ: **Genetic switchboard for synthetic biology applications**. *Proc Natl Acad Sci* 2012, **109**:5850-5855.

72. • Xiao Y, Bowen CH, Liu D, Zhang F: **Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis**. *Nat Chem Biol* 2016. [advance online publication].
This paper developed a PopQC genetic circuit to dynamically select hyper-producing cells during chemical production phase by coupling the heterologous pathway with antibiotics selection.

73. Taylor ND, Garruss AS, Moretti R, Chan S, Arbing MA, Cascio D, Rogers JK, Isaacs FJ, Kosuri S, Baker D *et al.*: **Engineering an allosteric transcription factor to respond to new ligands**. *Nat Meth* 2016, **13**:177-183.

74. Wang H, Isaacs F, Carr P, Sun Z, Xu G, Forest C, Church G: **Programming cells by multiplex genome engineering and accelerated evolution**. *Nature* 2009, **460**:894-898.

75. Wang HH, Kim H, Cong L, Jeong J, Bang D, Church GM: **Genome-scale promoter engineering by coselection MAGE**. *Nat Meth* 2012, **9**:591-593.

76. Klesmith JR, Bacik J-P, Michalczyk R, Whitehead TA: **comprehensive sequence-flux mapping of a levoglucosan utilization pathway in *E. coli***. *ACS Synth Biol* 2015, **4**:1235-1243.

77. Warner JR, Reeder PJ, Karimpour-Fard A, Woodruff LBA, Gill RT: **Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides**. *Nat Biotechnol* 2010, **28** 856-U138.

78. Zeitoun RI, Garst AD, Degen GD, Pines G, Mansell TJ, Glebes TY, Boyle NR, Gill RT: **Multiplexed tracking of combinatorial genomic mutations in engineered cell populations**. *Nat Biotechnol* 2015, **33**:631-637.

79. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA: **Multiplex genome engineering using CRISPR/Cas systems**. *Science* 2013, **339**:819-823.

80. Olson DG, McBride JE, Joe Shaw A, Lynd LR: **Recent progress in consolidated bioprocessing**. *Curr Opin Biotechnol* 2012, **23**:396-405.

81. Cress BF, Toparlak OD, Guleria S, Lebovich M, Stieglitz JT, Englaender JA, Jones JA, Linhardt RJ, Koffas MA: **CRISPathBrick: modular combinatorial assembly of type II-A CRISPR arrays for dCas9-mediated multiplex transcriptional repression in *E. coli***. *ACS Synth Biol* 2015, **4**:987-1000.

82. Zhu LJ, Holmes BR, Aronin N, Brodsky MH: **CRISPRseek: a bioconductor package to identify target-specific guide RNAs for CRISPR-Cas9 genome-editing systems**. *PLOS ONE* 2014, **9**:e108424.

83. Winkler JD, Halweg-Edwards AL, Gill RT: **The LASER database: formalizing design rules for metabolic engineering**. *Metab Eng Commun* 2015, **2**:30-38.

84. •• Trinh CT, Liu Y, Conner D: **Rational design of efficient modular cells**. *Metab Eng* 2015, **32**:220-231.
This paper laid out a blueprint for designing a modular cell.

85. Trinh CT, Unrean P, Srienc F: **Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses**. *Appl Environ Microbiol* 2008, **74**:3634-3643.

86. Trinh CT, Srienc F: **Metabolic engineering of *Escherichia coli* for efficient conversion of glycerol to ethanol**. *Appl Environ Microbiol* 2009, **75**:6696-6705.

87. Trinh CT: **Elucidating and reprogramming Escherichia coli metabolisms for obligate anaerobic n-butanol and isobutanol production**. *Appl Microbiol Biotechnol* 2012:1-12.

88. Xavier JC, Patil KR, Rocha I: **Systems biology perspectives on minimal and simpler cells**. *Microbiol Mol Biol Rev* 2014, **78**:487-509.

89. Castellanos M, Wilson DB, Shuler ML: **A modular minimal cell model: purine and pyrimidine transport and metabolism**. *Proc Natl Acad Sci U S A* 2004, **101**:6681-6686.

90. Mushegian AR, Koonin EV: **A minimal gene set for cellular life derived by comparison of complete bacterial genomes**. *Proc Natl Acad Sci U S A* 1996, **93**:10268-10273.

91. Forster AC, Church GM: **Towards synthesis of a minimal cell**. *Mol Syst Biol* 2006, **2**:45.

92. Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM *et al.*: **The minimal gene complement of *Mycoplasma genitalium***. *Science* 1995, **270**:397-403.

93. Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA 3rd, Smith HO, Venter JC: **Essential genes of a minimal bacterium**. *Proc Natl Acad Sci U S A* 2006, **103**:425-430.

94. Hutchison CA, Peterson SN, Gill SR, Cline RT, White O, Fraser CM, Smith HO, Venter JC: **Global transposon mutagenesis and a minimal Mycoplasma genome**. *Science* 1999, **286**:2165-2169.

95. Karr Jonathan R, Sanghvi Jayodita C, Macklin Derek N, Gutschow Miriam V, Jacobs Jared M, Bolival B Jr, Assad-Garcia N, Glass John I, Covert Markus W: **A whole-cell computational model predicts phenotype from genotype**. *Cell* 2012, **150**:389-401.

96. Mandell DJ, Lajoie MJ, Mee MT, Takeuchi R, Kuznetsov G, Norville JE, Gregg CJ, Stoddard BL, Church GM: **Biocontainment of genetically modified organisms by synthetic protein design**. *Nature* 2015.