frontiers | Research Topics

Insights in synthetic biology novel developments, current challenges, and future perspectives

2021

Edited by
Jean Marie François and Shota Atsumi

Published in
Frontiers in Bioengineering and Biotechnology

ISSN 1664-8714
ISBN 978-2-8325-2270-7
DOI 10.3389/978-2-8325-2270-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Insights in synthetic biology 2021: Novel developments, current challenges, and future perspectives

Topic editors

Jean Marie François — Institut Biotechnologique de Toulouse (INSA), France

Shota Atsumi — University of California, Davis, United States

Citation

François, J. M., Atsumi, S., eds. (2023). *Insights in synthetic biology 2021: Novel developments, current challenges, and future perspectives*.

Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2270-7

Table of contents

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Editorial
PUBLISHED 13 April 2023
DOI 10.3389/fbioe.2023.1200227

OPEN ACCESS

EDITED AND REVIEWED BY
Ranieri Cancedda,
Independent researcher, Italy

*CORRESPONDENCE
Jean Marie François,
fran_jm@insa-toulouse.fr

RECEIVED 04 April 2023
ACCEPTED 05 April 2023
PUBLISHED 13 April 2023

CITATION
François JM and Atsumi S (2023), Editorial: Insights into synthetic biology 2021: Novel developments, current challenges, and future perspectives. *Front. Bioeng. Biotechnol.* 11:1200227. doi: 10.3389/fbioe.2023.1200227

Editorial: Insights into synthetic biology 2021: Novel developments, current challenges, and future perspectives

Jean Marie François[1,2]* and Shota Atsumi[3]

[1]Toulouse Biotechnology Institute, UMR-INSA—CNRS-INRA, Toulouse, France, [2]Toulouse White Biotechnology Excellent Center, UMS-INSA-INRA-CNRS, Toulouse, France, [3]Department of Chemistry, University of California, Davis, Davis, CA, United States

KEYWORDS

synthetic biology, genetic circuits, modelling, synthetic cells, metabolic engineering, sustainability, bioproduction

Editorial on the Research Topic
Insights in synthetic biology 2021: Novel developments, current challenges, and future perspectives'

This "*Insights into Synthetic Biology 2021*" Research Topic aims to illustrate the remarkable progress of synthetic biology in many different research and application areas. These include the conceptual process of building synthetic cells from scratch, sustainable solutions to our fossil dependency that can now solve greenhouse gas emissions, proposing innovative therapeutic solutions to complex diseases, or fighting infectious diseases. SB's strength also depends on advances in methods, techniques, and algorithms, which are also presented in this Research Topic. Finally, SB is positioned in the field of education as a subject that motivates entrepreneurial aspirations.

The Research Topic "*Insight into Synthetic Biology 2021*"has been very satisfactory in providing 15 contributions, divided into 9 mini-reviews, four original papers, and two opinion papers, which together cover the current developments and challenges in Synthetic Biology (SB). Four articles were devoted to updating tools and methods for better insertion of novel functions or better control of functions in a biological system. Li et al. elegantly summarized the main pitfalls of CRISPR technology in genome editing and proposed some solutions to overcome these difficulties. Liao et al. demonstrated high antibiotic-free plasmid stability upon insertion of the *hok/sok* gene system into this plasmid, opening up the possibility of using it in large-scale process production. Regarding the methods required to implement and control new functions in biological systems, Abraha and Marchisio reported the implementation in yeast of a bacterial ClpXP protein degradation system and showed that SBML level 3 is perfectly suited to describe the modular function of this orthogonal synthetic gene circuit constructed in yeast. At a higher level of complexity are the systemic principles of gene regulatory networks, whose understanding in terms of architecture, organization, dynamics, and evolution is important for the optimal engineering of biological systems, as summarized in the study by Freyre-González et al. These authors coined the term "concilion," which corresponds to a group of structural genes with their local regulators

dedicated to a single well-defined function, as opposed to regulons or modulons, which are under the control of a specific or global regulator and that can be controlled by a diversity of functions. A biological example illustrating the concept of a concilion is provided by the response of bacteria or yeast to multi-stress, as this concilion will include several regulons organized in a regulatory cascade, each under the control of a specific regulator.

As SB is an engineering science several articles in this Research Topic are devoted to applications in different fields of biotechnology. In the field of industrial biotechnology, SB proposes sustainable alternatives aimed at reducing our dependence on fossil resources through a radical rewiring of microbial metabolism for the scalable production of a wide range of drop-in alternatives with carbon neutrality and a circular bioeconomy as the ultimate goals. Of particular relevance is the growing interest in identifying and engineering microbial systems capable of generating valuable chemicals from C1-carbon (CO_2, methanol, and methane), thereby reducing the use of sugar feedstock and capturing atmospheric CO_2, both of which, according to Carruthers and Lee, can transform microbial bioproduction into a more techno-economically sustainable industrial biotechnology. However, as pointed out by Treece et al., cyanobacteria would be remarkable cell factories producing chemicals from CO_2 and sunlight if major obstacles due to low RuBisCO-dependent carbon fixation efficiency and poor growth could be overcome, which could be accomplished in the near future with the tools of SB combined with the isolation of faster-growing cyanobacteria. In support of this claim, Ferreira et al. reported the metabolic engineering of the cyanobacterium *Synechocystis* sp. PCC 6803 for glycine-betaine, showing that the production of this compatible soluble comes mostly at the expense of glycogen degradation, suggesting fixed carbon and/or energy were likely the limiting factors for this production. Synthetic methylotrophy, which is the introduction of a non-native methanol utilization pathway into a model host microorganism, is another promising topic for sustainable chemical bioproduction that also faces several scientific and technological difficulties, including the inability to grow on methanol as the sole carbon source and the very high toxicity of the intermediate formaldehyde during methanol assimilation. A solution proposed by Peiro et al. which exists in naturally methylotrophic organisms, would be to isolate the initial metabolic methanol assimilation in a compartment from its central metabolic network. SB is also relevant in many other applications like food, feed, and medicine. This is illustrated in this Research Topic by two interesting papers. The first one, presented by Cruz et al. deals with the concept of bioengineered probiotics endowed with new functionalities able to control and/or kill foodborne pathogens, induce immunity against pathogens, and eventually neutralize pathogen toxins. While there may be several benefits of this live cell biotherapeutic, including high specificity to the target, self-limiting due to lack of selection, and cost-effectiveness, several technical and ethical barriers still need to be overcome: competing with resident microbes in the gut, the induction of pathogen resistance, and biocontainment. Bioengineered probiotics have also been proposed as a next-generation product for the treatment of inflammatory bowel disease (IBD). IBD is a complex, chronic inflammatory bowel disease that primarily includes Crohn's disease and ulcerative colitis. Although very complex, it is generally accepted that there is an association between intestinal microbiota-derived metabolites and IBD, which raises several possible clinical interventions that include delivery of soluble effector proteins inhibiting inflammatory effector genes by engineered microbes, fecal microbial transplantation, and bacteriophage therapy. As interesting and attractive as these possibilities are, the main obstacle remains the safety of genetically modified organisms in these innovative therapeutic approaches.

Finally, the demonstration that SB is an elective field in science is illustrated on the one hand by the article by Stano (2022), which presents his opinion on the design, function, and application of artificial cellular systems, and by the report of an iGEM-inspired applied program given to undergraduate students, showing improved skills, talent, and entrepreneurial motivation (Gill et al., 2022).

Author contributions

JF wrote the draft of the manuscript which was revised and improve by SA. Both authors approved the final version of the submitted manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Gill, H., Ahsan, M., Khalil, Y., Feng, V., Pearce, J., Sharma, T., et al. (2022). The BioExperience Research and Entrepreneurship Challenge: An iGEM-inspired applied research program for BIOSTEM talent and skills development. *Front. Bioeng. Biotechnol.* 10, 1046723. doi:10.3389/fbioe.2022.1046723

Stano, P. (2022). A four-track perspective for bottom-up synthetic cells. *Front. Bioeng. Biotechnol.* 10, 1029446. doi:10.3389/fbioe.2022.1029446

frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 07 January 2022
doi: 10.3389/fbioe.2021.821075

Heterologous Production of Glycine Betaine Using *Synechocystis* sp. PCC 6803-Based Chassis Lacking Native Compatible Solutes

Eunice A. Ferreira[1,2,3], *Catarina C. Pacheco*[1,2], *João S. Rodrigues*[1,2,4†], *Filipe Pinto*[1,2], *Pedro Lamosa*[5], *David Fuente*[6], *Javier Urchueguía*[6] *and Paula Tamagnini*[1,2,4*]

[1]*I3S—Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal,* [2]*IBMC—Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal,* [3]*ICBAS—Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal,* [4]*Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal,* [5]*Instituto de Tecnologia Química e Biológica António Xavier, ITQB NOVA, Oeiras, Portugal,* [6]*Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas, Universitat Politècnica de València, València, Spain*

OPEN ACCESS

Edited by:
Shota Atsumi,
University of California, Davis, United States

Reviewed by:
Xuefeng Lu,
Qingdao Institute of Bioenergy and Bioprocess Technology (CAS), China
Jianping Yu,
National Renewable Energy Laboratory (DOE), United States
Weiwen Zhang,
Tianjin University, China

***Correspondence:**
Paula Tamagnini
pmtamagn@ibmc.up.pt

†Present Address:
João S. Rodrigues,
Department of Chemistry—Ångström, Uppsala University, Uppsala, Sweden

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *23 November 2021*
Accepted: *15 December 2021*
Published: *07 January 2022*

Citation:
Ferreira EA, Pacheco CC, Rodrigues JS, Pinto F, Lamosa P, Fuente D, Urchueguía J and Tamagnini P (2022) Heterologous Production of Glycine Betaine Using Synechocystis *sp. PCC 6803-Based Chassis Lacking Native Compatible Solutes. Front. Bioeng. Biotechnol. 9:821075. doi: 10.3389/fbioe.2021.821075*

Among compatible solutes, glycine betaine has various applications in the fields of nutrition, pharmaceuticals, and cosmetics. Currently, this compound can be extracted from sugar beet plants or obtained by chemical synthesis, resulting in low yields or high carbon footprint, respectively. Hence, in this work we aimed at exploring the production of glycine betaine using the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 as a photoautotrophic chassis. *Synechocystis* mutants lacking the native compatible solutes sucrose or/and glucosylglycerol—Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS*—were generated and characterized. Under salt stress conditions, the growth was impaired and accumulation of glycogen decreased by ~50% whereas the production of compatible solutes and extracellular polymeric substances (capsular and released ones) increased with salinity. These mutants were used as chassis for the implementation of a synthetic device based on the metabolic pathway described for the halophilic cyanobacterium *Aphanothece halophytica* for the production of the compatible solute glycine betaine. Transcription of ORFs comprising the device was shown to be stable and insulated from *Synechocystis'* native regulatory network. Production of glycine betaine was achieved in all chassis tested, and was shown to increase with salinity. The introduction of the glycine betaine synthetic device into the Δ*ggpS* background improved its growth and enabled survival under 5% NaCl, which was not observed in the absence of the device. The maximum glycine betaine production [64.29 μmol/gDW (1.89 μmol/mg protein)] was reached in the Δ*ggpS* chassis grown under 3% NaCl. Taking into consideration this production under seawater-like salinity, and the identification of main key players involved in the carbon fluxes, this work paves the way for a feasible production of this, or other compatible solutes, using optimized *Synechocystis* chassis in a pilot-scale.

Keywords: compatible solutes, cyanobacteria, glycine betaine, glucosylglycerol, salt stress, sucrose, *Synechocystis*, synthetic biology

Abbreviations: CPS, capsular polysaccharides; CS, compatible solutes; DW, dry weight; EPS, extracellular polymeric substances; FW, fresh weight; GB, glycine betaine; GG, glucosylglycerol; NMR, nuclear magnetic resonance; RPS, released polysaccharides; RT-qPCR, reverse transcription quantitative PCR.

INTRODUCTION

Microorganisms can cope with environmental stresses such as temperature, salinity or drought via the production of compatible solutes (CS)—low-molecular weight organic compounds highly soluble in water that can accumulate intracellularly up to molar concentrations, without interfering with the cell metabolism (Klähn and Hagemann, 2011). CS belong to different chemical classes including sugars (e.g., sucrose, trehalose), polyols (e.g., glycerol, sorbitol), heterosides (e.g., glucosylglycerol, floridoside), and amino acids or their derivatives (e.g., proline, glutamate, glycine betaine, ectoine) (Kirsch et al., 2019). Glycine betaine (or *N*,*N*,*N*-trimethylglycine) is an ubiquitous solute that can be found in bacteria, plants and mammals, being mostly synthesized by the two-step oxidation of choline to betaine aldehyde and subsequently to glycine betaine (Landfald and Strøm, 1986; Grossman and Hebert, 1989; Rathinasabapathi et al., 1997). Later on, the synthesis of glycine betaine *via* a three-step methylation of glycine was described in extremely halophilic bacteria (Nyyssola et al., 2000; Waditee et al., 2003). This biosynthetic pathway involves two *N*-methyltransferases: the glycine-sarcosine-*N*-methyltransferase (GSMT) that catalyzes the methylation of glycine and sarcosine, and the dimethylglycine-*N*-methyltransferase (DMT) that converts dimethylglycine to glycine betaine. This CS has a strong stabilizing effect on biomolecules, by maintaining their structure and function (Guinn et al., 2011; Stadmiller et al., 2017), and thus conferring drought-, osmo- and thermo-protection to cells (Caldas et al., 1999; Holmström et al., 2000; Cleland et al., 2004; You et al., 2019). Moreover, glycine betaine plays an important physiological role as methyl group donor with beneficial stress-mitigating effects in humans (Lever and Slow, 2010; Day and Kempson, 2016), and up-regulating antioxidant defense systems in plants (Rady et al., 2018; Sun et al., 2020). Due to these interesting properties, glycine betaine is a value-added compound with applications in human nutrition, animal feed, agriculture, pharmaceuticals, and cosmetics (Eklund et al., 2005; Lever and Slow, 2010; Nsimba et al., 2010; Cholewa et al., 2014; Day and Kempson, 2016; Dikilitas et al., 2020). Most of the commercially available glycine betaine is extracted from sugar beets (*Beta vulgaris*) (Heikkila et al., 1982), resulting in relatively low yields and rendering this organic production an expensive process. Alternatively, glycine betaine can be produced by chemical synthesis, but this process, although cheaper, has a high environmental impact increasing the carbon footprint (DuPont, 2015; Kar et al., 2015). A more sustainable and cost-effective production is highly desirable, and thus cyanobacteria emerge as promising chassis for the production of compatible solutes and other products of interest. Their photoautotrophic metabolism enables the sequestration and conversion of atmospheric CO_2 into organic compounds using sunlight and water as energy and electron sources, respectively (Knoll, 2008; Ananya and Ahmad, 2014). Therefore, they are being increasingly studied to be used as solar-powered cell factories for many biotechnological applications including the production of e.g., alcohols, alkanes, hydrogen, sugars, and terpenoids (Hays and Ducat, 2015; Lindblad, 2018; Sadvakasova et al., 2020; Wang et al., 2020; Rodrigues and Lindberg, 2021). Among cyanobacteria, the unicellular *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*) is the best studied strain, and the vast array of data generated over the past decades allowed the construction of genome-scale metabolic models to predict system's behavior (Montagud et al., 2010; Montagud et al., 2011; Joshi et al., 2017; Gopalakrishnan et al., 2018). Moreover, various molecular and synthetic biology tools are now available for the genetic manipulation and engineering of this particular cyanobacterial strain (Huang et al., 2010; Heidorn et al., 2011; Pinto et al., 2015; Pacheco et al., 2021). *Synechocystis* is a freshwater strain and thus moderately halotolerant, relying on the biosynthesis of the compatible solutes sucrose, glutamate and glucosylglycerol to maintain the osmotic pressure under stress conditions (Klähn and Hagemann, 2011; Iijima et al., 2020).

In this study, *Synechocystis* knockout mutants in the biosynthetic pathways producing the main native compatible solutes sucrose or/and glucosylglycerol were generated to serve as chassis for the production of value-added compounds. The genome-scale metabolic model *i*Syn811 was used to simulate the production rates of different heterologous CS and the highest rate was predicted for glycine betaine. As a proof-of-concept, we explored the production of this compatible solute using our *Synechocystis*-based chassis. For this purpose, a synthetic device based on the biosynthetic gene cluster from the halophilic cyanobacterium *Aphanothece halophytica* was constructed and introduced in the different chassis. Besides showing the production of glycine betaine and validating the functionality of the synthetic device, the characterization of the strains contributes to a better understanding of the mechanisms used by the cells to maintain homeostasis and cope with different levels of salinity.

MATERIALS AND METHODS

Reagents and Enzymes

The media components and other reagents were obtained from Fisher Scientific (United States), Merck (Germany) or Sigma Aldrich (United States), and noble agar from Difco (United States). All DNA-modifying enzymes and polymerases were purchased from Thermo Fisher Scientific (United States) and Promega (United States), and standard molecular biology kits were obtained from NZY Tech (Portugal). The Sanger sequencing and oligo synthesis services were provided by STAB VIDA, Lda (Portugal).

Organisms and Culture Conditions

Wild-type and mutants of the unicellular, non-motile cyanobacterium *Synechocystis* sp. PCC 6803 substrain GT-Kazusa (Kanesaki et al., 2012; Trautmann et al., 2012)

(obtained from the Pasteur Culture Collection, France) were maintained in Erlenmeyer flasks batch cultures with BG11 medium (Stanier et al., 1971) at 30°C with orbital shaking (150 rpm) under a 12 h light/12 h dark regimen. Light intensity was 25 μE/m^2/s in all experiments and Cosine-corrected irradiance was measured using a Dual Solar/Electric Quantum Meter (Spectrum Technologies, Inc., United States). For solid BG11, the medium was supplemented with 1.5% (wt/vol) noble agar, 0.3% (wt/vol) sodium thiosulfate and 10 mM TES-KOH buffer, pH 8.2 (Stanier et al., 1971). For the selection and maintenance of mutants, BG11 medium was supplemented with chloramphenicol (Cm, 10–20 μg/ml). For cloning purposes, *E. coli* strains DH5α and XL1-blue were used. Cells were grown at 37°C in LB medium (Sambrook and Russel, 2001), supplemented with kanamycin (Km, 50 μg/ml) or Cm (34 μg/ml).

DNA and RNA Extraction

Cyanobacterial genomic DNA (gDNA) extraction was performed according to the procedure described previously (Tamagnini et al., 1997). For RNA extraction, 50 ml of *Synechocystis* culture at OD_{730} ≈ 1 was centrifuged for 10 min at 4,470 *g*; cell pellets were treated with RNAprotect Bacteria Reagent (Qiagen, Germany) according to instructions, and stored at −80°C. RNA was extracted using the TRIzol® Reagent (Ambion) according to the method described previously (Leitão et al., 2006) with the following adaptations: the cells were disrupted using a FastPrep®-24 (MP Biomedicals) in 2 cycles of 1 min at 4.0 m/s and the RNA samples were treated with 1 U of RQ1 RNase-Free DNase (Promega) according to manufacturer's instructions.

Glycine Betaine Device: Design, DNA Synthesis, and Assembly

The synthetic construction for the synthesis of glycine betaine (Ahbet) was designed based on *gsmt* (encoding the glycine-sarcosine-*N*-methyltransferase) and *dmt* (encoding the dimethylglycine-*N*-methyltransferase) Open Reading Frames (ORFs) from the cyanobacterium *Aphanothece halophytica*, and the *metX* (*sll0927*, encoding *S*-adenosyl-methionine synthase) ORF from *Synechocystis*. All the ORF sequences were codon optimized for *Synechocystis* using the Gene Designer 2.0 software (DNA 2.0, Inc., United States), restriction sites incompatible with the BioBrick™ standard RFC [10] were eliminated and double stop codons included. Each ORF is preceded by the BioBrick™ (BB) ribosome binding site (RBS) BBa_B0030 and the double terminator BBa_B0015 was included after the *metX* ORF. In addition, the synthetic construction is flanked by the prefix and suffix sequences of the BB RFC [10] standard. All the BB sequences were retrieved from the Registry of Standard Biological Parts (parts.igem.org). Subsequently, the sequence of the glycine betaine synthetic construction flanked by the BB prefix, the double terminator and BB suffix was synthesized and cloned into *Sma*I digested pBluescript II SK(-) (Epoch Life Science, Inc., United States).

To construct the glycine betaine device, the synthesized Ahbet construct was assembled with the synthetic promoter $P_{trc.x.lacO}$, previously characterized in *Synechocystis* (Ferreira et al., 2018). For this purpose, the Ahbet was PCR-amplified from the plasmid pBSK with the pUC primers (Additional file 1: **Supplementary Table S4**), using Phusion high-fidelity DNA polymerase, according to the manufacturer's instructions. The PCR product was purified using NZYGelpure kit, digested with *Xba*I and *Pst*I and cloned downstream of $P_{trc.x.lacO}$ in the pJ201 plasmid (digested with *Spe*I and *Pst*I, and dephosphorylated). The generated $P_{trc.x.lacO}$::Ahbet device was excised from the pJ201 plasmid with *Xba*I and *Spe*I, and transferred to pSEVA351 shuttle vector (Silva-Rocha et al., 2013), digested with *Xba*I. The pSEVA351 was obtained from the "Standard European Vector Architecture" repository and is comprised by the broad-host-range replicon RSF1010 and the chloramphenicol antibiotic marker. The assembly and transfer of the synthetic device was confirmed by PCR, restriction analysis and Sanger sequencing.

Construction of Integrative Plasmids for the Generation of CS Mutants

The construction of integrative plasmids for the knockout of *ggpS* (glucosylglycerol-phosphate synthase) and *sps* (sucrose-phosphate synthase) genes was performed as described previously (Pinto et al., 2012). Briefly, the plasmids were based on pGEM-T® Easy (Promega, United States) and contain the *Synechocystis* chromosomal regions flanking the *ggpS* or the *sps* gene. The 5′- and 3′-flanking regions were amplified from the cyanobacterium's genome using *Pfu* DNA polymerase and the primer pairs 5-O/5-I and 3-O/3-I (Additional file 1: **Supplementary Table S4**), respectively. Subsequently, the purified PCR fragments were fused by Overlap Extension PCR using primers 5-O/3-O and 80 ng of each amplicon. The resulting product was purified and cloned into the vector pGEM-T® Easy, according to the manufacturer's instructions, originating the pGDggpS, and the pGDsps plasmids (**Table 1**). A selection cassette, containing the *nptII* gene (conferring resistance to neomycin and kanamycin) and the *sacB* gene (conferring sensitivity to sucrose), was PCR amplified from the plasmid pK18mobsacB (Schafer et al., 1994) with specific primers (Additional file 1: **Supplementary Table S4**). The amplicon was then cloned into the *Age*I/*Sma*I restriction site of pGDggpS or pGDsps plasmids, generating the pGDggpS.KS and the pGDsps.KS plasmids, respectively (**Table 1**). All constructs were confirmed by sequencing.

Generation of *Synechocystis* CS Knockout Mutants

Synechocystis was transformed based on the protocol described by Williams (1988) with modifications. *Synechocystis* cultures were grown under standard conditions to an OD_{730} ≈ 0.5. Cells were harvested by centrifugation at 3,850 *g* for 10 min; and then resuspended in BG11 to a final OD_{730} ≈ 2.5. A 500 μL aliquot of these cells was used (per transformation) and incubated with

TABLE 1 | List of plasmids used to transform *Synechocystis*.

Designation	Plasmid	Description	Reference/ Source
$P_{trc.x.lacO}$:: Ahbet	pSEVA351	Ahbet synthetic construction under the control of the $P_{trc.x.lacO}$ promoter	This study
pGDggpS	pGEM-T® Easy	pGEM-T easy vector containing the two regions for double homologous recombination targeting the *ggpS locus*	Ferreira et al. (2018)
pGDggpS.KS	pGEM-T® Easy	pGEM-T easy vector containing the *nptII* and *sacB* genes flanked by the two regions for double homologous recombination targeting the *ggpS locus*	This study
pGDsps	pGEM-T® Easy	pGEM-T easy vector containing the two regions for double homologous recombination targeting the *sps locus*	This study
pGDsps.KS	pGEM-T® Easy	pGEM-T easy vector containing the *nptII* and *sacB* genes flanked by the two regions for double homologous recombination targeting the *sps locus*	This study

purified pGDggpS.KS or pGDsps.KS plasmids, at a final DNA concentration of 20 μg/ml, for 5 h in light at 30°C. Cells were then spread onto Immobilon™-NC membranes (0.45 μm pore size, 82 mm, Millipore, United States) resting on solid BG11 plates, incubated at 25°C under low light, and transferred to selective solid BG11 plates supplemented with 10 μg/mL km after 24 h. Transformants were observed after 1–2 weeks. For complete segregation, Km-resistant colonies were streaked on BG11 plates with increasing Km concentrations (up to 500 μg/ml), and finally transferred into liquid medium. Mutants were then tested for sucrose sensitivity and confirmed by PCR and Southern blot (for details see below). Subsequently, to remove the selection markers from the insertion mutants, cells were transformed as described above with the pGDggpS or the pGDsps plasmids, and the mutants were selected on solid BG11 containing 10% (wt/vol) sucrose. These mutants were also screened for Km-sensitivity. The double mutant Δ*sps*Δ*ggpS* was generated by deleting the *ggpS* gene from the Δ*sps* background following the abovementioned protocol. The full segregation of the mutants was confirmed by PCR using GoTaq® G2 Flexi DNA Polymerase, together with specific primers (Additional file 1: **Supplementary Table S4**), according to manufacturer's instructions. Mutant segregation was also confirmed by Southern blots that were performed using 4 μg of genomic DNA of the wild-type and mutants, digested with *Mun*I (wild-type, Δ*sps.KS*, Δ*sps*, Δ*ggpS.KS*, and Δ*ggpS*), and *Ava*II (wild-type, Δ*sps*Δ*ggpS.KS*, and Δ*sps*Δ*ggpS*). The DNA fragments were separated by electrophoresis on a 1% (wt/vol) agarose gel and blotted onto Hybond™-N membrane (GE Healthcare, United States). Probes covering the 5′ flanking region of the *ggpS* or 3′ flanking region of the *sps* genes were amplified by PCR (using primers indicated in Additional file 1: **Supplementary Table S4**), and labeled using the DIG DNA labelling kit (Roche Diagnostics GmbH, Germany), according to the manufacturer's instructions. Hybridization was performed overnight at 65°C, and digoxigenin-labelled probes were detected by chemiluminescence using CPD-star (Roche) in a Chemi Doc™ XRS+ Imager (Bio-Rad, United States).

Introduction of the Glycine Betaine Synthetic Device Into *Synechocystis*

The pSEVA351 plasmid containing the synthetic device $P_{trc.x.lacO}$::Ahbet (**Table 1**; sequence provided in **Supplementary Datasheet S2**) was introduced into *Synechocystis* by electroporation following the protocol described previously (Ferreira et al., 2018). The presence of the synthetic device was confirmed in *Synechocystis* transformants by PCR using specific primers (Additional file 1: **Supplementary Table S4**), as described by Ferreira et al. (2018).

Growth Experiments

Pre-cultures of *Synechocystis* wild-type and mutants were inoculated in BG11 medium (supplemented with 10 μg/ml Cm, when appropriate) and grown in an orbital shaker (150 rpm), at 30°C under a 12 h light (25 μE/m^2/s)/12 h dark regimen. The cultures were grown to an $OD_{730} \approx 2$ and, subsequently, diluted in fresh BG11 medium without antibiotic to a final $OD_{730} \approx 0.5$. Fifty milliliters of the dilution were transferred to 100 ml Erlenmeyer flasks without NaCl or containing 3, 5, or 7% (wt/vol) NaCl (510, 860, and 1,200 mM NaCl, respectively), previously sterilized. These cultures were maintained in the same conditions as the pre-cultures and growth was monitored for at least 16 days, by measuring the optical density at 730 nm (OD_{730}) and determining the chlorophyll *a* (chl *a*) content as described by Meeks and Castenholz (1971). All the growth experiments included, at least, three biological replicates with technical duplicates.

Total Carbohydrate Content, Released and Capsular Polysaccharides Measurements

Total carbohydrate content and RPS were determined as previously described (Mota et al., 2013). Briefly, 10 ml of culture samples were dialyzed (12–14 kDa molecular weight cutoff; CelluSepT4, Orange Scientific) against at least 10 volumes of distilled water, 3 or 5% (wt/vol) NaCl solutions (identical to the growth medium), for at least 24 h. One milliliter of the collected sample was used to spectrophotometrically quantify the total carbohydrate content by the phenol-sulfuric acid method (Dubois et al., 1956), whereas 5 ml of the dialyzed sample was centrifuged at 3,857 *g* for 10 min at RT, and the cell-free supernatant was used to determine the RPS. For CPS quantification, the procedure was performed as described previously (Pereira et al., 2019). Five milliliters of

dialyzed cultures were centrifuged at 3,857 *g* for 10 min at RT, the cell pellet was resuspended in water and boiled for 15 min. After centrifugation as described previously, the cell-free supernatant was used for CPS measurement by the phenol-sulfuric acid method (Dubois et al., 1956). Total carbohydrate content, RPS and CPS were normalized by chl *a* content. All experiments included, at least, three biological replicates with technical triplicates.

Glycogen Extraction and Quantification

Glycogen extraction was performed as described previously (Ernst et al., 1984). Ten milliliters of cell culture were collected 1 h after the transition between the dark and the light phase. Samples were centrifuged, and the cell pellets suspended in 100 µL of distilled water and 400 µL of 30% (wt/vol) KOH was added. The mixture was incubated at 100°C for 90 min and then quickly cooled on ice. Six hundred µL of ice-cold absolute ethanol were added, and the mixture was incubated on ice for 2 h. The mixture was centrifuged for 5 min at maximum speed and 4°C. The supernatant was discarded, and the isolated glycogen was washed three times with 500 µL of ice-cold absolute ethanol and dried at 60°C. Glycogen quantification was performed by the phenol-sulfuric acid method (Dubois et al., 1956), and normalized by chl *a* content. Experiments included, at least, three biological replicates with technical triplicates.

Optical Microscopy

Cultures of *Synechocystis* wild-type (WT) and the Δ*sps* mutant were grown in BG11 or BG11 supplemented with 5% (wt/vol) NaCl as stated above (initial $OD_{730} \approx 0.5$). Four days after inoculation, cells were stained with 0.5% (wt/vol) of Alcian Blue (in 3% (vol/vol) acetic acid) in 1:1 (culture:dye) ratio. This mixture was added to 10 µL of 1% (wt/vol) low-melting point agarose beds (dissolved in BG11 medium) and covered with a coverslip. The preparations were observed using the light microscope Olympus DP25 Camera software Cell B.

Transcription Analysis by RT-qPCR

After extraction (for details see above), RNA concentration and purity (the ratios A_{260}/A_{280} and A_{260}/A_{230}) were measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., United States). The quality and integrity of the RNA samples was also inspected in 1% (wt/vol) agarose gel performed by standard protocols using TAE buffer. The absence of genomic DNA contamination was checked by PCR, in reaction mixtures containing 0.5 U of GoTaq® G2 Flexi DNA Polymerase, 1x Green GoTaq Flexi buffer, 200 µM of each dNTP, 1.5 mM $MgCl_2$, 0.25 µM of each rnpB primer (Additional file 1: **Supplementary Table S4**), and 200 ng of total RNA. The PCR profile was: 5 min at 95°C followed by 25 cycles of 20 s at 95°C, 20 s at 56°C and 20 s at 72°C, and a final extension at 72°C for 5 min. The PCR reactions were run on 1% (wt/vol) agarose gel as described above. For cDNA synthesis, 1 µg of total RNA was transcribed with the iScript™ Reverse Transcription Supermix for RT-qPCR (Bio-Rad) in a final volume of 20 µL, following the manufacturer's instructions. A control PCR was performed using 1 µL of cDNA as a template, the BD16S primers (Additional file 1: **Supplementary Table S4**), and the same reaction conditions and PCR program described above. Five-fold standard dilutions of the cDNAs were made (1/5, 1/25, 1/125, and 1/625) and stored at −20°C. RT-qPCRs were performed on Hard-Shell 384-Well PCR Plates (thin wall, skirted, clear/white) covered with Microseal® B adhesive seal (Bio-Rad). The reactions (10 µL) were manually assembled and contained 0.125 µM of each primer (Additional file 1: **Supplementary Table S4**), 5 µL of iTaq™ Universal SYBR® Green Supermix (Bio-Rad) and 1 µL of template cDNA (dilution 1/5). The PCR protocol used was: 3 min at 95°C followed by 45 cycles of 30 s at 95°C, 30 s at 56°C, and 30 s at 72°C. In the end, a melting curve analysis of the amplicons (10 s cycles between 55 and 95°C with a 0.5°C increment per cycle) was conducted. Standard dilutions of the cDNA were used to check the relative efficiency and quality of primers, and negative controls (no template cDNA) included (for more details on RT-qPCR parameters see Additional file 1: **Supplementary Table S5**). RT-qPCRs were performed with three biological replicates and technical triplicates of each cDNA sample in the CFX384 Touch™ Real-Time PCR Detection System (Bio-Rad). The data obtained were analyzed using the Bio-Rad CFX Maestro™ 1.1 software, implementing an efficiency-corrected delta Cq method (ΔCq). This method was used since the target genes *gsmt*, *dmt*, and *metX* were validated as reference genes using the reference gene selection tool available in the Maestro™ software. For this reason, the relative expression of the targets is represented instead of the usual relative normalized expression. Statistical analysis was performed using a one-way ANOVA using the same software, and tests were considered significant if $p < 0.05$. The amplicon sizes were checked by agarose gel electrophoresis, and the DNA sequence was confirmed by Sanger sequencing. These experiments were compliant with the MIQE guidelines (Bustin et al., 2009), to promote the effort for experimental consistency and transparency, and to increase the reliability and integrity of the results obtained.

Compatible Solutes Quantification

Cultures of *Synechocystis* wild-type (WT), the Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS* mutants and the strains harboring the GB device (WT, Δ*ggpS*, and Δ*sps*Δ*ggpS* backgrounds) were grown in BG11 or BG11 supplemented with 3% or 5% (wt/vol) NaCl, as described above, at an initial $OD_{730} \approx 0.5$. The quantification of the CS—sucrose, glutamate, glucosylglycerol, and glycine betaine—was performed using 500 ml culture (distributed in 50 ml cultures in 100 ml Erlenmeyer flasks). Four days after inoculation, cells were harvested by centrifugation at 4,470 *g* for 10 min at room temperature (RT). In the case of the strains harboring the GB device, the extracellular medium was stored at −80°C, for further lyophilization and CS extraction. Cells were washed using 100 ml of cold distilled water, 3 or 5% (wt/vol) NaCl solutions (identical to the growth medium). Centrifugation was repeated and the cell pellets were resuspended in 50 ml of

the respective solutions. From this suspension, a 0.5 ml aliquot was centrifuged and stored at −20°C to be used later for protein quantification. The remaining cell suspension was centrifuged at 4°C and the cell pellet was stored at −20°C. Ethanol-chloroform extraction of the CS was performed as described in Ferreira et al. (2018) with adaptations. Briefly, cell pellets or lyophilized extracellular medium were suspended in 25 ml of 80% (vol/vol) ethanol and subsequently transferred to a 100 ml round flask containing a magnetic stir bar. The flask was connected to a coil condenser (circulating cold water) and heated at 100°C with stirring, for 10 min. The suspension was transferred to a 50 ml tube and centrifuged at 4,000 *g* for 10 min at RT. The supernatant was stored and the pellet resuspended in 20 ml of 80% (vol/vol) ethanol for a new extraction process. The remainder protocol was performed as described in Santos et al. (2006). Detection, identification and quantification of CS was performed by proton NMR. To that effect, freeze-dried extracts were dissolved in 1 ml of D_2O and a known amount of sodium formate was added to serve as an internal concentration standard. Spectra were acquired at 25°C on a Bruker Avance III 800 spectrometer (Bruker, Rheinstetten, Germany) working at a proton operating frequency of 800.33 MHz, equipped with a 5 mm, three channel, inverse detection cryoprobe TCI-z H&F/C/N with pulse-field gradients. A 3 s soft pulse was applied before the excitation pulse, to pre-saturate the water signal. Spectra were acquired under fully relaxed conditions (flip angle 60°; repetition delay of 60 s) so that the area of the NMR signals was proportional to the amount of the different protons in the sample. Integration of the signals was performed using the tools available in the TopSpin software (Bruker, Rheinstetten, Germany) version 3.6.2. The concentration of CS was expressed as μmol per mg of protein. Protein extracts were obtained by sonication as described by Pinto et al. (2015), and protein quantification was performed using the Bio-Rad Protein Assay. For cell dry weight (DW) determinations, 40 ml of culture at $OD_{730} \approx 1.0$ (or equivalent) was centrifuged at 3,857 *g* for 10 min at RT. Then, the cell pellet was dried at 60°C for 48 h. Experiments included, at least, three biological replicates.

In silico Analysis of CS Production

The genome-scale metabolic model of *Synechocystis*—*i*Syn811 (Montagud et al., 2011) –, was updated to include all the information required for the simulations. The manual curation process started with the addition of the final reaction in the synthesis of sucrose ("spp: H_2O+ sucrose 6-phosphate → phosphate + sucrose"), and also the metabolic precursors and the three reactions required for the synthesis of the heterologous CS, glycine betaine ("GSMT1: *S*-adenosyl-L-methionine + glycine ↔ *S*-adenosyl-L-homocysteine + sarcosine," "GSMT2: *S*-adenosyl-L-methionine + sarcosine ↔ *S*-adenosyl-L-homocysteine + *N*,*N*-dimethylglycine," and "DMT: *S*-adenosyl-L-methionine + *N*,*N*-dimethylglycine ↔ *S*-adenosyl-L-homocysteine + *N*,*N*,*N*-trimethylglycine"). In this process, the nomenclature was corrected and standardized (e.g., "glycerone" to "dihydroxy-acetone" or "GDP-mannose" to "GDP-D-mannose"), and the reversibility of some reactions changed (e.g., "sn-glycerol-3-phosphate → dihydroxy-acetone phosphate" to "sn-glycerol-3-phosphate ↔ dihydroxy-acetone phosphate"). Flux balance analysis (Orth et al., 2010) was performed to the *i*Syn811 genome-scale metabolic reconstruction of *Synechocystis* for the production assessment of four different CS: three heterologous (glycine betaine, ectoine, and mannosylglycerate), and three native (glucosylglycerol, glutamate and sucrose). The MATLAB software, COBRA Toolbox v3.0 (Heirendt et al., 2019) was used for quantitative prediction of cellular and multicellular biochemical networks with constraint-based modelling. Simulations were constrained to match an autotrophic specific growth rate of 0.09/h, which corresponds to a light input of 0.8 mE/gDW/h and to a net carbon flux of 3.4 mmol/gDW/h into the cell, with CO_2 as carbon source. The description of the *i*Syn811 model and further information on the simulation procedure are available in Montagud et al. (2010).

Statistical Analysis

The statistical analysis was performed by means of one- or two-way ANOVAs, using GraphPad Prism v6.01 (GraphPad Software Inc., United States).

RESULTS

Generation of *Synechocystis* Mutants Deficient in the Synthesis of Native Compatible Solutes

The sustainable production of heterologous compatible solutes using *Synechocystis* as a chassis was envisioned in this work. The starting step was the generation of mutants deficient in the production of one or both of the main native compatible solutes, sucrose, or glucosylglycerol (GG). For this purpose, the genes encoding the enzymes involved in the first step of sucrose or/and GG synthesis (*sps* and *ggpS*, respectively), were knockout by double homologous recombination generating the *Synechocystis* markerless mutants Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS* (for details see *Materials*). The complete segregation of the mutants was confirmed by PCR and Southern blot (Additional file 1: **Supplementary Figures S1, S2**).

Effect of NaCl on the Growth of *Synechocystis* Wild-Type and the CS Deficient Mutants

The growth of the CS deficient mutants under different salinities was analyzed. *Synechocystis* WT and mutants Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS* were grown in standard BG11 medium or in BG11 supplemented with 3, 5, and 7% (wt/vol) NaCl, corresponding to 510, 860, and 1,200 mM, respectively. The growth was monitored by measuring the OD_{730} and chlorophyll *a* (chl *a*) content (**Figure 1**). In the absence of NaCl, the three mutants exhibited growth similar to the WT

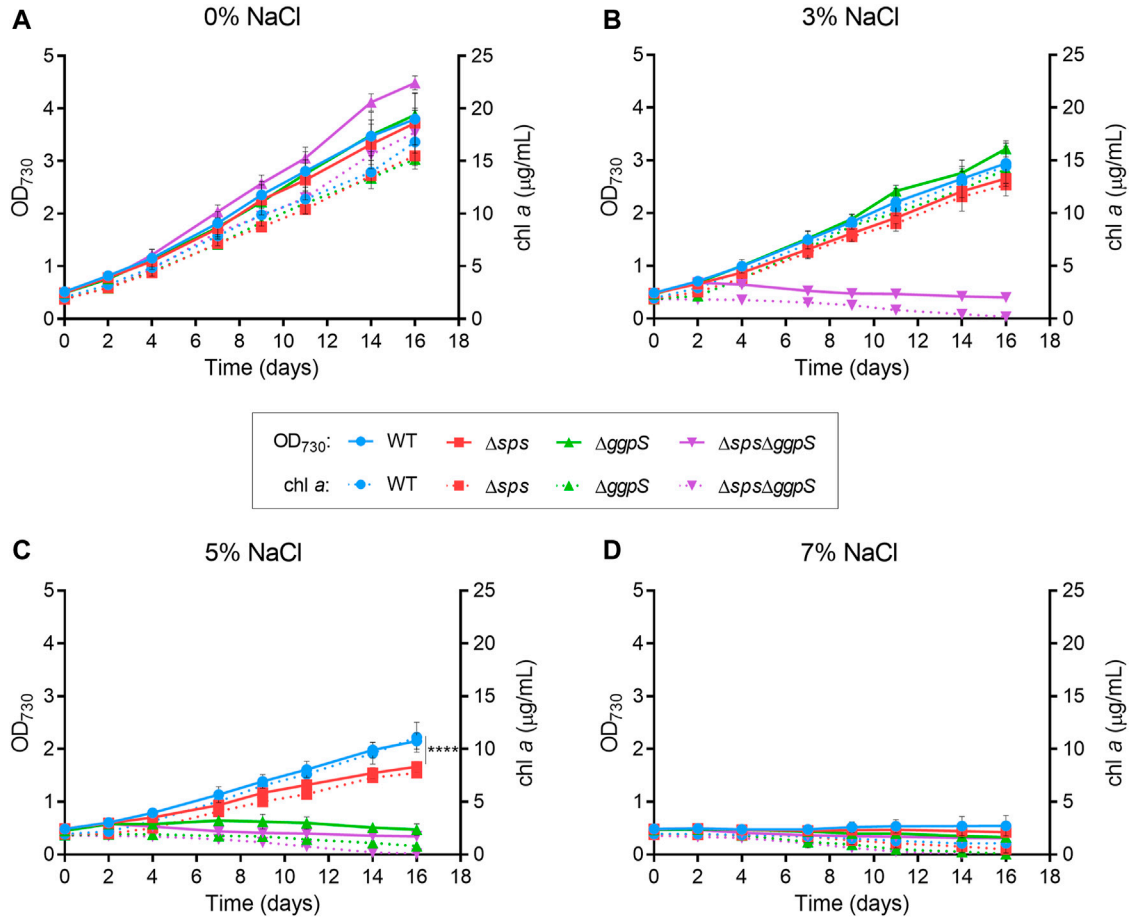


FIGURE 1 | Growth curves of *Synechocystis* wild-type (WT) and Δ*sps*, Δ*ggpS* and Δ*sps*Δ*ggpS* mutants in BG11 **(A)** or BG11 supplemented with 3% **(B)**, 5% **(C)** or 7% **(D)** (wt/vol) NaCl. Cultures were grown at 30°C with orbital shaking (150 rpm) under a 12 h light (25 µE/m^2/s)/12 h dark regimen. Growth was monitored by measuring optical density at 730 nm (OD_{730}), and chlorophyll *a* (chl *a*) (full and dotted lines, respectively). Error bars correspond to standard deviations from at least three biological replicates with technical duplicates. Statistical analysis was performed using a two-way ANOVA, and a significant difference in terms of OD_{730} is identified by **** ($p \leq 0.0001$).

(**Figure 1A**), indicating that the synthesis of sucrose and/or GG is nonessential under standard growth conditions. Nonetheless, challenging the cells with 3% NaCl had clear detrimental effects, with a ~23% growth decrease observed for the WT and CS single mutants Δ*sps* and Δ*ggpS* (**Figure 1B**; Additional file 1: **Supplementary Table S1**). The inactivation of both pathways in the Δ*sps*Δ*ggpS* mutant led to total growth arrest accompanied by a decline in chl *a* content (**Figure 1B**; purple lines and Additional file 1: **Supplementary Table S1**). A more pronounced impact was observed by increasing NaCl to 5%. The Δ*ggpS* could not grow in these conditions, while for the WT and Δ*sps*, a severe growth impairment (~49%) was observed (**Figure 1C**; Additional file 1: **Supplementary Table S1**). The growth of the latter two strains was similar up to day 7 however, after this period, the growth of Δ*sps* slowed down and by day 16 there was a significant difference ($p \leq 0.0001$) compared with the WT. The chl *a* content confirmed these observations (**Figure 1C**; red and blue lines). Further increase in the NaCl concentration to 7% (wt/vol) showed that none of the strains tested could withstand the stress imposed (**Figure 1D**; Additional file 1: **Supplementary Table S1**).

Quantification of CS in *Synechocystis* Wild-Type and the CS Deficient Mutants

The CS content was quantified in *Synechocystis* WT, Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS* mutants grown in BG11 or BG11 supplemented with NaCl (**Figure 2**), under salinity conditions in which each strain could sustain growth (**Figure 1**). In the WT, CS accumulation increased significantly in a salinity-dependent manner and GG was accumulated in higher amounts followed by glutamate and sucrose (**Figure 2A**). As expected, the two main compatible solutes sucrose and GG could only be detected in the presence of NaCl. The amino acid glutamate was detected in all backgrounds and conditions, and increased more than 1.8-fold in the presence of salinity (significant difference $p \leq 0.01$ for WT in 0 and 5% NaCl). In the Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS* mutants the absence of sucrose and/or GG production was confirmed

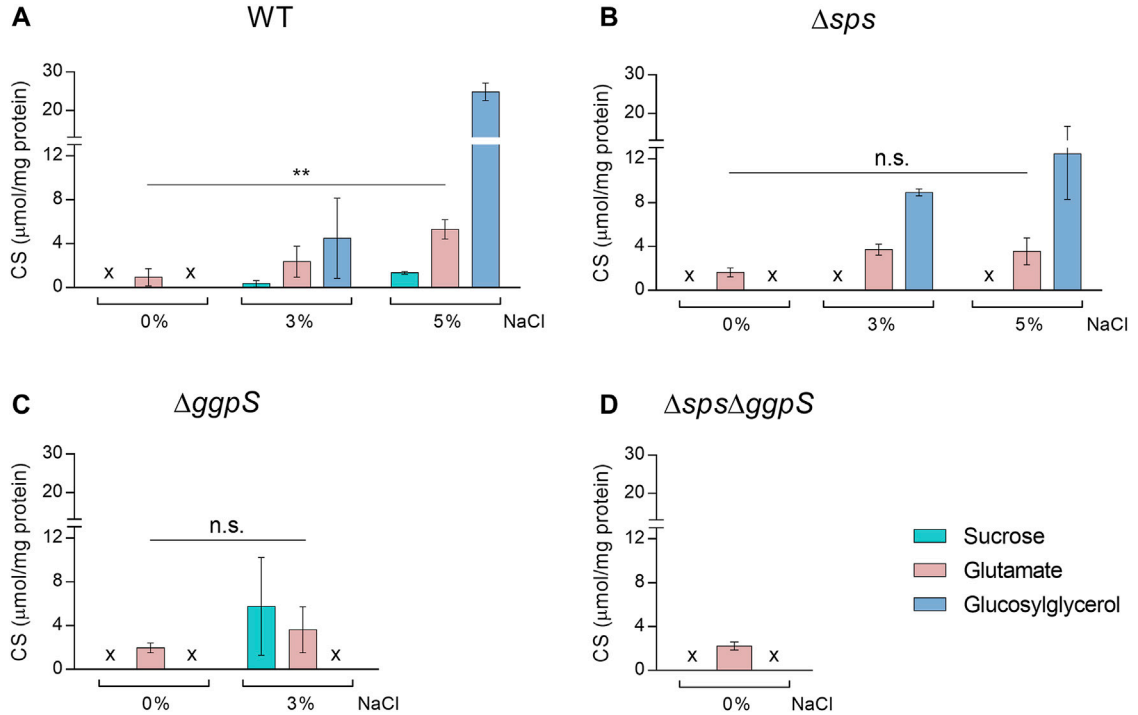


FIGURE 2 | Effect of NaCl on the synthesis of native compatible solutes sucrose, glutamate and glucosylglycerol by *Synechocystis* wild-type (WT) **(A)**, and the Δ*sps* **(B)**, Δ*ggpS* **(C),** and Δ*sps*Δ*ggpS* mutants **(D)**. Cultures were grown in BG11 or BG11 supplemented with 3% or 5% (wt/vol) NaCl, at 30°C with orbital shaking (150 rpm) under a 12 h light (25 μE/m^2/s)/12 h dark regimen, and cells were harvested 4 days after inoculation (initial OD_{730} ≈ 0.5). Compatible solutes were quantified by H-NMR and the results were normalized per mg of protein. x—not detected. Error bars correspond to standard deviations from at least three biological replicates. Statistical analysis was performed using two-way ANOVA. Statistically significant differences are identified: ** ($p \leq 0.01$) and n.s. (not significant).

(**Figures 2B–D**). For the Δ*sps*, a salinity-dependent accumulation of GG was also detected (**Figure 2B**), however, the GG concentration was 50% lower compared with the WT, under 5% NaCl. In contrast, the Δ*ggpS* mutant accumulated 17-fold more sucrose than the WT, under 3% NaCl (**Figure 2C**). All the proton NMR spectra are depicted in Additional file 1: **Supplementary Figure S3**.

Effect of NaCl on Total Carbohydrates, Glycogen, Capsular, and Released Polysaccharides

In addition to the CS pools, the total carbohydrate content was analyzed in *Synechocystis* WT and the CS deficient mutants (**Figure 3A**). Generally, the presence of salinity (3 or 5% NaCl) had no significant impact on the production of total carbohydrates, except for the Δ*sps* mutant that showed some fluctuation when exposed to different salinities (**Figure 3A**; $p \leq 0.05$). To further clarify the carbon distribution in response to salinity, the amount of glycogen as well as the production of extracellular polymeric substances, CPS and RPS, were also determined (**Figures 3B–D**). The presence of NaCl led to a significant decrease in the amount of glycogen in the WT, Δ*sps*, and Δ*ggpS*, with reductions of more than 56%, independent of the salinity concentration and the deletion of one of the CS pathways (**Figure 3B**). In terms of CPS, the opposite effect was observed with a 2.4-fold increase in CPS for the WT and the Δ*sps* at 5%, and a 2.1-fold increase for Δ*ggpS* at 3% NaCl, compared with 0% NaCl (**Figure 3C**). The amount of RPS produced by WT or Δ*ggpS* did not change significantly whereas for the Δ*sps* a 2.5-fold increase was registered under 5% NaCl (**Figure 3D**). Staining the WT and Δ*sps* cultures with Alcian Blue confirmed similar RPS production for the WT under 0 and 5% NaCl, while for the Δ*sps* the accumulation of RPS in 5% NaCl is evident, leading to the formation of cell aggregates (**Figure 3E**; black arrowhead).

In silico Prediction of Production Rates for Native and Heterologous CS Using *Synechocystis* Wild-Type

The genome-scale metabolic model of *Synechocystis*—*i*Syn811 (Montagud et al., 2011)—was updated to include all the information required for calculating CS production rates. The manual curation process started with the addition of the metabolic precursors and the reactions required for CS synthesis. The nomenclature was also corrected and standardized, and the reversibility of some reactions was changed (for more details see the *Materials* section). After the curation of the metabolic model was completed, the COBRA ("The COnstraint-Based Reconstruction and Analysis") Toolbox v3.0 (Heirendt et al., 2019), was used to simulate the compatible solute production rate as a function of *Synechocystis* wild-type growth under autotrophic conditions (**Figure 4**). The results show a linear tradeoff between the cell's

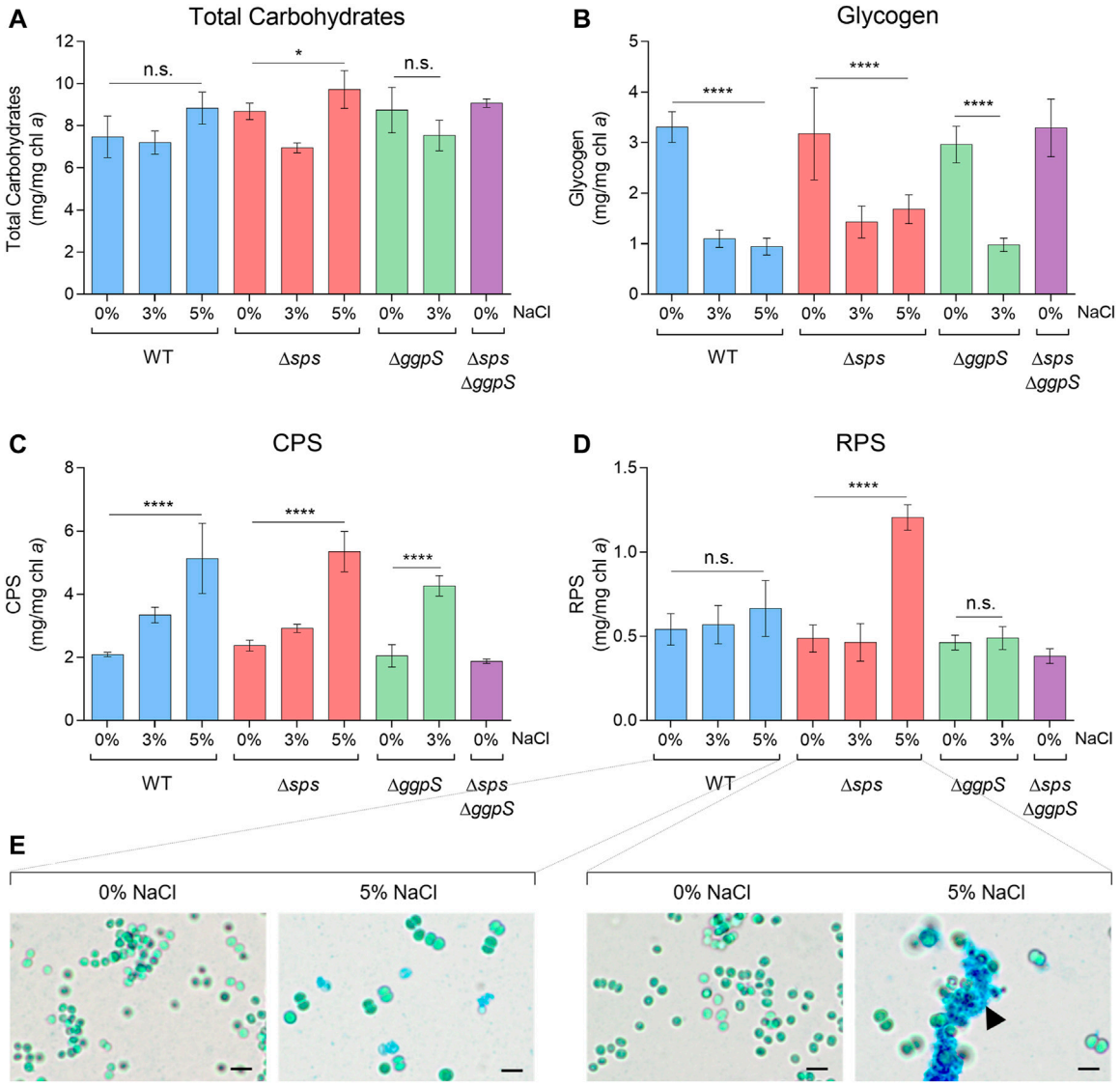


FIGURE 3 | Effect of NaCl on total carbohydrates **(A)**, glycogen **(B)**, CPS—capsular polysaccharides **(C)**, and RPS—released polysaccharides **(D)** produced by *Synechocystis* wild-type (WT) and the Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS* mutants. Cells were grown in BG11 or BG11 supplemented with 3% or 5% (wt/vol) NaCl, at 30°C with orbital shaking (150 rpm) under a 12 h light (25 μE/m^2/s)/12 h dark regimen for 16 days. Results are expressed as milligrams per milligram of chlorophyll *a* (chl *a*). Error bars correspond to standard deviations from at least three biological replicates, with technical triplicates. Statistical analysis was performed using one-way ANOVA. Statistically significant differences are identified: **** ($p \leq 0.0001$), * ($p \leq 0.05$), and n.s. (not significant). Light micrographs **(E)** of *Synechocystis* WT and Δ*sps* cultures grown in BG11 or BG11 supplemented with 5% NaCl and stained with Alcian Blue; the black arrowhead highlights RPS production by the Δ*sps* mutant under 5% NaCl. Scale bars: 5 μm.

resources toward growth or the production of the different CS. Regarding the production of the native CS, GG, and sucrose impose a higher metabolic burden showing lower *in silico* production rates (0.378 and 0.283 mmol/gDW/h, respectively), compared with glutamate (0.567 mmol/gDW/h). The simulation of the production of heterologous CS glycine betaine (GB) showed the best compromise between growth and production compared with the three native CS, with the highest predicted maximum production rate of 0.680 mmol/gDW/h. In addition to GB, the production rates of other heterologous CS, such as ectoine and mannosylglycerate were also simulated, revealing that the maximum production rate predicted for these two solutes is lower than the obtained for GB (Additional file 1: **Supplementary Figure S4**). Hence, GB was chosen as the heterologous CS to be produced using the *Synechocystis* chassis developed (CS-deficient mutants).

Design and Assembly of the Synthetic Device for the Production of Glycine Betaine

Envisaging the heterologous production of GB, a synthetic device was designed based on the metabolic pathway described for the

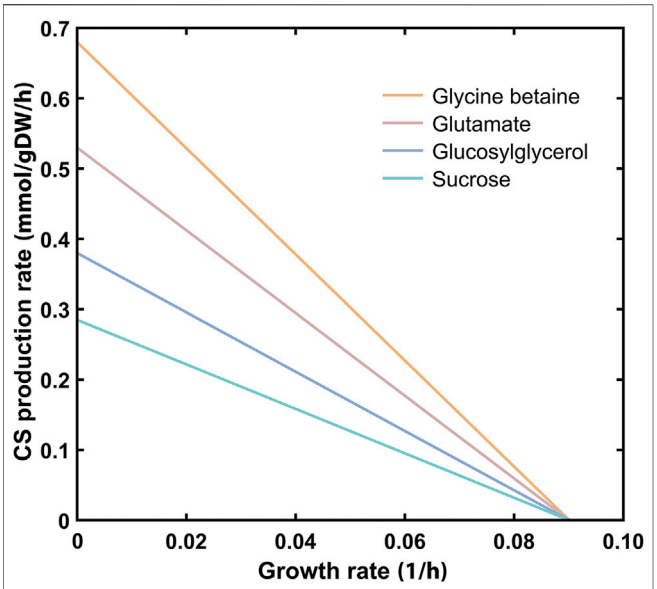


FIGURE 4 | Theoretical productivity of the heterologous compatible solute glycine betaine and the native ones (glutamate, glucosylglycerol, and sucrose), as predicted by the updated version of the genome-scale metabolic model *i*Syn811. The lines represent the compatible solute production rate as a function of *Synechocystis* wild-type growth under autotrophic conditions.

halophilic cyanobacterium *Aphanothece halophytica* (**Figure 5A**) (Nyyssola et al., 2000; Waditee et al., 2003). This device comprises two Open Reading Frames (ORFs) encoding the enzymes involved in the three-step methylation of glycine to glycine betaine: glycine-sarcosine-*N*-methyltransferase (*gsmt*), and the dimethylglycine-*N*-methyltransferase (*dmt*). In these reactions, *S*-adenosylmethionine (SAM) is the source of methyl groups for the synthesis of GB, and it can be synthesized from L-methionine by the *S*-adenosyl-methionine synthase (MetX) (**Figure 5A**). To prevent the shortage of the SAM precursor, the ORF encoding *Synechocystis'* native MetX (*metX*, *sll0927*) was also included in the device. The sequences of the three ORFs (*gsmt*, *dmt*, and *metX*) were codon-optimized and restriction sites incompatible with the BioBrick standard RFC [10] were eliminated. Subsequently, the ribosome binding site (RBS) BBa_B0030 and double stop codons (TAATAA) were included before and after each ORF, respectively. A double terminator (BBa_B0015) was also included at the end of the synthetic construction (Ahbet). Additionally, the designed DNA sequence was flanked by the prefix and suffix of the BioBrick RFC [10] standard (Canton et al., 2008), enabling the use of the standard assembly method to include the regulatory element (promoter). After DNA synthesis, the Ahbet construction was cloned downstream of the promoter $P_{trc.x.lacO}$, originating the $P_{trc.x.lacO}$::Ahbet synthetic device (hereafter GB device) (**Figure 5B**). The $P_{trc.x.lacO}$ is a constitutive promoter in *Synechocystis*, previously characterized by our group and is 41-fold stronger than the reference cyanobacterial promoter P_{rnpB} (Ferreira et al., 2018).

Effect of the Implementation of the GB Device Into the *Synechocystis* Chassis

The GB synthetic device was implemented into the *Synechocystis* WT and the CS deficient Δ*ggpS* and Δ*sps*Δ*ggpS* chassis described above, using the replicative plasmid pSEVA351. The device was not introduced into

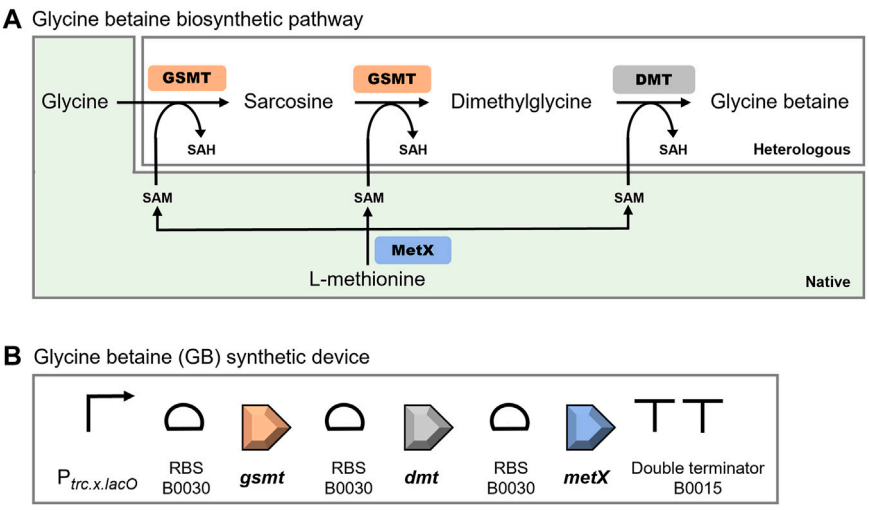


FIGURE 5 | Glycine betaine biosynthetic pathway **(A)** and schematic representation of the glycine betaine (GB) synthetic device for the production of this heterologous compatible solute in *Synechocystis* **(B)**. The glycine betaine biosynthetic pathway comprises the GSMT (glycine-sarcosine-*N*-methyltransferase) and DMT (dimethylglycine-*N*-methyltransferase) from *Aphanothece halophytica* (heterologous), and MetX (*S*-adenosyl-methionine synthase) from *Synechocystis* (native). SAM—*S*-adenosyl-methionine; SAH—*S*-adenosyl-homocysteine. The GB device includes the promoter $P_{trc.x.lacO}$, the ribosomal binding site (RBS) B0030, the open reading frames *gsmt*, *dmt*, and *metX* codon optimized for *Synechocystis*, and the double terminator B0015.

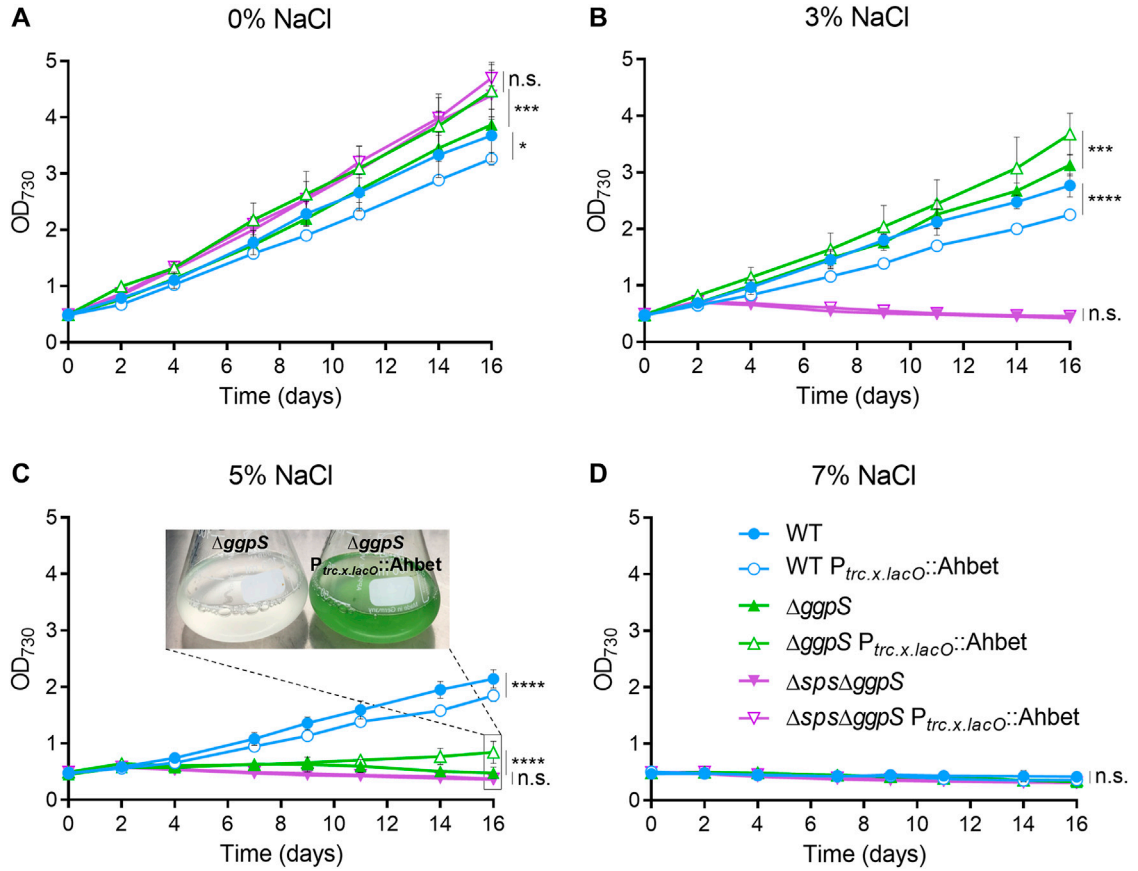


FIGURE 6 | Growth of *Synechocystis* wild-type (WT), Δ*ggpS*, and Δ*sps*Δ*ggpS* without or with the GB device ($P_{trc.x.lacO}$::Ahbet). Cultures were grown in BG11 **(A)** or BG11 supplemented with 3% **(B)**, 5% **(C)**, or 7% **(D)** (wt/vol) NaCl, at 30°C with orbital shaking (150 rpm) under a 12 h light (25 μE/m^2/s)/12 h dark regimen. Growth was monitored by measuring the optical density at 730 nm (OD_{730}). Error bars correspond to standard deviations from at least three biological replicates with technical duplicates. Statistical analysis was performed using two-way ANOVA. Statistically significant differences are identified: **** ($p \leq 0.0001$), *** ($p \leq 0.001$), * ($p \leq 0.05$), and n.s. (not significant). In **Figure 6C**, the liquid cultures of *Synechocystis* Δ*ggpS* **(left)** and Δ*ggpS* $P_{trc.x.lacO}$::Ahbet **(right)**, in BG11 supplemented with 5% NaCl after 16 days of cultivation, are shown.

Δ*sps* since the characterization showed that this mutant is similar to the WT in terms of growth, total carbohydrates, glycogen content and CPS (**Figure 1**, **Figures 3A–C**, respectively). The presence of the GB device in the cells was confirmed by PCR (Additional file 1: **Supplementary Figure S5**), and the growth and chl *a* content of the transformants were monitored in absence/presence of salinity and compared with the respective backgrounds (**Figure 6**; Additional file 1: **Supplementary Figure S6**). As shown in **Figure 6**, the introduction of the synthetic device had distinct effects depending on the genetic background. The implementation of GB device into Δ*ggpS* led to a significant improvement of growth (16%) in BG11 and BG11 supplemented with 3% NaCl (**Figures 6A,B**; green lines), and supported its survival under 5% NaCl (**Figure 6C**; green lines). After 16 days of cultivation under 5% NaCl, the batch culture of Δ*ggpS* showed clear signs of chlorosis/necrosis while the culture of the Δ*ggpS* mutant harboring the GB device remained green (**Figure 6C**; insert). In agreement, the chl *a* content was 0.8 μg/ml and 1.8 μg/ml, respectively (Additional file 1: **Supplementary Figure S6**; green lines). Notably, this survival phenotype was observed for at least 25 days (data not shown). In the WT background the presence of the device had a detrimental effect on growth (~15% decrease) in all conditions tested (**Figure 6**; blue lines). The growth of the double mutant Δ*sps*Δ*ggpS* in the absence of NaCl was not affected by the introduction of the synthetic device (**Figure 6A**; purple lines). Moreover, the mutant harboring the GB device was unable to survive under saline conditions, similarly to what happened to Δ*sps*Δ*ggpS* background (**Figures 6B–D**; purple lines).

Analysis of Transcript Levels in *Synechocystis* Strains Harboring the GB Device

The next step in the characterization of the different *Synechocystis* strains harboring the GB device was the

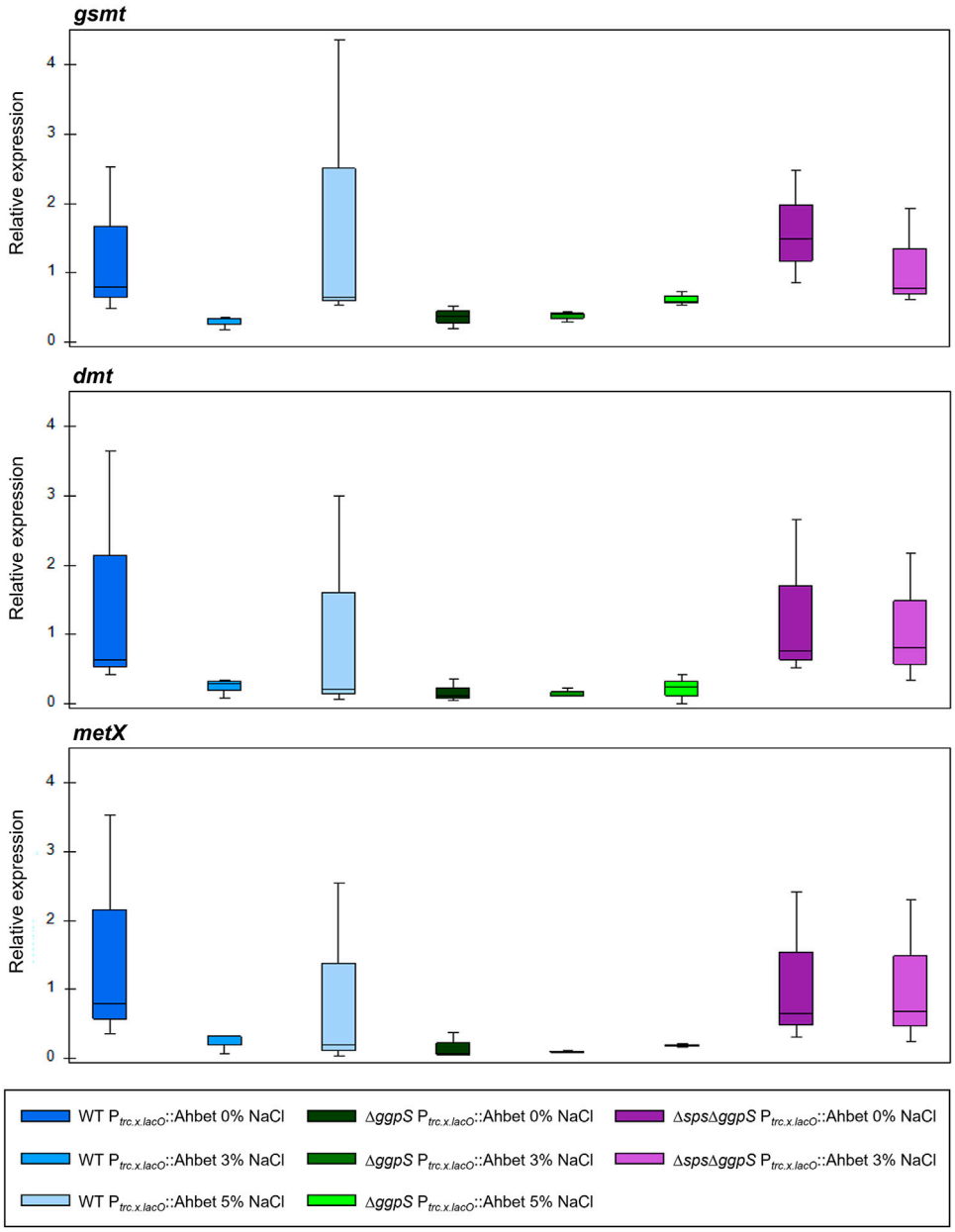


FIGURE 7 | RT-qPCR analysis of *gsmt*, *dmt*, and *metX* relative expression in *Synechocystis* strains (WT, Δ*ggpS*, and Δ*sps*Δ*ggpS*) harboring the GB device. RNA was extracted from cells grown in BG11 or BG11 supplemented with 3 and 5% (wt/vol) NaCl, at 30°C with orbital shaking (150 rpm) under a 12 h light (25 μE/m^2/s)/12 h dark regimen. The box-whisker plots represent the expression of the target genes relative to WT $P_{trc.x.LacO}$::Ahbet in absence of salt (0% NaCl). Data were obtained from three biological replicates and three technical replicates, and the whiskers represent the minimum and maximum non-outlier values in the dataset. One-way ANOVA was performed no significant differences could be detected.

evaluation of the transcript levels of the three ORFs comprised in the device (*gsmt*, *dmt*, and *metX*) by RT-qPCR (for more details see the *Materials* section). As shown in **Figure 7**, the transcripts of the three genes (*gsmt*, *dmt*, and *metX*) were detected in all samples and the relative expression was reasonably stable independently of the background strain. Additionally, the relative expression remained similar under salinity conditions and, even though some variation could be detected, it was not statistically significant (Additional file 1: **Supplementary Table S2**).

Quantification of Native and Heterologous CS in *Synechocystis* Chassis Harboring the GB Device

The CS pool of the different *Synechocystis* strains (WT, Δ*ggpS*, and Δ*sps*Δ*ggpS*) harboring the GB device was analyzed in

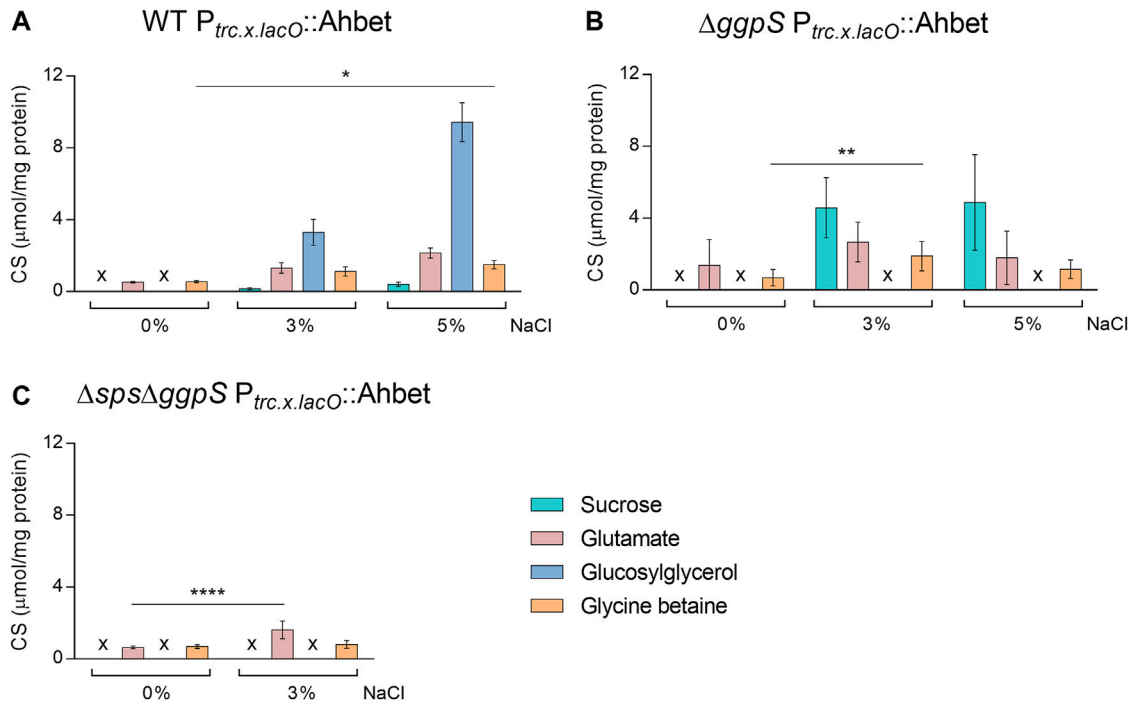


FIGURE 8 | Production of glycine betaine (GB) and native compatible solutes in different *Synechocystis* strains harboring the GB device: WT (wild-type) $P_{trc.x.lacO}$::Ahbet **(A)**, Δ*ggpS* $P_{trc.x.lacO}$::Ahbet **(B)**, and Δ*sps*Δ*ggpS* $P_{trc.x.lacO}$::Ahbet **(C)**. Cultures were grown in BG11 or BG11 supplemented with 3% or 5% (wt/vol) NaCl, at 30°C with orbital shaking (150 rpm) under a 12 h light (25 μE/m^2/s)/12 h dark regimen; and cells were harvested 4 days after inoculation (initial OD_{730} ≈ 0.5). Compatible solutes were quantified by H-NMR and the results were normalized per mg of protein. x—not detected. Error bars correspond to standard deviations from three biological replicates. Statistical analysis was performed using two-way ANOVA. Statistically significant differences are identified: **** ($p \leq 0.0001$), ** ($p \leq 0.01$), and * ($p \leq 0.05$).

absence/presence of NaCl after 4 days of cultivation (**Figure 8**). The results obtained confirmed that the implementation of the pathway for the synthesis of heterologous CS was successful, since glycine betaine could be detected in all strains and conditions analyzed. Under 0 and 3% NaCl, the presence of the GB device in the WT background did not significantly influence the synthesis of native CS and heterologous production of GB is not significantly influenced by salinity. However, under 5% NaCl, there was an impact on the synthesis of glutamate and GG that decreased by 59 and 62%, respectively, and the synthesis of GB increased 2.7-fold compared with 0% NaCl (**Figures 2A**, **8A**). Similarly, the implementation of the device in the Δ*ggpS* did not affect the production of the native CS under 0 and 3% NaCl compared with Δ*ggpS* chassis (**Figures 2C**, **8B**). However, the introduction of the device into the Δ*ggpS* background allowed this strain to survive under 5% NaCl and, therefore the quantification of CS was also performed. The results obtained showed that in Δ*ggpS* $P_{trc.x.lacO}$::Ahbet, besides glycine betaine production, the levels of sucrose and glutamate were similar to the ones observed for 3% NaCl (**Figure 8B**). An analysis of the Δ*ggpS* harboring the GB device after 16 days of cultivation suggested that the production of all CS is maintained for at least 2 weeks of cultivation (Additional file 1: **Supplementary Figure S7**). For the double mutant, the presence of the device led to a significant decrease (71%) of glutamate in the absence of salinity (**Figures 2D**, **8C**). In contrast, under 3% NaCl, there was a 2.5-fold increase in glutamate ($p \leq 0.0001$) compared with 0% NaCl, while the glycine betaine content remained similar (**Figure 8C**). All the proton NMR spectra are depicted in Additional file 1: **Supplementary Figure S8**. Furthermore, the CS quantification was also performed for the extracellular medium, and the results showed that none of the native CS could be detected, while GB was detected in negligible amounts in all strains harboring the device and conditions tested (Additional file 1: **Supplementary Table S3**).

DISCUSSION

The sustainable production of compatible solutes (CS) is essential for pharmaceutical and cosmetic industries. The current microbiological processes have a significant negative impact on the environment, which could be mitigated by the use of photoautotrophic chassis such as cyanobacteria. For the synthesis of heterologous CS in *Synechocystis*, the construction of customized chassis is required and our strategy was to eliminate competing or redundant pathways. Therefore, in this work we have generated three *Synechocystis* mutants deficient in the production of native compatible solutes (namely, sucrose, or/and glucosylglycerol). These strains—Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS*—were characterized under different salinity concentrations, expanding the knowledge that will allow

further optimization of the chassis for the increased production of heterologous CS, such as glycine betaine (GB). In this context, an updated version of the genome-scale metabolic model of *Synechocystis*—*i*Syn811 (Montagud et al., 2011)—was used to predict the production rates for native and heterologous CS using *Synechocystis* wild-type. The simulations show a linear tradeoff between deviating resources toward cellular growth or toward the production of the solutes. Since energy and carbon uptake are limited, any extra need of ATP or carbon molecules for compatible solute synthesis will impair cell growth. Whether carbon or light uptake is limiting the synthesis of each CS is difficult to predict, since alternative routes with different energetic efficiencies can be simultaneously active under different growth conditions. From the CS evaluated, the predictions indicate that the synthesis of native sucrose, and glucosylglycerol (GG) has a higher impact on cell growth than glutamate or the heterologous solute glycine betaine (GB) (**Figure 4**). The production of sucrose and GG require glucose that drains more cell resources than the reported for the synthesis of an amino acid (Kaleta et al., 2013), such as glutamate or glycine (the latter required for GB production). The results also suggest that the production of GB has a smaller restraining effect on growth than glutamate or other heterologous CS, like ectoine and mannosylglycerate (Additional file 1: **Supplementary Figure S4**).

In parallel, the evaluation of the CS levels of the wild-type and the mutants (Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS*) confirmed the salt-induced accumulation of sucrose and GG, which is well documented in the literature [for reviews see e.g., Klähn and Hagemann (2011); Hagemann (2013); Kirsch et al. (2019)]. In contrast, glutamate could be detected in the absence and presence of NaCl (**Figure 2**). These results are in agreement with the reported accumulation of this amino acid in *Synechocystis* grown in artificial seawater medium (ASW; 340 mM NaCl), and in BG11 supplemented with 12 mM KCl (Iijima et al., 2015; Iijima et al., 2020).

Previous works have also reported a tradeoff between the pools of different compatible solutes and other carbon sinks, such as glycogen or extracellular polymeric substances (EPS) in cyanobacteria (Du et al., 2013; Baran et al., 2017; Kirsch et al., 2017). In our work, the total carbohydrate content of *Synechocystis* WT and CS-deficient mutants remained unchanged when cells were exposed to NaCl, whereas a significant decrease in glycogen was observed (**Figure 3**). Concomitantly, the accumulation of capsular polysaccharides (CPS) was observed in a salinity-dependent manner and for the strains tested. In line with these observations is the increase in the levels of proteins involved in glycogen degradation, reported when *Synechocystis* cells were grown in ASW medium (Iijima et al., 2015). The protective role of EPS against salt stress was also demonstrated in a *Synechocystis* Δ*sll1581*Δ*slr1875* double mutant, showing that a decrease in CPS content increases NaCl sensitivity (Jittawuttipoka et al., 2013). Altogether, these results strongly suggest that under saline conditions, *Synechocystis* breaks down glycogen and redirects carbon fluxes toward the production of CS and extracellular polysaccharides, promoting cell homeostasis and contributing to cell protection.

From the three *Synechocystis* CS-deficient mutants generated in this work, the Δ*sps* was the only one able to grow in 5% NaCl. We also observed that this mutant's growth gets impaired over time, suggesting that the presence of sucrose is of additional importance for long-term cultivation. Accordingly, a previous work showed that Δ*sps* cells in stationary phase were unable to survive a salt shock, which was not observed for cells in exponential phase; this effect could be prevented by sucrose supplementation (Desplats et al., 2005). Moreover, we show that the absence of sucrose leads to a severe reduction in the accumulation of GG whereas the released polysaccharides (RPS) increase significantly (1.8-fold), implying that extracellular polysaccharides production is crucial for the survival of the Δ*sps* mutant under 5% NaCl. These results also suggest that sucrose role might go beyond osmoprotection, being involved in the regulation of metabolic pathways, e.g., triggering signaling cascades, as it was previously hypothesized by Desplats et al. (2005). For the Δ*ggpS* mutant an increased sucrose level was detected under 3% NaCl, showing that this sugar can sustain *Synechocystis*' survival under sea salt conditions for at least 16 days. Previously (Miao et al., 2003), generated a *Synechocystis* Δ*agp* mutant unable to synthesize ADP-glucose (a precursor required for GG synthesis) that was also shown to accumulate high levels of sucrose and could survive upon a salt shock of 900 mM (5.2% NaCl). Notably, in the latter work and here, the mutant's sucrose levels were similar to GG accumulated in the WT cultivated under the same conditions. Taken together, these studies imply that GG and sucrose can have comparable osmoprotectant capacity when accumulated in similar levels. Additionally, the Δ*sps*Δ*ggpS* mutant was unable to survive in any salt concentration tested and glutamate was the only CS that could be detected. Thus, this amino acid seems to have a minor contribution to the salt acclimation process in *Synechocystis*, similar to what was reported for the halophilic bacterium *Salinibacter ruber* (Oren et al., 2002).

Considering the metabolic model simulation, a synthetic device for the production of glycine betaine (GB) was designed and implemented into the *Synechocystis* wild-type and our customized chassis (CS-deficient mutants). Besides the ORFs required for GB production (*gsmt* and *dmt*) and *metX* (to prevent SAM shortage), this device comprises well-characterized regulatory elements: the synthetic promoter $P_{trc.x.lacO}$ (Ferreira et al., 2018), the RBS BBa_B0030, and the double terminator BBa_B0015. This design ensured the stable constitutive transcription observed for the GB device ORFs, regardless of the chassis or salinity conditions (**Figure 7**), reinforcing that the use of orthogonal regulatory components is crucial to ensure the proper insulation of synthetic devices from the regulatory network of the chassis (Costello and Badran, 2021). Unlike transcription, the synthesis of the solute was not independent of cultivation conditions, and GB levels increased with salinity. In agreement, higher levels of glycine have been reported for *Synechocystis* cells grown in ASW medium compared with those grown in BG11 (Iijima et al., 2015). Since this amino acid is a precursor of GB, the high levels of glycine under salinity conditions most probably favor the synthesis of GB. In addition, glycogen degradation and carbon fluxes redirection

TABLE 2 | Native and heterologous production of glycine betaine *via* the three-step glycine methylation pathway.

Native production				
Strain	**Salinity (mM)**	**Production capacity**	**Cultivation time**	**Reference**
Aphanothece halophytica	2,000	0.06 μmol/mg protein	1 h	Ishitani et al. (1993)
Aphanothece halophytica	2,000	~0.4 μmol/10^7 cells	7 days	Incharoensakdi and Waditee (2000)
Aphanothece halophytica	1,500	~40,000 μmol/gFW	10 days	Waditee et al. (2007)
Aphanothece halophytica	2,000	20.1 μmol/gFW	15 days	Waditee-Sirisattha et al. (2015)

Heterologous production					
Production strain	**Native strain**	**Salinity (mM)**	**Production capacity**	**Cultivation time**	**Reference**
Arabidopsis thaliana	*Aphanothece halophytica*	100	~2 μmol/gFW	15 days	Waditee et al. (2005)
Nicotiana tabacum	*Aphanothece halophytica*	0	0.4 μmol/gFW	28 days	He et al. (2011)
Synechococcus PCC 7942	*Aphanothece halophytica*	500	~1.5 μmol/gFW	NA	Waditee et al. (2005)
Anabaena sp. PCC 7120	*Aphanothece halophytica*	140	0.04 μmol/gFW	7 days	Waditee-Sirisattha et al. (2012)
Anabaena doliolum	*Aphanothece halophytica*	500	12.92 μmol/gDW	10 days	Singh et al. (2013)
Synechocystis sp. PCC 6803	*Aphanothece halophytica*	510	64.29 μmol/gDW	4 days	This work
Escherichia coli XL1-Blue	*Ectothiorhodospira halochloris*	300	78 μmol/gDW	NA	Nyyssola et al. (2000)
Escherichia coli BL21	*Aphanothece halophytica*	300	~23 μmol/gDW	3 h	Waditee et al. (2003)
Escherichia coli BL21	*Aphanothece halophytica*	300	~2,000 μmol/mg protein	3 h	Waditee et al. (2007)
Escherichia coli BL21	*Aphanothece halophytica*	300	~80 μmol/L	2 h	He et al. (2011)
Escherichia coli DH5α	*Aphanothece halophytica*	500	6 μmol/gDW	24 h	Waditee-Sirisattha et al. (2012)
Escherichia coli DH5α	*Aphanothece halophytica*	0	80.62 μmol/gDW	ON	Singh et al. (2013)
Pseudomonas denitrificans	*Aphanothece halophytica*	0	NA*	1 day	Shkryl et al. (2020)

*ON, overnight; NA, not available; DW, dry weight; FW, fresh weight; *—Identified by HPLC-MS.*

toward the production of CS could also explain the increased amount of GB produced in the presence of NaCl.

The implementation of the GB device into *Synechocystis* wild-type led to a small decrease in growth in all conditions tested (**Figure 6**). As predicted by the metabolic flux model, the device may drain the cell's resources imposing a metabolic burden, causing growth impairment. This can be explained by the redirection of part of the photosynthetically fixed carbon to the synthesis of CS, which is no longer available for biomass formation, similarly to what was reported for the production of mannitol (Wu et al., 2020). In contrast to what was observed for the WT, the introduction of the GB device into the Δ*ggpS* mutant resulted in an increased salt tolerance with the concomitant growth improvement, enabling its survival under 5% NaCl. This phenotype was maintained under long-term cultivation periods up to 25 days (data not shown), suggesting that GB can compensate for the absence of GG. Conversely, the implementation of the GB device in the Δ*sps*Δ*ggpS* mutant did not improve its performance under salinity conditions. However, it remains unclear if this outcome is due to: 1) insufficient production of glycine betaine to allow cell survival or 2) the absence of both native compatible solutes (sucrose and GG).

In terms of production, the highest GB amount was obtained for the Δ*ggpS* cultivated in BG11 supplemented with 3% NaCl for 4 days (1.89 μmol GB/mg protein, corresponding to 64.29 μmol/gDW, and a volumetric productivity of 13.67 μg/L/h) (**Figure 8B**). Unexpectedly, the production of GB was not higher at 5% NaCl, which may be due to the limited capacity of the cells to survive in such conditions. Extending the cultivation period up to 16 days does not seem to affect the synthesis of GB in any condition tested (**Supplementary Figure S7**), suggesting that the process is stable. In addition, the negligible GB amounts detected in the extracellular medium show that this CS can be exported. Since product secretion facilitates recovery and reduces costs, this aspect should be addressed in the establishment of a GB-cell factory.

The production of GB using native organisms and heterologous hosts (with the synthesis of the solute mainly based on the metabolic pathway described for *A. halophytica*), has been previously reported (**Table 2**). However, a direct comparison is difficult since different normalization methods were used, and the cultivation conditions/time periods need also to be taken into consideration. Generally, the use of native GB producers such as the hypersaline cyanobacterium *A. halophytica* can render high amounts of the solute. This entails major disadvantages related to the high salt concentrations required, such as the reduced durability of the bioreactors, long processes, and detrimental impact on the environment. In contrast, with heterologous hosts the salinity concentrations used are at least 1/3 of those employed for *A. halophytica* (**Table 2**—heterologous production). Considering the photoautotrophic organisms, the amounts obtained using plants are low and require rather long cultivation periods. The most promising results were obtained using the filamentous cyanobacterium *Anabaena doliolum* that produced 12.92 μmol GB/gDW after 10 days of cultivation (Singh et al., 2013). Using our *Synechocystis* Δ*ggpS* chassis, we report a production level ~5-fold higher than that of *A. doliolum* in just 4 days of cultivation (64.29 μmol GB/gDW). Regarding the heterotrophic chassis, the GB amounts obtained using different *Escherichia coli* strains were only up to 1.25-fold higher than with *Synechocystis* Δ*ggpS*. Cultivation times are significantly reduced for heterotrophic bacteria, but the use of photoautotrophic

chassis enables CO_2 fixation promoting bio-mitigation and surpassing the need to supply a carbon source. Additionally, the highest GB production by our *Synechocystis* Δ*ggpS* chassis was achieved under 510 mM NaCl, opening up the possibility of large-scale cultivation with seawater (salinity range 3.1–3.8%). This does not seem as viable with *E. coli* since increasing the salt concentration to 500 mM has a substantial detrimental impact on the GB production (Waditee-Sirisattha et al., 2012).

CONCLUSION

The heterologous production of the compatible solute glycine betaine (GB) was successfully achieved in different *Synechocystis*-based chassis. The characterization of these compatible solutes (CS) deficient chassis (Δ*sps*, Δ*ggpS*, and Δ*sps*Δ*ggpS*) revealed that under saline conditions, the carbon fluxes are redirected from the synthesis of glycogen toward the production of CS and extracellular polysaccharides. In fact, the maximum amount of GB was obtained in Δ*ggpS* harboring the GB device, under 3% NaCl (64.29 μmol/gDW). This production level is promising and not far from applications using *E. coli*. Considering that the knowledge generated by the characterization of the CS deficient mutants will allow the identification of potential targets to optimize our chassis, that our GB production is based on sunlight and CO_2 fixation, and that there is the possibility of using seawater, *Synechocystis* emerges as a feasible photoautotrophic chassis for large-scale heterologous production of GB or other CS.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

CCP, EAF, FP, and PT were involved in the conceptual design of the work. The experimental design and strain engineering were carried out by EAF, CCP, JSR, and FP. Characterization and data analysis were performed by EAF and CCP. CS quantification by proton NMR was carried out by PL. EAF, DF, and JU were involved in the curation, simulations, and analyses involving the *i*Syn metabolic model. Data interpretation and article preparation was performed by EAF, CCP, DF (*i*Syn metabolic model), and PT. All authors have revised and approved the submitted version of the article.

FUNDING

This work was financed by Portuguese funds through the Fundação para a Ciência e a Tecnologia (FCT)/Ministério da Ciência, Tecnologia e Ensino Superior within the scope of UIDB/04293/2020, and UIDP/04293/2020. We also greatly acknowledge FCT for the scholarship SFRH/BD/117508/2016 (EAF) and the Assistant Researcher contracts CEECIND/00259/2017 (CCP) and 2020.01953.CEECIND (FP). The NMR data were acquired at CERMAX, ITQB-NOVA, Oeiras, Portugal with equipment funded by FCT, project AAC 01/SAICT/2016.

ACKNOWLEDGMENTS

The authors acknowledge the technical and scientific support of the i3S Scientific Platform "Cell Culture and Genotyping" in the RT-qPCR experiments. We are grateful to: Prof. Paula Gomes, Dr. Luísa Aguiar, and Mélanie Fonte (LAQV-REQUIMTE, Faculdade de Ciências, Universidade do Porto) for providing the equipment and technical assistance regarding compatible solutes extraction; Dr. Paulo Oliveira (i3S) for all the support provided, the scientific discussions, and suggestions; and Dr. Maria Siurana (UPV) for the scientific discussions regarding the *i*Syn metabolic model. The content of this article is available as a preprint at: https://doi.org/10.21203/rs.3.rs-1031622/v1 (Ferreira et al., 2021).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.821075/full#supplementary-material

REFERENCES

Ananya, A. K., and Ahmad, I. Z. (2014). Cyanobacteria" the Blue green Algae" and its Novel Applications: A Brief Review. *IJIAS* 7, 251–261.

Baran, R., Lau, R., Bowen, B. P., Diamond, S., Jose, N., Garcia-Pichel, F., et al. (2017). Extensive Turnover of Compatible Solutes in Cyanobacteria Revealed by Deuterium Oxide (D_2O) Stable Isotope Probing. *ACS Chem. Biol.* 12, 674–681. doi:10.1021/acschembio.6b00890

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* 55, 611–622. doi:10.1373/clinchem.2008.112797

Caldas, T., Demont-Caulet, N., Ghazi, A., and Richarme, G. (1999). Thermoprotection by glycine Betaine and Choline. *Microbiol* 145, 2543–2548. doi:10.1099/00221287-145-9-2543

Canton, B., Labno, A., and Endy, D. (2008). Refinement and Standardization of Synthetic Biological Parts and Devices. *Nat. Biotechnol.* 26, 787–793. doi:10.1038/nbt1413

Cholewa, J. M., Guimarães-Ferreira, L., and Zanchi, N. E. (2014). Effects of Betaine on Performance and Body Composition: a Review of Recent Findings and Potential Mechanisms. *Amino Acids* 46, 1785–1793. doi:10.1007/s00726-014-1748-5

Cleland, D., Krader, P., McCree, C., Tang, J., and Emerson, D. (2004). Glycine Betaine as a Cryoprotectant for Prokaryotes. *J. Microbiol. Methods* 58, 31–38. doi:10.1016/j.mimet.2004.02.015

Costello, A., and Badran, A. H. (2021). Synthetic Biological Circuits within an Orthogonal Central Dogma. *Trends Biotechnol.* 39, 59–71. doi:10.1016/j.tibtech.2020.05.013

Day, C. R., and Kempson, S. A. (2016). Betaine Chemistry, Roles, and Potential Use in Liver Disease. *Biochim. Biophys. Acta (BBA) - Gen. Subjects* 1860, 1098–1106. doi:10.1016/j.bbagen.2016.02.001

Desplats, P., Folco, E., and Salerno, G. L. (2005). Sucrose May Play an Additional Role to that of an Osmolyte in *Synechocystis* sp. PCC 6803 Salt-Shocked Cells. *Plant Physiol. Biochem.* 43, 133–138. doi:10.1016/j.plaphy.2005.01.008

Dikilitas, M., Simsek, E., and Roychoudhury, A. (2020). "Role of Proline and glycine Betaine in Overcoming Abiotic Stresses," in *Protective Chemical Agents in the Amelioration of Plant Abiotic Stress: Biochemical and Molecular Perspectives*. Editors A. Roychoudhury and D. K. Tripathi (John Wiley & Sons), 1–23. doi:10.1002/9781119552154.ch1

Du, W., Liang, F., Duan, Y., Tan, X., and Lu, X. (2013). Exploring the Photosynthetic Production Capacity of Sucrose by Cyanobacteria. *Metab. Eng.* 19, 17–25. doi:10.1016/j.ymben.2013.05.001

Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F. (1956). Colorimetric Method for Determination of Sugars and Related Substances. *Anal. Chem.* 28, 350–356. doi:10.1021/ac60111a017

DuPont (2015). Betafin® Natural Betaine Sustainable and Substantiated. Available at: http://animalnutrition.dupont.com/fileadmin/user_upload/live/animal_nutrition/documents/open/Betafin__Natural_Betaine_Lifecycle_Analysis.pdf (Accessed October 14, 2021).

Eklund, M., Bauer, E., Wamatu, J., and Mosenthin, R. (2005). Potential Nutritional and Physiological Functions of Betaine in Livestock. *Nutr. Res. Rev.* 18, 31–48. doi:10.1079/nrr200493

Ernst, A., Kirschenlohr, H., Diez, J., and Böger, P. (1984). Glycogen Content and Nitrogenase Activity in *Anabaena Variabilis*. *Arch. Microbiol.* 140, 120–125. doi:10.1007/bf00454913

Ferreira, E. A., Pacheco, C. C., Pinto, F., Pereira, J., Lamosa, P., Oliveira, P., et al. (2018). Expanding the Toolbox for *Synechocystis* sp. PCC 6803: Validation of Replicative Vectors and Characterization of a Novel Set of Promoters. *Synth. Biol. (Oxf)* 3, ysy014. doi:10.1093/synbio/ysy014

Ferreira, E. A., Pacheco, C. C., Rodrigues, J. S., Pinto, F., Lamosa, P., Fuente, D., et al. (2021). Heterologous Production of glycine Betaine Using *Synechocystis* sp. PCC 6803-based Chassis Lacking Native Compatible Solutes, Preprint. *Res. Square* (Accessed November 22, 2021). doi:10.21203/rs.3.rs-1031622/v1

Gopalakrishnan, S., Pakrasi, H. B., and Maranas, C. D. (2018). Elucidation of Photoautotrophic Carbon Flux Topology in *Synechocystis* PCC 6803 Using Genome-Scale Carbon Mapping Models. *Metab. Eng.* 47, 190–199. doi:10.1016/j.ymben.2018.03.008

Grossman, E. B., and Hebert, S. C. (1989). Renal Inner Medullary Choline Dehydrogenase Activity: Characterization and Modulation. *Am. J. Physiology-Renal Physiol.* 256, F107–F112. doi:10.1152/ajprenal.1989.256.1.f107

Guinn, E. J., Pegram, L. M., Capp, M. W., Pollock, M. N., and Record, M. T. (2011). Quantifying Why Urea Is a Protein Denaturant, whereas glycine Betaine Is a Protein Stabilizer. *Proc. Natl. Acad. Sci.* 108, 16932–16937. doi:10.1073/pnas.1109372108

Hagemann, M. (2013). "Genomics of Salt Acclimation," in *Adv Bot Res*. Editor C. C.-C. Franck Chauvat (Academic Press), 65, 27–55. doi:10.1016/b978-0-12-394313-2.00002-0

Hays, S. G., and Ducat, D. C. (2015). Engineering Cyanobacteria as Photosynthetic Feedstock Factories. *Photosynth Res.* 123, 285–295. doi:10.1007/s11120-014-9980-0

He, Y., He, C., Li, L., Liu, Z., Yang, A., and Zhang, J. (2011). Heterologous Expression of *ApGSMT2* and *ApDMT2* Genes from *Aphanothece Halophytica* Enhanced Drought Tolerance in Transgenic Tobacco. *Mol. Biol. Rep.* 38, 657–666. doi:10.1007/s11033-010-0152-9

Heidorn, T., Camsund, D., Huang, H.-H., Lindberg, P., Oliveira, P., Stensjö, K., et al. (2011). "Synthetic Biology in Cyanobacteria," in *Methods Enzymol*. Editor C. Voigt (Academic Press), 497, 539–579. doi:10.1016/b978-0-12-385075-1.00024-x

Heikkila, H. O., Melaja, J. A., Millner, D. E., and Virtanen, J. J. (1982). Betaine Recovery Process. US4359430A.

Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S. N., Richelle, A., Heinken, A., et al. (2019). Creation and Analysis of Biochemical Constraint-Based Models Using the COBRA Toolbox v.3.0. *Nat. Protoc.* 14, 639–702. doi:10.1038/s41596-018-0098-2

Holmström, K. O., Somersalo, S., Mandal, A., Palva, T. E., and Welin, B. (2000). Improved Tolerance to Salinity and Low Temperature in Transgenic Tobacco Producing glycine Betaine. *J. Exp. Bot.* 51, 177–185. doi:10.1093/jexbot/51.343.177

Huang, H.-H., Camsund, D., Lindblad, P., and Heidorn, T. (2010). Design and Characterization of Molecular Tools for a Synthetic Biology Approach towards Developing Cyanobacterial Biotechnology. *Nucleic Acids Res.* 38, 2577–2593. doi:10.1093/nar/gkq164

Iijima, H., Nakaya, Y., Kuwahara, A., Hirai, M. Y., and Osanai, T. (2015). Seawater Cultivation of Freshwater Cyanobacterium *Synechocystis* sp. PCC 6803 Drastically Alters Amino Acid Composition and Glycogen Metabolism. *Front. Microbiol.* 6, 326. doi:10.3389/fmicb.2015.00326

Iijima, H., Watanabe, A., Sukigara, H., Shirai, T., Kondo, A., and Osanai, T. (2020). Simultaneous Increases in the Levels of Compatible Solutes by Cost-effective Cultivation of *Synechocystis* sp. PCC 6803. *Biotechnol. Bioeng.* 117, 1649–1660. doi:10.1002/bit.27324

Incharoensakdi, A., and Waditee, R. (2000). Degradation of Glycinebetaine by Betaine-Homocysteine Methyltransferase in *Aphanothece halophytica*: Effect of Salt Downshock and Starvation. *Curr. Microbiol.* 41, 227–231. doi:10.1007/s002840010125

Ishitani, M., Takabe, T., Kojima, K., and Takabe, T. (1993). Regulation of Glycinebetaine Accumulation in the Halotolerant Cyanobacterium *Aphanothece Halophytica*. *Funct. Plant Biol.* 20, 693–703. doi:10.1071/pp9930693

Jittawuttipoka, T., Planchon, M., Spalla, O., Benzerara, K., Guyot, F., Cassier-Chauvat, C., et al. (2013). Multidisciplinary Evidences that *Synechocystis* PCC6803 Exopolysaccharides Operate in Cell Sedimentation and protection against Salt and Metal Stresses. *PLoS One* 8, e55564. doi:10.1371/journal.pone.0055564

Joshi, C. J., Peebles, C. A. M., and Prasad, A. (2017). Modeling and Analysis of Flux Distribution and Bioproduct Formation in *Synechocystis* sp. PCC 6803 Using a New Genome-Scale Metabolic Reconstruction. *Algal Res.* 27, 295–310. doi:10.1016/j.algal.2017.09.013

Kaleta, C., Schäuble, S., Rinas, U., and Schuster, S. (2013). Metabolic Costs of Amino Acid and Protein Production in *Escherichia coli*. *Biotechnol. J.* 8, 1105–1114. doi:10.1002/biot.201200267

Kanesaki, Y., Shiwa, Y., Tajima, N., Suzuki, M., Watanabe, S., Sato, N., et al. (2012). Identification of Substrain-specific Mutations by Massively Parallel Whole-Genome Resequencing of *Synechocystis* sp. PCC 6803. *DNA Res.* 19, 67–79. doi:10.1093/dnares/dsr042

Kar, J. R., Hallsworth, J. E., and Singhal, R. S. (2015). Fermentative Production of glycine Betaine and Trehalose from Acid Whey Using *Actinopolyspora Halophila* (MTCC 263). *Environ. Tech. Innovation* 3, 68–76. doi:10.1016/j.eti.2015.02.001

Kirsch, F., Klähn, S., and Hagemann, M. (2019). Salt-regulated Accumulation of the Compatible Solutes Sucrose and Glucosylglycerol in Cyanobacteria and its Biotechnological Potential. *Front. Microbiol.* 10, 2139. doi:10.3389/fmicb.2019.02139

Kirsch, F., Pade, N., Klähn, S., Hess, W. R., and Hagemann, M. (2017). The Glucosylglycerol-Degrading Enzyme GghA Is Involved in Acclimation to Fluctuating Salinities by the Cyanobacterium *Synechocystis* sp. Strain PCC 6803. *Microbiol* 163, 1319–1328. doi:10.1099/mic.0.000518

Klähn, S., and Hagemann, M. (2011). Compatible Solute Biosynthesis in Cyanobacteria. *Environ. Microbiol.* 13, 551–562. doi:10.1111/j.1462-2920.2010.02366.x

Knoll, A. H. (2008). "Cyanobacteria and Earth History," in *The Cyanobacteria: Molecular Biology, Genomics, and evolution*. Editors F. G. Flores and A. Herrero (Caister Academic Press), 1–19.

Landfald, B., and Strøm, A. R. (1986). Choline-glycine Betaine Pathway Confers a High Level of Osmotic Tolerance in *Escherichia coli*. *J. Bacteriol.* 165, 849–855. doi:10.1128/jb.165.3.849-855.1986

Leitão, E., Pereira, S., Bondoso, J., Ferreira, D., Pinto, F., Moradasferreira, P., et al. (2006). Genes Involved in the Maturation of Hydrogenase(s) in the

Nonheterocystous Cyanobacterium *Lyngbya Majuscula* CCAP 1446/4. *Int. J. Hydrogen Energ.* 31, 1469–1477. doi:10.1016/j.ijhydene.2006.06.012

Lever, M., and Slow, S. (2010). The Clinical Significance of Betaine, an Osmolyte with a Key Role in Methyl Group Metabolism. *Clin. Biochem.* 43, 732–744. doi:10.1016/j.clinbiochem.2010.03.009

Lindblad, P. (2018). "Hydrogen Production Using Novel Photosynthetic Cell Factories. Cyanobacterial Hydrogen Production: Design of Efficient Organisms," in *Microalgal Hydrogen Production: Achievements and Perspectives*. Editors G. Torzillo and M. Seibert (UK: G. Royal Society of Chemistry), 323–334.

Meeks, J. C., and Castenholz, R. W. (1971). Growth and Photosynthesis in an Extreme Thermophile, *Synechococcus Lividus* (Cyanophyta). *Archiv. Mikrobiol.* 78, 25–41. doi:10.1007/bf00409086

Miao, X., Wu, Q., Wu, G., and Zhao, N. (2003). Sucrose Accumulation in Salt-Stressed Cells Ofagpgene Deletion-Mutant in cyanobacterium *Synechocystis* sp. PCC 6803. *FEMS Microbiol. Lett.* 218, 71–77. doi:10.1111/j.1574-6968.2003.tb11500.x

Montagud, A., Navarro, E., Fernández de Córdoba, P., Urchueguía, J. F., and Patil, K. R. (2010). Reconstruction and Analysis of Genome-Scale Metabolic Model of a Photosynthetic Bacterium. *BMC Syst. Biol.* 4, 156. doi:10.1186/1752-0509-4-156

Montagud, A., Zelezniak, A., Navarro, E., de Córdoba, P. F., Urchueguía, J. F., and Patil, K. R. (2011). Flux Coupling and Transcriptional Regulation within the Metabolic Network of the Photosynthetic Bacterium *Synechocystis* sp. PCC6803. *Biotechnol. J.* 6, 330–342. doi:10.1002/biot.201000109

Mota, R., Guimarães, R., Büttel, Z., Rossi, F., Colica, G., Silva, C. J., et al. (2013). Production and Characterization of Extracellular Carbohydrate Polymer from *Cyanothece* sp. CCY 0110. *Carbohydr. Polym.* 92, 1408–1415. doi:10.1016/j.carbpol.2012.10.070

Nsimba, Z. F., Paquot, M., Mvumbi, L. G., and Deleu, M. (2010). Glycine Betaine Surfactant Derivatives: Synthesis Methods and Potentialities of Use. *Biotechnol. Agron. Soc. Envir* 14, 737–748.

Nyyssölä, A., Kerovuo, J., Kaukinen, P., von Weymarn, N., and Reinikainen, T. (2000). Extreme Halophiles Synthesize Betaine from glycine by Methylation. *J. Biol. Chem.* 275, 22196–22201. doi:10.1074/jbc.m910111199

Oren, A., Heldal, M., Norland, S., and Galinski, E. (2002). Intracellular Ion and Organic Solute Concentrations of the Extremely Halophilic Bacterium *Salinibacter Ruber*. *Extremophiles* 6, 491–498. doi:10.1007/s00792-002-0286-3

Orth, J. D., Thiele, I., and Palsson, B. Ø. (2010). What Is Flux Balance Analysis? *Nat. Biotechnol.* 28, 245–248. doi:10.1038/nbt.1614

Pacheco, C. C., Ferreira, E. A., Oliveira, P., and Tamagnini, P. (2021). "Synthetic Biology of Cyanobacteria," in *The Autotrophic Biorefinery: Raw Materials from Biotechnology*. Editors R. Kourist and S. Schmidt, 131–172. doi:10.1515/9783110550603-006

Pereira, S. B., Santos, M., Leite, J. P., Flores, C., Eisfeld, C., Büttel, Z., et al. (2019). The Role of the Tyrosine Kinase Wzc (Sll0923) and the Phosphatase Wzb (Slr0328) in the Production of Extracellular Polymeric Substances (EPS) by *Synechocystis* PCC 6803. *Microbiologyopen* 8, e00753. doi:10.1002/mbo3.753

Pinto, F., Pacheco, C. C., Ferreira, D., Moradas-Ferreira, P., and Tamagnini, P. (2012). Selection of Suitable Reference Genes for RT-qPCR Analyses in Cyanobacteria. *PLoS One* 7, e34983. doi:10.1371/journal.pone.0034983

Pinto, F., Pacheco, C. C., Oliveira, P., Montagud, A., Landels, A., Couto, N., et al. (2015). Improving a *Synechocystis*-Based Photoautotrophic Chassis through Systematic Genome Mapping and Validation of Neutral Sites. *DNA Res.* 22, 425–437. doi:10.1093/dnares/dsv024

Rady, M. O. A., Semida, W. M., Abd El-Mageed, T. A., Hemida, K. A., and Rady, M. M. (2018). Up-regulation of Antioxidative Defense Systems by glycine Betaine Foliar Application in Onion Plants Confer Tolerance to Salinity Stress. *Scientia Horticulturae* 240, 614–622. doi:10.1016/j.scienta.2018.06.069

Rathinasabapathi, B., Burnet, M., Russell, B. L., Gage, D. A., Liao, P.-C., Nye, G. J., et al. (1997). Choline Monooxygenase, an Unusual Iron-Sulfur Enzyme Catalyzing the First Step of glycine Betaine Synthesis in Plants: Prosthetic Group Characterization and cDNA Cloning. *Proc. Natl. Acad. Sci.* 94, 3454–3458. doi:10.1073/pnas.94.7.3454

Rodrigues, J. S., and Lindberg, P. (2021). "Engineering Cyanobacteria as Host Organisms for Production of Terpenes and Terpenoids," in *Cyanobacteria Biotechnology*. Editors S. L. Jens Nielsen, G. Stephanopoulos, and P. Hudson (WILEY-VCH GmbH), 267–300. doi:10.1002/9783527824908.ch9

Sadvakasova, A. K., Kossalbayev, B. D., Zayadan, B. K., Bolatkhan, K., Alwasel, S., Najafpour, M. M., et al. (2020). Bioprocesses of Hydrogen Production by Cyanobacteria Cells and Possible Ways to Increase Their Productivity. *Renew. Sust. Energ. Rev.* 133, 110054. doi:10.1016/j.rser.2020.110054

Sambrook, J., and Russel, D. (2001). *Molecular Cloning: A Laboratory Manual*. 3rd ed. NY, USA: Cold Spring Harbor Laboratory Press.

Santos, H., Lamosa, P., and Borges, N. (2006). "Characterization and Quantification of Compatible Solutes in (Hyper)thermophilic Microorganisms," in *Methods in Microbiology*. Editors F. A. Rainey and A. Oren (Academic Press), 35, 173–199. Extremophiles. doi:10.1016/s0580-9517(08)70011-4

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and Pühler, A. (1994). Small Mobilizable Multi-Purpose Cloning Vectors Derived from the *Escherichia coli* Plasmids pK18 and pK19: Selection of Defined Deletions in the Chromosome of *Corynebacterium Glutamicum*. *Gene* 145, 69–73. doi:10.1016/0378-1119(94)90324-7

Shkryl, Y., Degtyarenko, A., Grigorchuk, V., Balabanova, L., and Tekutyeva, L. (2020). Betaine Biosynthesis in a Heterologous Expression System Based on the B12 Producer Pseudomonas Denitrificans. Paper presented at the E3S Web of Conferences .

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de las Heras, A., et al. (2013). The Standard European Vector Architecture (SEVA): a Coherent Platform for the Analysis and Deployment of Complex Prokaryotic Phenotypes. *Nucleic Acids Res.* 41, D666–D675. doi:10.1093/nar/gks1119

Singh, M., Sharma, N. K., Prasad, S. B., Yadav, S. S., Narayan, G., and Rai, A. K. (2013). The Freshwater Cyanobacterium *Anabaena Doliolum* Transformed with *ApGSMT-DMT* Exhibited Enhanced Salt Tolerance and protection to Nitrogenase Activity, but Became Halophilic. *Microbiol* 159, 641–648. doi:10.1099/mic.0.065078-0

Stadmiller, S. S., Gorensek-Benitez, A. H., Guseman, A. J., and Pielak, G. J. (2017). Osmotic Shock Induced Protein Destabilization in Living Cells and its Reversal by glycine Betaine. *J. Mol. Biol.* 429, 1155–1161. doi:10.1016/j.jmb.2017.03.001

Stanier, R. Y., Kunisawa, R., Mandel, M., and Cohen-Bazire, G. (1971). Purification and Properties of Unicellular Blue-green Algae (Order *Chroococcales*). *Bacteriol. Rev.* 35, 171–205. doi:10.1128/br.35.2.171-205.1971

Sun, H., Luo, M., Zhou, X., Zhou, Q., Sun, Y., Ge, W., et al. (2020). Exogenous glycine Betaine Treatment Alleviates Low Temperature-Induced Pericarp browning of 'Nanguo' Pears by Regulating Antioxidant Enzymes and Proline Metabolism. *Food Chem.* 306, 125626. doi:10.1016/j.foodchem.2019.125626

Tamagnini, P., Troshina, O., Oxelfelt, F., Salema, R., and Lindblad, P. (1997). Hydrogenases in *Nostoc* sp. Strain PCC 73102, a Strain Lacking a Bidirectional Enzyme. *Appl. Environ. Microbiol.* 63, 1801–1807. doi:10.1128/aem.63.5.1801-1807.1997

Trautmann, D., Voss, B., Wilde, A., Al-Babili, S., and Hess, W. R. (2012). Microevolution in Cyanobacteria: Re-sequencing a Motile Substrain of *Synechocystis* sp. PCC 6803. *DNA Res.* 19, 435–448. doi:10.1093/dnares/dss024

Waditee, R., Bhuiyan, M. N. H., Rai, V., Aoki, K., Tanaka, Y., Hibino, T., et al. (2005). Genes for Direct Methylation of glycine Provide High Levels of glycinebetaine and Abiotic-Stress Tolerance in *Synechococcus* and *Arabidopsis*. *Proc. Natl. Acad. Sci.* 102, 1318–1323. doi:10.1073/pnas.0409017102

Waditee, R., Bhuiyan, N. H., Hirata, E., Hibino, T., Tanaka, Y., Shikata, M., et al. (2007). Metabolic Engineering for Betaine Accumulation in Microbes and Plants. *J. Biol. Chem.* 282, 34185–34193. doi:10.1074/jbc.m704939200

Waditee, R., Tanaka, Y., Aoki, K., Hibino, T., Jikuya, H., Takano, J., et al. (2003). Isolation and Functional Characterization of N-Methyltransferases that Catalyze Betaine Synthesis from Glycine in a Halotolerant Photosynthetic Organism *Aphanothece halophytica*. *J. Biol. Chem.* 278, 4932–4942. doi:10.1074/jbc.m210970200

Waditee-Sirisattha, R., Kageyama, H., Fukaya, M., Rai, V., and Takabe, T. (2015). Nitrate and Amino Acid Availability Affects glycine Betaine and Mycosporine-2-glycine in Response to Changes of Salinity in a Halotolerant cyanobacterium *Aphanothece halophytica*. *FEMS Microbiol. Lett.* 362, fnv198. doi:10.1093/femsle/fnv198

Waditee-Sirisattha, R., Singh, M., Kageyama, H., Sittipol, D., Rai, A. K., and Takabe, T. (2012). *Anabaena* sp. PCC7120 Transformed with glycine Methylation Genes from *Aphanothece Halophytica* Synthesized glycine

Betaine Showing Increased Tolerance to Salt. *Arch. Microbiol.* 194, 909–914. doi:10.1007/s00203-012-0824-z

Wang, L., Chen, L., Yang, S., and Tan, X. (2020). Photosynthetic Conversion of Carbon Dioxide to Oleochemicals by Cyanobacteria: Recent Advances and Future Perspectives. *Front. Microbiol.* 11, 634. doi:10.3389/fmicb.2020.00634

Williams, J. G. K. (1988). [85] Construction of Specific Mutations in Photosystem II Photosynthetic Reaction center by Genetic Engineering Methods in *Synechocystis* 6803. *Methods Enzymol.* 167, 766–778. doi:10.1016/0076-6879(88)67088-1

Wu, W., Du, W., Gallego, R. P., Hellingwerf, K. J., van der Woude, A. D., and Branco dos Santos, F. (2020). Using Osmotic Stress to Stabilize Mannitol Production in *Synechocystis* sp. PCC6803. *Biotechnol. Biofuels* 13, 117. doi:10.1186/s13068-020-01755-3

You, L., Song, Q., Wu, Y., Li, S., Jiang, C., Chang, L., et al. (2019). Accumulation of glycine Betaine in Transplastomic Potato Plants Expressing Choline Oxidase Confers Improved Drought Tolerance. *Planta* 249, 1963–1975. doi:10.1007/s00425-019-03132-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Bioengineering and Biotechnology

MINI REVIEW
published: 11 March 2022
doi: 10.3389/fbioe.2022.869195

Synthetic Biology Approaches for Improving Chemical Production in Cyanobacteria

Tanner R. Treece[1], *Jake N. Gonzales*[1,2], *Joseph R. Pressley*[1] *and Shota Atsumi*[1,2]*

[1]Department of Chemistry, University of California, Davis, Davis, CA, United States, [2]Plant Biology Graduate Group, University of California, Davis, Davis, CA, United States

Biological chemical production has gained traction in recent years as a promising renewable alternative to traditional petrochemical based synthesis. Of particular interest in the field of metabolic engineering are photosynthetic microorganisms capable of sequestering atmospheric carbon dioxide. CO_2 levels have continued to rise at alarming rates leading to an increasingly uncertain climate. CO_2 can be sequestered by engineered photosynthetic microorganisms and used for chemical production, representing a renewable production method for valuable chemical commodities such as biofuels, plastics, and food additives. The main challenges in using photosynthetic microorganisms for chemical production stem from the seemingly inherent limitations of carbon fixation and photosynthesis resulting in slower growth and lower average product titers compared to heterotrophic organisms. Recently, there has been an increase in research around improving photosynthetic microorganisms as renewable chemical production hosts. This review will discuss the various efforts to overcome the intrinsic inefficiencies of carbon fixation and photosynthesis, including rewiring carbon fixation and photosynthesis, investigating alternative carbon fixation pathways, installing sugar catabolism to supplement carbon fixation, investigating newly discovered fast growing photosynthetic species, and using new synthetic biology tools such as CRISPR to radically alter metabolism.

Keywords: CO_2 fixation, cyanobacteria, photosynthesis, RuBisCO, CRISPR

OPEN ACCESS

Edited by:
M. Kalim Akhtar,
United Arab Emirates University,
United Arab Emirates

Reviewed by:
Weiwen Zhang,
Tianjin University, China

***Correspondence:**
Shota Atsumi
satsumi@ucdavis.edu

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *04 February 2022*
Accepted: *25 February 2022*
Published: *11 March 2022*

Citation:
Treece TR, Gonzales JN, Pressley JR and Atsumi S (2022) Synthetic Biology Approaches for Improving Chemical Production in Cyanobacteria. Front. Bioeng. Biotechnol. 10:869195. doi: 10.3389/fbioe.2022.869195

INTRODUCTION

It is well established that rising atmospheric CO_2 levels are the primary cause for unprecedented climate change impacting the globe (Solomon et al., 2009). Despite this, chemical production still relies mostly on petroleum-based synthesis (Levi and Cullen, 2018). In response to the growing concern over greenhouse gasses, research with a focus on more sustainable chemical production has become high priority. The fields of synthetic biology and metabolic engineering aim to achieve a more sustainable method for chemical production using engineered organisms. These efforts include the use of both heterotrophic and photosynthetic microorganisms. Heterotrophic chemical production involves a carbon input of a sugar feedstock to a microorganism to generate a biochemical product as an output. Alternatively, using photosynthetic microorganisms, such as cyanobacteria, offers the advantage of eliminating the need for sugar feedstocks and the ability to generate valuable chemical commodities from CO_2 and sunlight. The two predominant categories of photosynthetic microorganisms being investigated for chemical production are microalgae and cyanobacteria. Microalgae are a diverse group of photosynthetic eukaryotes that have been shown to be viable production hosts for a wide array of useful chemical commodities ranging from biofuels to lipids and vitamins (Sproles et al., 2021). Cyanobacteria are a group of prokaryotic microorganisms with some of the fastest carboxylation rates present in photosynthetic organisms

(Flamholz et al., 2019). The focus of this review will be on recent synthetic biology research in cyanobacteria which have garnered interest as efficient photosynthetic chemical production hosts.

Despite the burgeoning interest around these photosynthetic microorganisms, the field still faces many challenges that have yet to be addressed. The primary concern is the inefficient nature of photosynthesis and CO_2 fixation. Attempts to improve upon the central carbon fixation enzyme ribulose-1,5-bisphosphate carboxylase-oxygenase (RuBisCO) have been met with little success (Erb and Zarzycki, 2018; Flamholz et al., 2019). Part of the challenge with RuBisCO is its inability to distinguish between CO_2 and O_2 with high specificity and notably, the oxygenase activity of RuBisCO results in an energetically costly pathway known as photorespiration which has widespread effects on the growth and metabolic needs of many species of photosynthetic organisms (Hagemann and Bauwe, 2016). Recent research suggests that photorespiration is a symptom of RuBisCO evolving in a high CO_2 environment where enzymatic specificity was not as vital, with this in mind, recent studies are investigating the possibilities of by reviving ancestral forms of the protein and subjecting it to new environments in the hopes of generating biologically important variants (Shih et al., 2016). Other efforts are looking towards natural adaptations of CO_2 fixation for inspiration with a focus on the carboxysome, a bacterial microcompartment that acts to localize RuBisCO with high concentrations of CO_2 (Kerfeld and Melnicki, 2016). These studies aim to avoid photorespiration by engineering synthetic protein structures to mimic cyanobacterial carboxysomes to concentrate CO_2 near RuBisCO and competitively inhibit the reaction with oxygen (Borden and Savage, 2021). Other efforts to further improve these photosynthetic organisms as chemical production hosts include engineering superior light delivery systems for bioreactors and engineering the light harvesting complexes to take advantage of the entire visible light spectrum (Stephens et al., 2021).

The focus of this review will be on the recent methods employed to overcome the supposed shortcomings of photosynthetic organisms ranging from rewiring carbon metabolism and photosynthesis, introducing additional carbon substrates, generating other chassis organisms capable of superior carbon sequestration, studying faster growing variants of cyanobacteria, and developing new tools via synthetic biology.

Rewiring Photosynthetic Metabolism

Efforts have been made to overcome the intrinsic shortcomings of photosynthetic microorganisms by rewiring metabolism related to carbon fixation and photosynthesis. While efforts to improve RuBisCO have not been met with much success, current research has shifted focus towards rerouting metabolism to improve overall photosynthetic efficiency by focusing on key aspects of the Calvin-Benson cycle or the photosynthetic electron transport chain (PETC). One strategy used to harness the excess energy being lost by the PETC in cyanobacteria involved overexpressing the protein OmcS (Meng et al., 2021). This strategy coupled the excess electrons from the PETC to NADH production and was shown to increase intracellular ATP and NADH allowing for a fourfold improvement of D-lactate production in the cyanobacterium *Synechococcus elongatus* UTEX 2973 (hereon 2973) (Meng et al., 2021).

An inherent drawback of RuBisCO is its promiscuous nature, when RuBisCO undergoes oxygenase activity a costly side pathway known as photorespiration occurs where the oxygenase product is recycled back into usable metabolism consuming energy and losing CO_2 in the process. As much as 30% of energy produced by photosynthesis has been observed to be lost through photorespiration in plants (Hagemann and Bauwe, 2016). Rewiring or preventing photorespiration represents a promising way to improve the overall efficiency of carbon fixation in photosynthetic organisms. Efforts to rewire photorespiration generally involve deleting energetically costly steps, circumventing steps where CO_2 is lost, and rerouting metabolites towards central carbon metabolism (Hagemann and Bauwe, 2016). One of the more ambitious efforts to ameliorate the cost of photorespiration was the expression of a synthetic carbon capture pathway to serve as both a photorespiratory bypass and as a supplement to the Calvin-Benson cycle, this was shown to be a viable use of synthetic biology to counteract the costly natural photorespiration pathway (Shih et al., 2014).

It should be noted that photorespiration is not the sole pathway responsible for carbon inefficiencies, many metabolic processes include steps where CO_2 is lost to the environment. An important way to engineer microorganisms for sustainability involves carbon conservation, focusing on rerouting metabolism to circumvent decarboxylation reactions (François et al., 2020). Of the more notable strategies is the non-oxidative glycolysis pathway (NOG) which has been shown to function in *Escherichia* coli and which can effectively conserve all carbon associated with sugar catabolism to acetyl-CoA (Bogorad et al., 2013). While carbon conservation is a powerful methodology for engineering metabolism, the field is still in its infancy and further work is required to evaluate the industrial viability of many carbon conservation strategies. Additionally, *de novo* carbon fixation pathways, which will be addressed later in this review, are currently being developed and may prove to be a better methodology for the development of sustainable production hosts.

Non-RuBisCO Carbon Fixation

In contrast to research centering on canonical CO_2 fixation, investigations into *de-novo* CO_2 fixation pathways have been explored and theorized in recent years as more efficient alternatives to traditional RuBisCO based CO_2 assimilation. These pathways may provide advantages in chemical production hosts by offering insight into carboxylation reactions that could work in tandem with RuBisCO. The expression of formate dehydrogenase in the cyanobacterium, *Anabaena* sp. PCC 7120, was shown to successfully increase intracellular formate concentration, representing an alternative to the photo-reduction of CO_2 and can act to supplement natural carbon fixation pathways (Ihara et al., 2013).

Many of these *de novo* CO_2 fixation pathways have had limited success when installed into model organisms such as *E. coli* and yeast and it has yet to be shown if these pathways can function effectively in photosynthetic hosts. One pathway of note that has been shown to work in *E. coli* is the reductive glycine pathway, hereafter RGP (Tashiro et al., 2018). This pathway leverages the native glycine cleavage system in the reverse direction to combine one equivalent of CO_2 with 5,10-methylenetetrahydrofolate that has been produced from formate to produce pyruvate. This

method allows *E. coli* to directly assimilate CO_2 into central metabolic pathways and is a more efficient method for CO_2 fixation than traditional RuBisCO (Bar-Even et al., 2013). This inorganic carbon can then be leveraged for biochemical synthesis. Additional work has also recently shown that the expression of formate dehydrogenase confers further renewable characteristics to strains harboring the RGP by removing the need for glucose supplementation (Bang et al., 2020). Other notable CO_2 fixation pathways include the crotonyl-CoA/ethylmalonyl-CoA/hydroxybutyryl-CoA (CETCH) cycle and the tartronyl-CoA (TaCo) pathways (Scheffen et al., 2021). While the RGP has proven to be a viable carbon fixation pathway that was shown to function in *E. coli,* the growth exhibited by this CO_2 fixing *E. coli* is slower than its traditional heterotrophic phenotype (Tashiro et al., 2018; Bang et al., 2020). Advancements in modeling and metabolomics may allow for an increase in the creation of *de-novo* carbon fixation pathways that may prove to be both more efficient than traditional pathways and capable of functioning in a wider array of chemical production hosts.

Photomixotrophy

Another approach to increase chemical production capacity is to supplement CO_2 with carbohydrates as an auxiliary carbon source for the Calvin-Benson cycle, thereby making the organism photomixotrophic. By re-engineering glucose catabolism to direct carbon flux into the Calvin-Benson cycle, more ribulose-1,5-bisphosphate can be supplied to RuBisCO, accelerating CO_2 fixation. This ultimately results in faster growth and production of downstream targets, as well as a six-fold increase in titer once metabolism was rewired to accommodate photomixotrophy (Kanno et al., 2017). The addition of a heterotrophic mode also allows for CO_2 fixation in darkness, resulting in a 24 h production period under natural diurnal conditions. Glucose can be readily obtained from the acid hydrolysis of agricultural waste products such as corn stover, in conjunction with other sugars: xylose, arabinose, and galacturonic acid (Mourtzinis et al., 2016). By installing catabolic pathways for these non-glucose sugars, these agricultural waste products can be used more efficiently. Xylose, the second most abundant sugar in corn stover lysate, has successfully been used to achieve photomixotrophic production of 2,3-butanediol in light and dark conditions with significant improvements in growth and product titer over the equivalent photoautotrophic organism (McEwen et al., 2016). A similar strategy has also recently been used to improve the production of 3-hydroxypropionic acid by 4.1 fold in cyanobacteria through the installation of a xylose photomixotrophic module along with other modifications to help assimilate the additional carbon source (Yao et al., 2022). It should also be noted that the cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter 6803) natively possesses photomixotrophic machinery to assimilate glucose. A recent study was able to improve these photomixotrophic conditions through the installation of the NOG pathway along with targeted gene knockouts to increase the intracellular concentration of acetyl-CoA, thereby improving the growth phenotype of 6803 (Song et al., 2021).

Discovery of Fast-Growing Cyanobacteria

Most research done in this field focuses on using just a handful of species that have traditionally been used as model organisms to study the mechanics of photosynthesis. With increasing interest in using photosynthetic organisms for industrial production there have been efforts to uncover new species that are faster growing and more receptive to engineering. In the realm of cyanobacterial chemical production, species like 2973 and *S. elongatus* PCC 11802 (hereon 11802) have risen in popularity as they are faster growing than the traditional *S. elongatus* PCC 7942 (hereon 7942) and have been shown to produce higher titers of target chemical products under certain circumstances (Yu et al., 2015; Sengupta et al., 2020). Additionally, these fast-growing organisms are providing inspiration for how to better engineer existing model organisms. The fast-growing cyanobacterium 2973 has relatively little differences genetically when compared to the model cyanobacterium 7942 (Yu et al., 2015). However, a notable difference in 2973 is an increase in the expression levels of PSI, cytochrome b_6f, and plastocyanin on a per cell basis which improves the downstream flux of electrons from PSII, which helps the faster growing cyanobacteria to better utilize photosynthetic energy (Ungerer et al., 2018). The discovery of new fast-growing cyanobacteria may enhance our understanding of photosynthesis and characterizing the differences between these new species with current model organisms.

Genome Engineering Tools

In cyanobacteria, traditional genomic modifications are a labor-intensive task and limiting in nature due to the polyploidal nature of these organisms and the need for antibiotic resistance markers (Griese et al., 2011). The current methodology for genomic integration involves constructing a plasmid with an antibiotic selection marker in a plasmid host such as *E. coli.* After introduction of this plasmid to cyanobacteria, several rounds of antibiotic screening are required to ensure complete genome segregation (Golden et al., 1987). This process generally limits the number of modifications that can be performed in a single strain due to the physiological constraints of expressing multiple different antibiotic resistance genes.

The overall task of metabolic engineering in cyanobacteria has been made dramatically more efficient thanks to the advent of CRISPR gene editing which allows for markerless edits (Behler et al., 2018). However, the protein Cas9 is toxic to a number of cyanobacteria species (Wendt et al., 2016). Researchers have recently uncovered other endonucleases that are similarly capable of CRISPR gene editing. The main endonuclease of interest is Cpf1 which, while similar to Cas9, is better tolerated by photosynthetic hosts (Ungerer and Pakrasi, 2016). As the body of research grows around Cpf1, more engineering strategies will be made available in the realm of photosynthetic chemical production and should offer a boon towards the viability of these organisms to begin replacing their non-CO_2 fixing brethren in the realm of biochemical production (Bishé et al., 2019; Niu et al., 2019). Additionally, having the ability to perform markerless genomic modifications unlocks the potential to engineer these microorganisms far more ambitiously than what was previously possible.

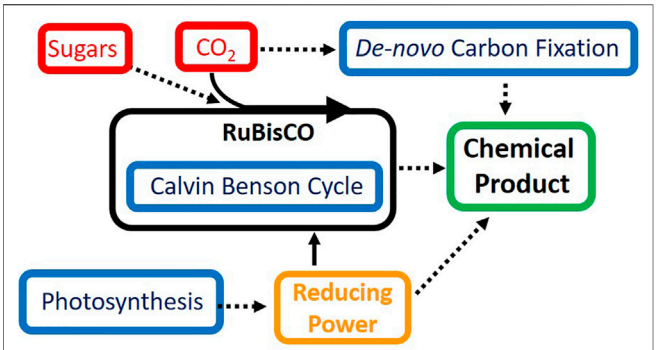


FIGURE 1 | Strategies for improving efficiency of carbon fixation and photosynthesis in cyanobacteria. Shown in red are carbon sources for metabolism in cyanobacteria, in blue are pathways of interest to synthetic biology and metabolic engineering, solid lines represent native pathways and dashed lines represent pathways of interest for improvement using synthetic biology.

Other work on CRISPR technologies in cyanobacteria includes the use of CRISPR inhibition (CRISPRi) by using dead Cas9 (dCas9) (Qi et al., 2013). While the endonuclease activity of the intact Cas9 protein seems to be toxic to these production hosts, dCas9 is able to function in the same manner in photosynthetic hosts as it is able to in heterotrophic hosts such as *E. coli* (Santos et al., 2021). While the use of dCas9 may not be as broadly useful as traditional CRISPR, dCas9 has been shown to be invaluable in certain chemical production applications where more traditional gene knockouts would otherwise be toxic.

CONCLUDING REMARKS

While many challenges remain and must be overcome to enable widespread adoption of photosynthetic chemical production hosts, the above studies suggest that there are myriad avenues of research that can get closer to this goal. The renewed interest in the field due to the ongoing climate crisis has spurred efforts to improve and adopt these microorganisms as a sustainable alternative for traditional petroleum-based synthesis. Many of the challenges in this field revolve around the intrinsic inefficiencies of carbon fixation and photosynthesis. While engineering RuBisCO remains an interesting target for improving carbon fixation, it has proven to be highly resistant to traditional engineering and decades of research would suggest that it is next to impossible to improve. Focusing on engineered carbon fixation pathways is a more promising route towards improving the carbon sequestration ability of cyanobacteria. Other research into improving the efficiency of photosynthesis by introducing alternative pathways downstream of the PETC for the production of chemical products is a prime example of how we can engineer these microbes to make full use of excess reducing potential from the PETC. The aforementioned approaches aim to enhance our understanding of the inefficiencies related to carbon fixation and photosynthesis while also representing some of the more novel approaches being undertaken by the field of synthetic biology. The discovery of new synthetic biology tools and investigation into faster growing cyanobacteria is also expanding the field of photosynthetic microbial research to make photosynthetic microorganisms a more viable alternative to petroleum based chemical production.

Of the discussed challenges for synthetic biology in cyanobacteria, improving the rate and efficiency of carbon fixation seems to be the most difficult, however, this task also holds the most promise. While RuBisCO is resistant to direct engineering strategies, adding additional carbon fixation modules can enhance the viability of cyanobacteria as a chemical production chassis. Further research into *de novo* carbon fixation pathways capable of operating in parallel to the Calvin-Benson cycle and RuBisCO holds great promise for circumventing the inefficiencies of carbon fixation in cyanobacteria. Multiple carbon fixation pathways operating in tandem could exponentially increase the amount of CO_2 sequestered by cyanobacteria and greatly enhance growth and product formation. The process of carbon fixation is a highly regulated process, and this strategy will likely face further challenges before successful implementation. Overall, it is highly likely that the optimal route for improving the conversion of CO_2 into valuable chemical commodities in cyanobacteria lies in exploiting multiple of the aforementioned strategies contained within this review (**Figure 1**).

As new synthetic biology tools become available for cyanobacteria, high throughput screening will allow for rapid progress to be made within this field. The advent of CRISPR technology has had profound effects on research in a wide variety of fields but is relatively new to cyanobacteria. Additionally, faster growing species of cyanobacteria are rising in popularity and more have yet to be discovered. While these newly discovered cyanobacteria are efficient production hosts in their own right, they also inform the field on future targets for modification. Understanding the inherent differences among these organisms is vital to improving our understanding of carbon fixation and photosynthesis. It stands to reason that research into discovering more of these faster growing species, as well as studying the known cyanobacterial variants will provide insight and guidance for future work in this field. While improving photosynthetic production hosts has been historically difficult, the studies described in this work point to a promising future.

AUTHOR CONTRIBUTIONS

TT, JG, JP, and SA wrote the manuscript.

FUNDING

This work was supported by the National Science Foundation (CBET-1902014).

REFERENCES

Bang, J., Hwang, C. H., Ahn, J. H., Lee, J. A., and Lee, S. Y. (2020). *Escherichia coli* Is Engineered to Grow on CO2 and Formic Acid. *Nat. Microbiol.* 5, 1459–1463. doi:10.1038/s41564-020-00793-9

Bar-Even, A., Noor, E., Flamholz, A., and Milo, R. (2013). Design and Analysis of Metabolic Pathways Supporting Formatotrophic Growth for Electricity-dependent Cultivation of Microbes. *Biochim. Biophys. Acta Bioenerg.* 1827, 1039–1047. doi:10.1016/j.bbabio.2012.10.013

Behler, J., Vijay, D., Hess, W. R., and Akhtar, M. K. (2018). CRISPR-based Technologies for Metabolic Engineering in Cyanobacteria. *Trends Biotechnol.* 36, 996–1010. doi:10.1016/j.tibtech.2018.05.011

Bishé, B., Taton, A., and Golden, J. W. (2019). Modification of RSF1010-Based Broad-Host-Range Plasmids for Improved Conjugation and Cyanobacterial Bioprospecting. *iScience* 20, 216–228. doi:10.1016/j.isci.2019.09.002

Bogorad, I. W., Lin, T.-S., and Liao, J. C. (2013). Synthetic Non-oxidative Glycolysis Enables Complete Carbon Conservation. *Nature* 502, 693–697. doi:10.1038/nature12575

Borden, J. S., and Savage, D. F. (2021). New Discoveries Expand Possibilities for Carboxysome Engineering. *Curr. Opin. Microbiol.* 61, 58–66. doi:10.1016/j.mib.2021.03.002

Erb, T. J., and Zarzycki, J. (2018). A Short History of RubisCO: the Rise and Fall (?) of Nature's Predominant CO2 Fixing Enzyme. *Curr. Opin. Biotechnol.* 49, 100–107. doi:10.1016/j.copbio.2017.07.017

Flamholz, A. I., Prywes, N., Moran, U., Davidi, D., Bar-On, Y. M., Oltrogge, L. M., et al. (2019). Revisiting Trade-Offs between Rubisco Kinetic Parameters. *Biochemistry* 58, 3365–3376. doi:10.1021/acs.biochem.9b00237

François, J. M., Lachaux, C., and Morin, N. (2020). Synthetic Biology Applied to Carbon Conservative and Carbon Dioxide Recycling Pathways. *Front. Bioeng. Biotechnol.* 7, 1–16. doi:10.3389/fbioe.2019.00446

Golden, S. S., Brusslan, J., and Haselkorn, R. (1987). Genetic Engineering of the Cyanobacterial Chromosome. *Methods Enzymol.* 153, 215–231. doi:10.1016/0076-6879(87)53055-5

Griese, M., Lange, C., and Soppa, J. (2011). Ploidy in Cyanobacteria. *FEMS Microbiol. Lett.* 323, 124–131. doi:10.1111/j.1574-6968.2011.02368.x

Hagemann, M., and Bauwe, H. (2016). Photorespiration and the Potential to Improve Photosynthesis. *Curr. Opin. Chem. Biol.* 35, 109–116. doi:10.1016/j.cbpa.2016.09.014

Ihara, M., Kawano, Y., Urano, M., and Okabe, A. (2013). Light Driven CO2 Fixation by Using Cyanobacterial Photosystem I and NADPH-dependent Formate Dehydrogenase. *PLoS ONE* 8, e71581–8. doi:10.1371/journal.pone.0071581

Kanno, M., Carroll, A. L., and Atsumi, S. (2017). Global Metabolic Rewiring for Improved CO2 Fixation and Chemical Production in Cyanobacteria. *Nat. Commun.* 8, 1–11. doi:10.1038/ncomms14724

Kerfeld, C. A., and Melnicki, M. R. (2016). Assembly, Function and Evolution of Cyanobacterial Carboxysomes. *Curr. Opin. Plant Biol.* 31, 66–75. doi:10.1016/j.pbi.2016.03.009

Levi, P. G., and Cullen, J. M. (2018). Mapping Global Flows of Chemicals: From Fossil Fuel Feedstocks to Chemical Products. *Environ. Sci. Technol.* 52, 1725–1734. doi:10.1021/acs.est.7b04573

McEwen, J. T., Kanno, M., and Atsumi, S. (2016). 2,3 Butanediol Production in an Obligate Photoautotrophic Cyanobacterium in Dark Conditions via Diverse Sugar Consumption. *Metab. Eng.* 36, 28–36. doi:10.1016/j.ymben.2016.03.004

Meng, H., Zhang, W., Zhu, H., Yang, F., Zhang, Y., Zhou, J., et al. (2021). Over-expression of an Electron Transport Protein OmcS Provides Sufficient NADH for D-Lactate Production in Cyanobacterium. *Biotechnol. Biofuels* 14, 1–16. doi:10.1186/s13068-021-01956-4

Mourtzinis, S., Cantrell, K. B., Arriaga, F. J., Balkcom, K. S., Novak, J. M., Frederick, J. R., et al. (2016). Carbohydrate and Nutrient Composition of Corn stover from Three southeastern USA Locations. *Biomass Bioenergy* 85, 153–158. doi:10.1016/j.biombioe.2015.11.031

Niu, T.-C., Lin, G.-M., Xie, L.-R., Wang, Z.-Q., Xing, W.-Y., Zhang, J.-Y., et al. (2019). Expanding the Potential of CRISPR-Cpf1-Based Genome Editing Technology in the CyanobacteriumAnabaenaPCC 7120. *ACS Synth. Biol.* 8, 170–180. doi:10.1021/acssynbio.8b00437

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., et al. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-specific Control of Gene Expression. *Cell* 152, 1173–1183. doi:10.1016/j.cell.2013.02.022

Santos, M., Pacheco, C. C., Yao, L., Hudson, E. P., and Tamagnini, P. (2021). CRISPRi as a Tool to Repress Multiple Copies of Extracellular Polymeric Substances (EPS)-Related Genes in the Cyanobacterium Synechocystis Sp. PCC 6803. *Life* 11, 1198. doi:10.3390/life11111198

Scheffen, M., Marchal, D. G., Beneyton, T., Schuller, S. K., Klose, M., Diehl, C., et al. (2021). A New-To-Nature Carboxylation Module to Improve Natural and Synthetic CO2 Fixation. *Nat. Catal.* 4, 105–115. doi:10.1038/s41929-020-00557-y

Sengupta, S., Jaiswal, D., Sengupta, A., Shah, S., Gadagkar, S., and Wangikar, P. P. (2020). Metabolic Engineering of a Fast-Growing Cyanobacterium *Synechococcus Elongatus* PCC 11801 for Photoautotrophic Production of Succinic Acid. *Biotechnol. Biofuels* 13, 1–18. doi:10.1186/s13068-020-01727-7

Shih, P. M., Zarzycki, J., Niyogi, K. K., and Kerfeld, C. A. (2014). Introduction of a Synthetic CO2-fixing Photorespiratory Bypass into a Cyanobacterium. *J. Biol. Chem.* 289, 9493–9500. doi:10.1074/jbc.C113.543132

Shih, P. M., Occhialini, A., Cameron, J. C., Andralojc, P. J., Parry, M. A. J., and Kerfeld, C. A. (2016). Biochemical Characterization of Predicted Precambrian RuBisCO. *Nat. Commun.* 7, 1–11. doi:10.1038/ncomms10382

Solomon, S., Plattner, G.-K., Knutti, R., and Friedlingstein, P. (2009). Irreversible Climate Change Due to Carbon Dioxide Emissions. *PNAS* 106, 1704–1709. doi:10.1073/pnas.0812721106

Song, X., Diao, J., Yao, J., Cui, J., Sun, T., Chen, L., et al. (2021). Engineering a Central Carbon Metabolism Pathway to Increase the Intracellular Acetyl-CoA Pool in *Synechocystis* Sp. PCC 6803 Grown under Photomixotrophic Conditions. *ACS Synth. Biol.* 10, 836–846. doi:10.1021/acssynbio.0c00629

Sproles, A. E., Fields, F. J., Smalley, T. N., Le, C. H., Badary, A., and Mayfield, S. P. (2021). Recent Advancements in the Genetic Engineering of Microalgae. *Algal Res.* 53, 102158. doi:10.1016/j.algal.2020.102158

Stephens, S., Mahadevan, R., and Allen, D. G. (2021). Engineering Photosynthetic Bioprocesses for Sustainable Chemical Production: A Review. *Front. Bioeng. Biotechnol.* 8, 1–15. doi:10.3389/fbioe.2020.610723

Tashiro, Y., Hirano, S., Matson, M. M., Atsumi, S., and Kondo, A. (2018). Electrical-biological Hybrid System for CO2 Reduction. *Metab. Eng.* 47, 211–218. doi:10.1016/j.ymben.2018.03.015

Ungerer, J., and Pakrasi, H. B. (2016). Cpf1 Is A Versatile Tool for CRISPR Genome Editing across Diverse Species of Cyanobacteria. *Sci. Rep.* 6, 1–9. doi:10.1038/srep39681

Ungerer, J., Lin, P.-C., Chen, H.-Y., and Pakrasi, H. B. (2018). Adjustments to Photosystem Stoichiometry and Electron Transfer Proteins Are Key to the Remarkably Fast Growth of the Cyanobacterium *Synechococcus Elongatus* UTEX 2973. *MBio* 9, e02327. doi:10.1128/mBio.02327-17

Wendt, K. E., Ungerer, J., Cobb, R. E., Zhao, H., and Pakrasi, H. B. (2016). CRISPR/Cas9 Mediated Targeted Mutagenesis of the Fast Growing Cyanobacterium *Synechococcus Elongatus* UTEX 2973. *Microb. Cel Fact.* 15, 1–8. doi:10.1186/s12934-016-0514-7

Yao, J., Wang, J., Ju, Y., Dong, Z., Song, X., Chen, L., et al. (2022). Engineering a Xylose-Utilizing *Synechococcus Elongatus* UTEX 2973 Chassis for 3-Hydroxypropionic Acid Biosynthesis under Photomixotrophic Conditions. *ACS Synth. Biol.* 11, 678–688. doi:10.1021/acssynbio.1c00364

Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., et al. (2015). Synechococcus Elongatus UTEX 2973, a Fast Growing Cyanobacterial Chassis for Biosynthesis Using Light and CO2. *Sci. Rep.* 5, 8132. doi:10.1038/srep08132

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Bioengineering and Biotechnology

BRIEF RESEARCH REPORT
published: 11 March 2022
doi: 10.3389/fbioe.2022.845240

NOT Gates Based on Protein Degradation as a Case Study for a New Modular Modeling via SBML Level 3—Comp Package

Biruck Woldai Abraha and Mario Andrea Marchisio *†

School of Pharmaceutical Science and Technology, Tianjin University, Tianjin, China

In 2008, we were among the first to propose a method for the visual design and modular modeling of synthetic gene circuits, mimicking the way electronic circuits are realized *in silico*. Basic components were DNA sequences that could be composed, first, into transcription units (TUs) and, then, circuits by exchanging fluxes of molecules, such as PoPS (polymerase per second) and RiPS (ribosomes per seconds) as suggested by Drew Endy. However, it became clear soon that such fluxes were not measurable, which highlighted the limit of using some concepts from electronics to represent biological systems. SBML Level 3 with the *comp* package permitted us to revise circuit modularity, especially for the modeling of eukaryotic networks. By using the libSBML Python API, TUs—rather than single parts—are encoded in SBML Level 3 files that contain species, reactions, and *ports*, i.e., the interfaces that permit to wire TUs into circuits. A circuit model consists of a collection of SBML Level 3 files associated with the different TUs plus a "main" file that delineates the circuit structure. Within this framework, there is no more need for any flux of molecules. Here, we present the SBML Level 3-based models and the wet-lab implementations of Boolean NOT gates that make use, in the yeast *Saccharomyces cerevisiae*, of the bacterial ClpX-ClpP system for protein degradation. This work is the starting point towards a new piece of software for the modular design of eukaryotic gene circuits and shows an alternative way to build genetic Boolean gates.

Keywords: SBML level 3, *S. cerevisiae*, ClpP-ClpX, protein degration, Boolean gates

OPEN ACCESS

Edited by:
Jean Marie François, Institut Biotechnologique de Toulouse (INSA), France

Reviewed by:
Monika Heiner, Brandenburg University of Technology Cottbus-Senftenberg, Germany
Lorenzo Pasotti, University of Pavia, Italy

***Correspondence:**
Mario Andrea Marchisio
mario@tju.edu.cn
mamarchisio@yahoo.com

†ORCID:
Mario Andrea Marchisio
orcid.org/0000-0002-5102-1069

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *29 December 2021*
Accepted: *21 February 2022*
Published: *11 March 2022*

Citation:
Abraha BW and Marchisio MA (2022) NOT Gates Based on Protein Degradation as a Case Study for a New Modular Modeling via SBML Level 3—Comp Package. Front. Bioeng. Biotechnol. 10:845240. doi: 10.3389/fbioe.2022.845240

INTRODUCTION

Like electronics, Synthetic Biology deals with circuits. Thus, over the last twenty years, efforts have been made to adapt electrical engineering concepts and methods to the modular design and modeling of circuits made of DNA. In his famous 2005 paper (Endy, 2005), Drew Endy depicted a possible correspondence between the electrical current—easily measurable with an ammeter—and biological currents that could be responsible for the working of synthetic gene circuits. Following the dogma of molecular biology, RNA polymerases and ribosomes, which lead the synthesis of mRNA and proteins, respectively, appeared to be the biological counterparts of the electrons. Moreover, basic circuit components (in eukaryotic cells) were identified with promoters, coding regions (CDSs), and terminators, i.e., DNA pieces with a well-defined function either in transcription or translation. As the electric current permits to wire together resistors, batteries, solenoids and all the other basic electric components, the fluxes

of RNA polymerase (PoPS: polymerases per second) and ribosomes (RiPS: ribosomes per second) should be the biological currents (a shared input/output) that permit to assemble together, first, biological parts into transcription units (TUs) and then TUs into circuits. Hence, biological fluxes permitted to define concepts such as *part composability* and *abstraction hierarchy* in Synthetic Biology.

Computational biologists liked these innovatory ideas because they indicated how to deal with biological circuits in a systematic way, whereas wet-lab biologists were skeptical since they were well-aware of the fact that a disordered motion like that of RNA polymerases and ribosomes could not be measured with any instruments. Therefore, PoPS/RiPS-based modeling was never completely accepted by the Synthetic Biology community and, in the end, it showed—in our opinion—the limits of using electronics concepts to study biological systems.

We were among the first to develop a piece of software (Marchisio and Stelling, 2008)—later named "Parts and Pools" (Marchisio, 2014b)—based on an extension of the ideas of Drew Endy. We pointed out that PoPS and RiPS allowed, indeed, parts' composability but were not enough to describe the wiring among transcription units, which is what really makes a circuit work. Therefore, we had to introduce other *signal carriers* (transcription factors, small RNAs, and chemicals) and their corresponding fluxes to mediate the interaction among TUs. Our piece of software was an add-on of ProMoT (process modeling tool) (Mirschel et al., 2009), a program for the visual, modular design of complex systems. ProMoT demands to program in MDL (model definition language), a Lisp-based language that permits to define modules that communicate via fluxes (of molecules, in biological systems) calculated and exchanged via the so-called module *terminals*. ProMoT provides a Graphical User Interface as well where modules are displayed in a drag-and-drop way and terminals are wired together.

Our ProMoT add-on is a collection of Perl and Python scripts each generating one or more MDL modules. The latest version (Marchisio et al., 2013) was successfully applied to the design and modeling of eukaryotic circuits. The usage of fluxes made it straightforward to build models based on mass-action kinetics. Moreover, eukaryotic promoters and coding regions can be characterized by a high number of species and reactions that are calculated by means of BioNeTGen (Blinov et al., 2004). The CDS script generates not only the MDL file for the CDS part, but also those corresponding to the pools of its mRNA and protein. Nucleus and cytoplasm are designed separately (due to their complexity) and then connected by running another script. Finally, a circuit is encoded by a collection of MDL files associated with parts and pools plus a main one that contains the connections among all the modules. ProMoT is able to export the circuit into the SBML2 format that permits circuit analysis and simulations—with, for instance, COPASI (Hoops et al., 2006) as we generally did—but loses the circuit modularity.

SBML Level 3 (Keating et al., 2020) with the hierarchical model composition package (*comp*) (Smith et al., 2015) allows for a direct modular modeling of genetic circuit without the need for the help of another language, such as MDL, nor the usage of fluxes of molecules to establish the wiring among the modules present in the circuit. Furthermore, SBML Level 3-comp (concisely SBML3-comp) files can be easily generated by means of the libSBML (Bornstein et al., 2008).

In this paper we describe the wet-lab implementation, in the yeast *Saccharomyces cerevisiae*, of two NOT gates based on protein degradation—a design strategy advanced 15 years ago (Grilly et al., 2007) but never extensively exploited nor optimized—and then we show how they can be modeled, in a modular way, by means of SBML3-comp. We will not explain all the libSBML commands required to generate the full circuit model (for this we refer the reader to our book chapter (Marchisio, 2021) and the files available at GitHub—see Data Availability below) but we will focus on the way to establish connections among circuit modules, which no longer demands to define fluxes of molecules that cannot be measured in the lab. SBML3-comp is already supported by several computational tools such as COPASI, iBioSim (Watanabe et al., 2019), Virtual Part Repository 2 (Misirli et al., 2021), Tellerium (Choi et al., 2018), and BMSS (Yeoh et al., 2019). For this reason, we think it will become the standard language to model genetic circuit in the very next future.

MATERIALS AND METHODS

Wet-Lab Experiments

Backbones for all plasmids constructed in this work were the yeast integrative shuttle-vector pRSII40X available at Addgene (a gift from Steven Haase) (Chee and Haase, 2012). The plasmids containing ClpP, the yeast codon-optimized version of ClpX, and yEGFP_ssrA were kindly provided by Jeff Hasty (University of California, San Diego, United States). Each new plasmid was assembled via the Gibson method (Gibson et al., 2009)—they are listed in **Supplementary Table S5**. The DNA sequences of the DNA parts used in this work are written in **Supplementary Material** as well. Plasmids were integrated into the genome of the yeast *S. cerevisiae* strain CEN.PK2-1C (MATa; his3D1; leu2-3_112; ura3-52; trp1-289; MAL2-8c; SUC2)—Euroscarf (Johann Wolfgang Goethe University, Frankfurt, Germany)—via the lithium-acetate protocol (Gietz and Woods, 2002).

Green fluorescence was measured with a BD FACSVerse™ Flow Cytometer (blue laser-488 nm, emission filter-527/32 nm). Each strain was measured in three replicas. During each experiment, 30,000 events were recorded. Data from the flow cytometer was analyzed with the flowcore R-Bioconductor package (Hahne et al., 2009).

The performance of a Boolean gate was characterized by the ON/OFF ratio, i.e., the the ratio between the maximal (ON) and minimal (OFF) fluorescence level expressed by the circuit. For many applications, an ON/OFF ratio bigger than or equal to 2 is enough to claim that a Boolean gate works properly (Yu and Marchisio, 2021; Zhang and Marchisio, 2022).

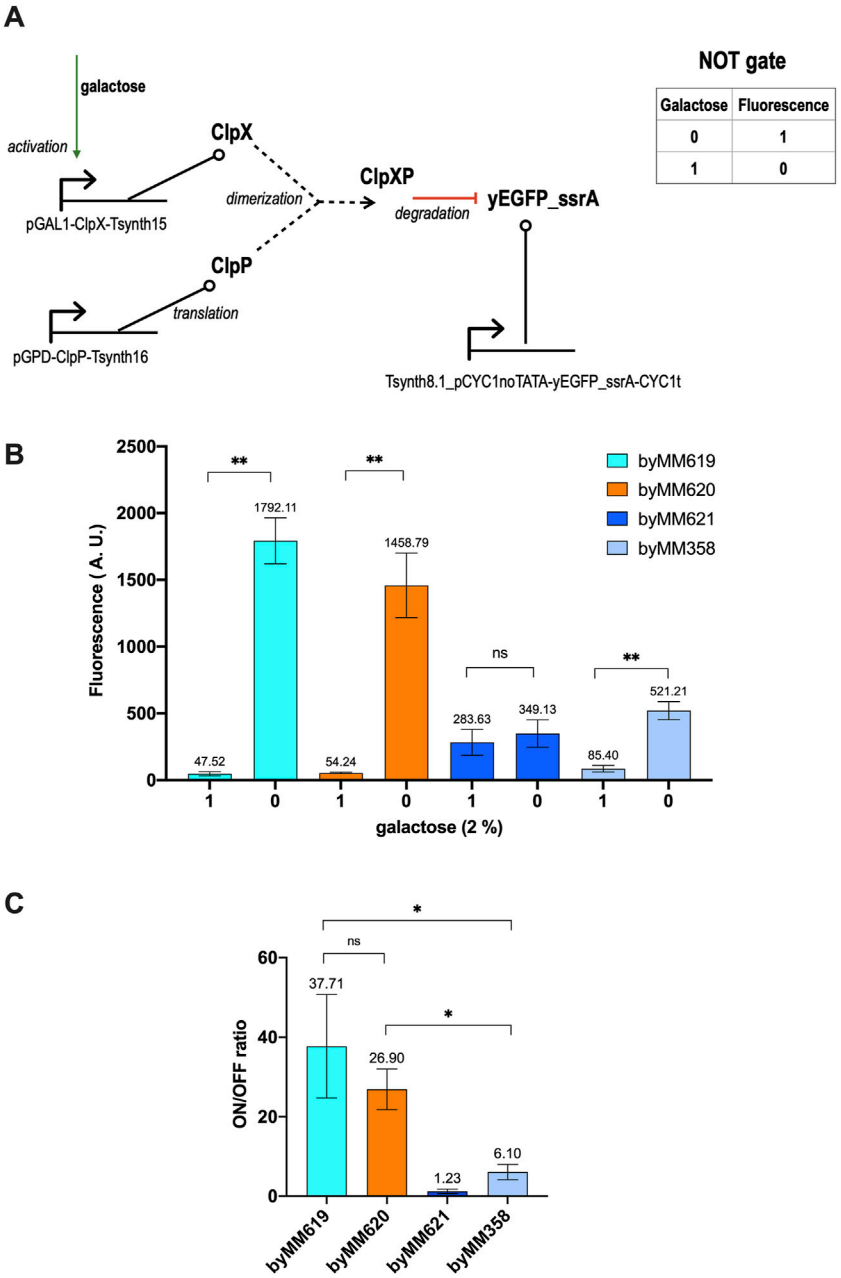


FIGURE 1 | Galactose-sensing NOT gate. **(A)** Circuit scheme. The reporter protein (yEGFP_ssrA) and ClpP are constitutively expressed by the rather weak synthetic promoter Tsynth8.1_pCYC1noTATA and the strong yeast *GPD* promoter, respectively. ClpX (a yeast codon optimized version) synthesis is switched ON only in the presence of galactose. Differently from the network in (Grilly et al., 2007), only the expression of one of the two components of the ClpXP dimer is controlled by an input. Moreover, the production of the tagged green fluorescent protein cannot be stopped by any external signal (see **Supplementary Figure S1** for a comparison). **(B)** Average fluorescence level (together with its standard deviation) corresponding to four strains hosting a NOT gate. byMM619 and byMM620 show the highest separation between the 0 and 1 output level. byMM621, in contrast, turned out to be unfunctional. byMM358 differs from the previous three gates for the synthetic promoter DEG1t_pCYC1noTATA that expresses yEGFP_ssrA instead of Tsynth8.1_pCYC1noTATA. DEG1t_pCYC1noTATA is 1.37-fold stronger than Tsynth8.1_pCYC1noTATA. **(C)** Gate performance in term of fluorescence gain. byMM619 and byMM620 turned out not to be significantly different under statistical analysis. Their mean gain is much higher than that of byMM358, which is nevertheless reasonably elevated for a logic circuit. Each experiment was made in three replicas. Data has been analyzed via two-sided Welch's *t*-test (*: *p*-value < 0.05; **: *p*-value < 0.01; ns: no significant difference).

Computational Experiments

Scripts to generate circuit modules were written in Python (version 3.8) and run on a MacBook Air (macOS Catalina 10.15.7, 1.6 GHz Dual-Core Intel Core i5, 8 GB RAM). Circuit simulations and analysis were carried out with COPASI (version 4.34, build 251).

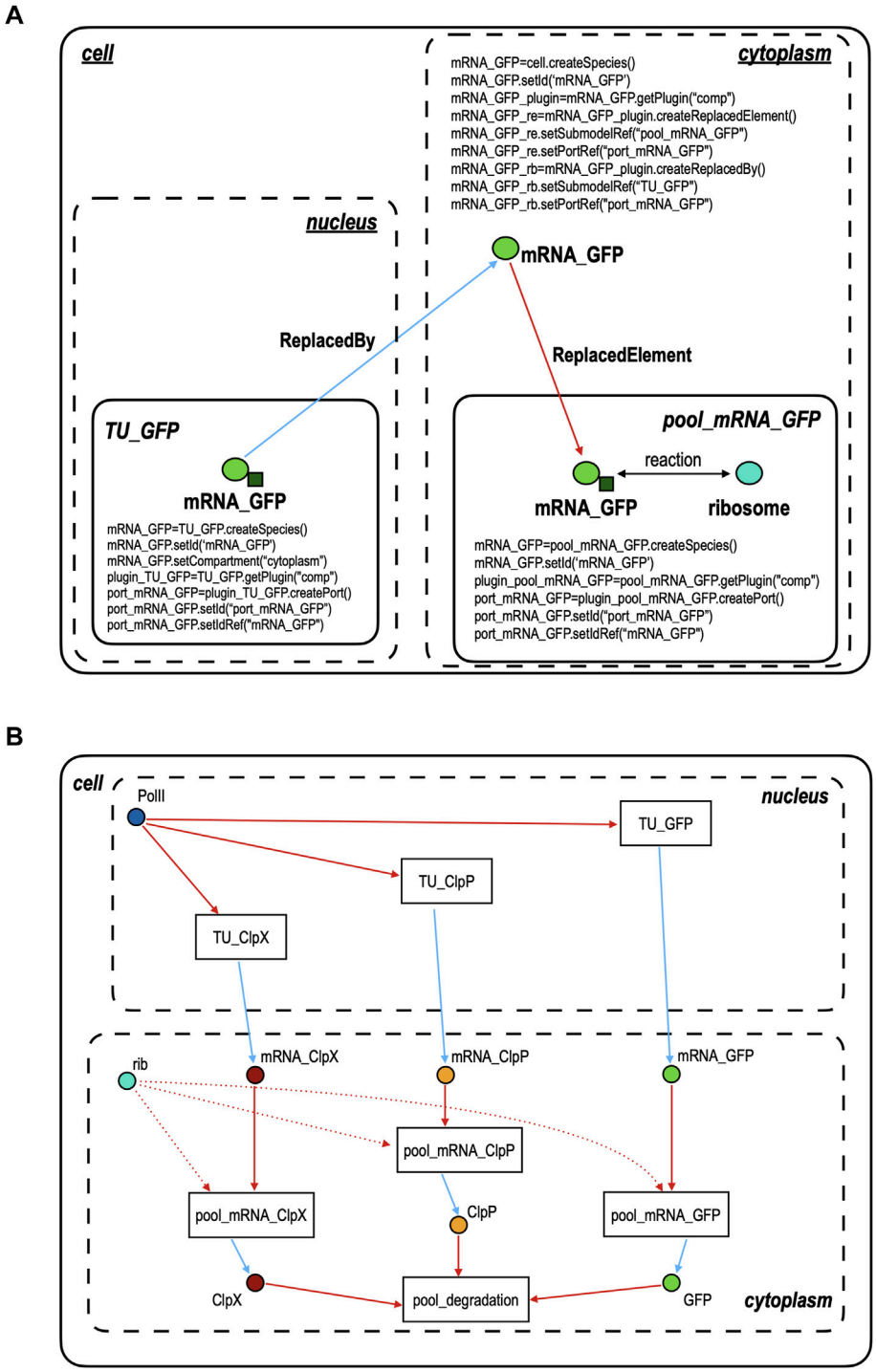


FIGURE 2 | Connecting modules in SBML3-comp. **(A)** Port objects (represented as small dark green squares) are instantiated in the TU_GFP and pool_mRNA_GFP modules and refer to the species mRNA_GFP (a light green circle)—a copy of which is present in each of the two modules. Ports permit to link the two modules indirectly, i.e., via a third copy of mRNA_GFP that lies in the cytoplasm (cell module) and is connected to the mRNA_GFP inside TU_GFP via a ReplacedBy object (cyan arrow) and to the mRNA_GFP inside pool_mRNA_GFP via a ReplacedElement object (red arrow). **(B)** Modular model of the galactose responding NOT gate shown in **Figure 1A**. Species, like mRNA_GFP, instantiated in the cell model are responsible for the connection between at least two modules. As in **Figure 2A**, all cyan arrows represent a ReplacedBy relation and the red arrows (straight or dotted) a ReplacedElement one. Galactose, the circuit input, is not shown since its presence or absence is determined by changing the value of the transcription initiation rate of pGAL1, which belongs to the module TU_ClpX. It should be noted that the circuit design still follows the idea to have two main kinds of module, like in our piece of software "Parts and Pools". However, by using SBML3-comp, some Pools become redundant and are replaced by single species, whereas Parts are no longer modeled separately because, as explained in the main text, TUs represent now the smallest circuit components. This assumption permits to remove several species (compounds) and reactions that were necessary to describe the interactions between RNA polymerases II and the DNA Parts—and also between the ribosomes and the mRNA segments.

RESULTS

In *E. coli*, the ClpX and ClpP proteins work in tandem (ClpXP) in order to carry out protein degradation. Upon recognition of a degron sequence, ClpX proceeds to unfold and translocate a protein into the ClpP proteolytic compartment with the release of small peptides (Baker and Sauer, 2012). This bacterial system was first shown to work efficiently in *S. cerevisiae* by Grilly and co-authors (Grilly et al., 2007)—and later re-proposed by Macia *et al.* (Macia et al., 2016)—who constructed a four-gene circuit where the yeast enhanced green fluorescent protein—yEGFP (Sheff and Thorn, 2004)—was fused to the ssrA degradation tag and the expression of both ClpX and ClpP was controlled by the LacI/IPTG system. Moreover, yEGFP was produced only in the presence of galactose in the cell culture (see **Supplementary Figure S1**). Taking a cue from this work, we made use of the ClpXP system to construct two different NOT gates, one responding to galactose, the other to beta-estradiol.

Galactose-Responsive NOT Gate

The galactose responsive NOT gate requires three transcription units only. We implemented two versions of it by changing the synthetic constitutive promoter upstream of *yEGFP_ssrA*: Tsynth8.1_pCYC1noTATA and DEG1t_pCYC1noTATA (Song et al., 2016). The former configuration returned the best results with an around 30-fold ON/OFF ratio (see **Figure 1**).

Modular Modeling With SBML3-Comp

The circuit scheme in **Figure 1A** shows only the DNA sequences required to build the gate and the proteins that make the circuit work. Several important features are, however, missing such as the cell compartments and the mRNA corresponding to each protein. By using SBML3-comp, we made a model for this NOT gate that consists of 16 species and 23 reactions distributed over 7 modules: 3 TUs (in the nucleus), 3 mRNA pools, and a *degradation* pool where ClpP and ClpX first dimerize and then degrade, quickly, yEGFP_ssrA. All pools lie in the cytoplasm since none of the proteins involved in the circuit interacts with the DNA.

Connecting Modules via Ports

The main feature we want to explain about modeling gene circuits in SBML3-comp is how to realize the connection between two (or more) modules without the need for fluxes. SBML3-comp permits to associate species with *ports*. They are objects that have the function to establish a link among *copies* of the same species that lie in different modules. However, ports cannot be connected directly but they need a "helper" species, i.e., a third copy of the same species that lies, however, in a *compartment* (or another module). This helper species does not have its own port but has the capability to finalize the junction between two ports. **Figure 2A** clarifies how ports, species, and modules work together by showing how the mRNA corresponding to yEGFP_ssrA (mRNA_GFP) is the species that permits to link the transcription unit encoding for yEGFP_ssrA in the nucleus (TU_GFP) to the pool, in the cytoplasm, where yEGFP_ssrA is synthesized (pool_mRNA_GFP). In this representation, we have three kinds of modules: the cell, which represents a *model* in SBML3-comp, TU_GFP and pool_mRNA_GFP, which are both *submodels* of the cell model, and the compartments (the nucleus and the cytoplasm). A model is a global container of submodels and compartments. TU_GFP is included in the nucleus and, from **Figure 2B**, we can see that it gets RNA polymerase II (PolII) as an input and delivers mRNA_GFP as an output. mRNA_GFP is created, as a species, in each of the three modules where it is contained, i.e., TU_GFP, the cytoplasm, and pool_mRNA_GFP. It should be noted, however, that the TU_GFP code contains the line "mRNA_GFP.setCompartment ("cytoplasm")" which means that, even though the module lies in the nucleus, its output will stay in the cytoplasm. Every module gets access to the package comp via the command "getPlugin ("comp")", that permits the creation of ports in the submodels and references to species in the main model. In our example, we have that the submodel TU_GFP instantiates a new "comp" object—here called "plugin_TU_GFP"—that allows the creation of a port object via the instruction port_mRNA_GFP = plugin_TU_GFP.createPort(). The new port is then given an ID (i.e. a name: port_mRNA_GFP.setId ("port_mRNA_GFP") and is associated with the mRNA_GFP species through the command: port_mRNA_GFP.setIdRef("mRNA_GFP"). By using the same syntax, a port named "port_mRNA_GFP" that refers to the species mRNA_GFP is generated also in the other submodel, pool_mRNA_GFP. These two ports are linked thanks to the mRNA_GFP species created in the cell model. Here, no new ports are made. mRNA_GFP invokes the getPlugin ("comp") command to establish two new objects. One belongs to the *ReplacedBy* class and sets up a bond with the TU_GFP submodel via the local port_mRNA_GFP. The other is a *ReplacedElement* object that "points" to the port_mRNA_GFP in the submodel pool_mRNA_GFP. In this way, a connection among the three mRNA_GFP species is established. As shown in **Figure 2B**, 5 more "helper" species are present in the cell model to realize connections between pairs of modules. The species "PolII" and "rib" do not belong to any submodel. Thus, our SBML3-comp model does not contain the RNA polymerase II and the ribosome pools. These species are linked to TUs and pool_mRNAs simply via ReplacedElement objects. Every submodel plus the cell model is saved as a separate xml file. In order to simulate the circuit with COPASI, it is enough to import the "cell.xml" file that calls all the other xml files and *flattens* the overall model such that, for instance, only one species called mRNA_GFP is present and lies in the cytoplasm.

Comparing Computational and Experimental Results

The mathematical model for the whole gate assumes that promoters and mRNAs are activated by RNA polymerases II and ribosomes, respectively, via a Hill function without cooperativity. The ClpXP complex is supposed to bind irreversibly to yEGFP_ssrA and induce its fast degradation. Parameters values were taken, where possible, from our previous works (Marchisio, 2014a; 2021), estimated from lab

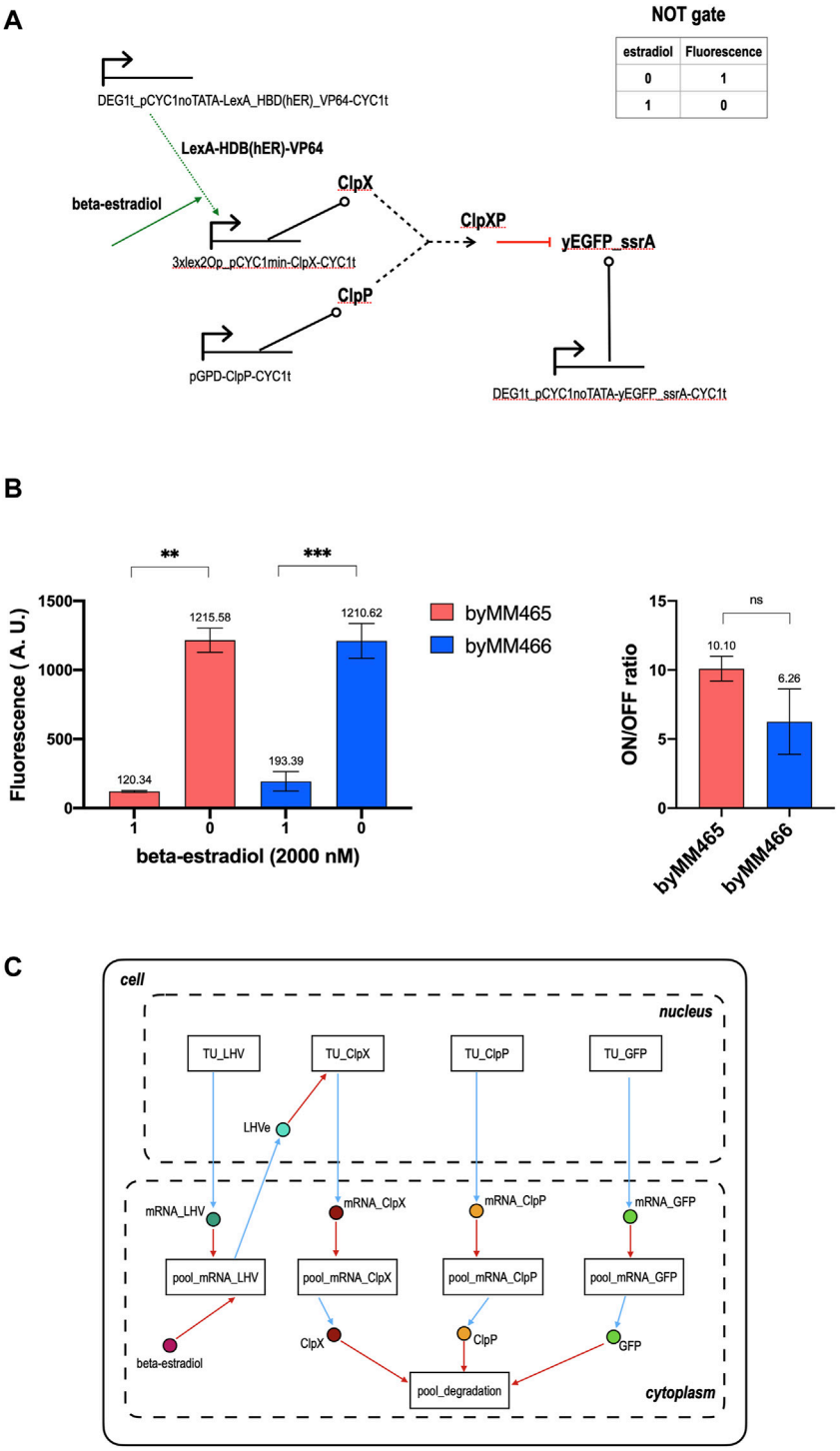


FIGURE 3 | NOT gate sensing beta-estradiol. **(A)** Circuit scheme. The synthesis of ClpX depends on the presence of the hormone beta-estradiol, the gate input. The additional TU with respect to the gate in **Figure 1** synthesizes, constitutively, a chimeric activator that, in our model, we called LHV since it consists of the bacterial protein LexA (acting as a DNA-binding domain), the hormone-binding domain of the human estrogen receptor (HBD(ER)), and the viral strong activation domain VP64. In the absence of beta-estradiol, LHV is kept in the cytoplasm because of the action of Hsp90 that binds HBD(hER). As a consequence, the synthetic activated promoter upstream of ClpX, which is made of three copies of lex2Op (the LexA operator) in front of a minimal weak *CYC1* promoter (Ottoz et al., 2014), is incapable of producing ClpX in high quantity such that the gate fluorescent output is high. A concentration of 1-2 μM beta-estradiol, which bind HBD (hER), nullifies the effect of Hsp90 and lets the complex LHVe (LexA_HBD(hER)_VP64-beta-estradiol) get into the nucleus and enhance the synthesis of ClpX, upon binding the three lex2Ops upstream of pCYC1min. Under these conditions, ClpX is highly expressed and, as a consequence, the fluorescence emitted by the gate drops to a very low level. **(B)**

(*Continued*)

FIGURE 3 | Mean fluorescence level and ON/OFF ratio of our two NOT gate implementations responsive to beta-estradiol. Every value corresponds to the mean of three replicas. The standard deviation of the mean is shown as well. Data has been analyzed via two-sided Welch's *t*-test (**: *p*-value < 0.01; ***: *p*-value < 0.001; ns: no significant difference). **(C)** NOT gate modular modeling. The gate input is now a species in the cell cytoplasm connected, via a ReplacedElement object, to the pool_mRNA_LHV, where it binds the chimeric activator LHV and forms the complex LHVe. This species lies in the nucleus and permits a direct connection between pool_mRNA_LHV and TU_ClpX. These two modules contain all the reactions where LHVe is involved in.

measurements, and optimized with COPASI to reach an ON/OFF ratio equal to 29.56, i.e. in reasonable agreement with our experimental results. A detailed description of the model of the galactose-sensing NOT gate is given in the **Supplementary Model 1**.

Beta-Estradiol Responsive NOT Gate

With respect to the galactose-sensing NOT gate, the scheme of that responding to beta-estradiol demands one more TU encoding for a chimeric activator that interacts with the circuit input (see **Figure 3A**). In the absence of beta-estradiol, ClpX is expressed at a very low quantity and, consequently, the fluorescence level of the circuit is high. In contrast, concentrations of beta-estradiol between 1 and 2 μM trigger an elevated production of ClpX that induces a drop in the output signal. Our two implementations gave an ON/OFF ratio equal to 10.10 and 6.26. However, they did not show any significant statistical difference (see **Figure 3B** and **Supplementary Figures S1, S2** for the corresponding titration curves). We modeled this circuit in a slightly different way with respect to the previous one, i.e. we neglected the presence of RNA polymerase II and ribosomes in the cells—a commonly used simplification that fully excludes fluxes such as PoPS and RiPS. Overall, we needed 9 modules (4 TUs in the nucleus, 4 mRNA pools and one degradation pool in the cytoplasm) and as many "connecting" species instantiated in the cell model: 8 in the cytoplasm and only one in the nucleus (see **Figure 3C**). On the whole, the model consists of 21 species and 35 reactions. After parameter optimization, we got a simulated ON/OFF ratio equal to 13.64, not too distant from the measured one (more details are given in **Supplementary Model 2**).

DISCUSSION

SBML3-comp represents a huge improvement of SBML2 since it permits to construct models of biological systems and, therefore, of synthetic gene circuits, in a modular way without being assisted by other programs. libSBML permits to divide the model of a circuit in different scripts, each generating one or more modules. We think that the best way is to treat transcription units, rather than single DNA parts, as basic modules. A TU can be encoded in the same file together with its corresponding pool of mRNA and, if required, that of the protein it encodes for. The direct use of SBML3 allows also to choose easily the best kinetics for every reaction, without the constraint of using the same in the whole circuit.

In this work we have shown how Boolean gates characterized by high ON/OFF ratio can be built, in *S. cerevisiae*, by using an almost forgotten, though powerful, yeast-orthogonal system, i.e., the bacterial ClpXP dimer to induce protein degradation. Our Boolean gates served as a case study to explain how to use SBML3-comp to model genetic circuits in a new modular way. Importantly, it was apparent how Python scripts, which generate the SBML3-comp files, can be recycled for the same TUs inside different circuits or just slightly modified to construct models for similar TUs.

We think that the joint use of the System Biology Open Language—SBOL (Galdzicki et al., 2014) and SBML3-comp seems to be the best way to combine, in the future, a detailed description of the actual DNA circuit components (sequences) with modular modeling (Watanabe et al., 2019; Misirli et al., 2021).

DATA AVAILABILITY STATEMENT

FACS data (fcs files) relating to the results presented in this work are accessible at FlowRepository (http://flowrepository.org). Python files to generate the model of the two different NOT gates, together with COPASI files containing the optimized parameter values, are publicly available at GitHub (https://github.com/mamarchisio/ClpXP-Python-COPASI).

AUTHOR CONTRIBUTIONS

BWA: circuit construction, data collection and analysis, manuscript writing. MAM: conceptualization, modeling, manuscript writing.

ACKNOWLEDGMENTS

We want to thank all the students of the Synthetic Biology lab-SPST, TJU-for their general help together with Zhi Li and Xiangyang Zhang for their assistance at the FACS machine. Moreover, we shall thank Wenjiang Song and Wen Chang for their preliminary tests at the Harbin Institute of Technology, School of Life Science and Technology. Finally, MAM wants to thank his former advisor Joerg Stelling and his mates at the ETHZ, D-BSSE yeast lab: Fabian Rudolf, Robert Gnuegge, Kristina Elfstroem, and Diana S. M. Ottoz for valuable discussions and suggestions about gene Boolean gates based on protein degradation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.845240/full#supplementary-material

REFERENCES

Baker, T. A., and Sauer, R. T. (2012). ClpXP, an ATP-Powered Unfolding and Protein-Degradation Machine. *Biochim. Biophys. Acta (Bba) - Mol. Cel Res.* 1823 (1), 15–28. doi:10.1016/j.bbamcr.2011.06.007

Blinov, M. L., Faeder, J. R., Goldstein, B., and Hlavacek, W. S. (2004). BioNetGen: Software for Rule-Based Modeling of Signal Transduction Based on the Interactions of Molecular Domains. *Bioinformatics* 20 (17), 3289–3291. doi:10.1093/bioinformatics/bth378

Bornstein, B. J., Keating, S. M., Jouraku, A., and Hucka, M. (2008). LibSBML: an API Library for SBML. *Bioinformatics* 24 (6), 880–881. doi:10.1093/bioinformatics/btn051

Chee, M. K., and Haase, S. B. (2012). New and Redesigned pRS Plasmid Shuttle Vectors for Genetic Manipulation of *Saccharomyces cerevisiae*. *G3 (Bethesda)* 2 (5), 515–526. doi:10.1534/g3.111.001917

Choi, K., Medley, J. K., König, M., Stocking, K., Smith, L., Gu, S., et al. (2018). Tellurium: An Extensible Python-Based Modeling Environment for Systems and Synthetic Biology. *Biosystems* 171, 74–79. doi:10.1016/j.biosystems.2018.07.006

Daniel Gietz, R., and Woods, R. A. (2002). Transformation of Yeast by Lithium Acetate/single-Stranded Carrier DNA/polyethylene Glycol Method. *Methods Enzymol.* 350 (350), 87–96. doi:10.1016/s0076-6879(02)50957-5

Endy, D. (2005). Foundations for Engineering Biology. *Nature* 438 (7067), 449–453. doi:10.1038/nature04342

Galdzicki, M., Clancy, K. P., Oberortner, E., Pocock, M., Quinn, J. Y., Rodriguez, C. A., et al. (2014). The Synthetic Biology Open Language (SBOL) Provides a Community Standard for Communicating Designs in Synthetic Biology. *Nat. Biotechnol.* 32 (6), 545–550. doi:10.1038/nbt.2891

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009). Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* 6 (5), 343–345. doi:10.1038/nmeth.1318

Grilly, C., Stricker, J., Pang, W. L., Bennett, M. R., and Hasty, J. (2007). A Synthetic Gene Network for Tuning Protein Degradation in *Saccharomyces cerevisiae*. *Mol. Syst. Biol.* 3, 127. doi:10.1038/msb4100168

Hahne, F., LeMeur, N., Brinkman, R. R., Ellis, B., Haaland, P., Sarkar, D., et al. (2009). flowCore: a Bioconductor Package for High Throughput Flow Cytometry. *BMC Bioinformatics* 10, 106. doi:10.1186/1471-2105-10-106

Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., et al. (2006). COPASI--a COmplex PAthway SImulator. *Bioinformatics* 22 (24), 3067–3074. doi:10.1093/bioinformatics/btl485

Keating, S. M., Waltemath, D., König, M., Zhang, F., Dräger, A., Chaouiya, C., et al. (2020). SBML Level 3: an Extensible Format for the Exchange and Reuse of Biological Models. *Mol. Syst. Biol.* 16 (8), e9110. doi:10.15252/msb.20199110

Macia, J., Manzoni, R., Conde, N., Urrios, A., de Nadal, E., Solé, R., et al. (2016). Implementation of Complex Biological Logic Circuits Using Spatially Distributed Multicellular Consortia. *Plos Comput. Biol.* 12 (2), e1004685. doi:10.1371/journal.pcbi.1004685

Marchisio, M. A. (2014a). In Silico design and *In Vivo* Implementation of Yeast Gene Boolean gates. *J. Biol. Eng.* 8 (1), 6. doi:10.1186/1754-1611-8-6

Marchisio, M. A. (2021). Modular Modeling of Genetic Circuits in SBML Level 3. *Methods Mol. Biol.* 2189, 45–63. doi:10.1007/978-1-0716-0822-7_4

Marchisio, M. A. (2014b). Parts & Pools: A Framework for Modular Design of Synthetic Gene Circuits. *Front. Bioeng. Biotechnol.* 2, 42. doi:10.3389/fbioe.2014.00042

Marchisio, M. A., and Stelling, J. (2008). Computational Design of Synthetic Gene Circuits with Composable Parts. *Bioinformatics* 24 (17), 1903–1910. doi:10.1093/bioinformatics/btn330

Marchisio, M., Colaiacovo, M., Whitehead, E., and Stelling, J. (2013). Modular, Rule-Based Modeling for the Design of Eukaryotic Synthetic Gene Circuits. *BMC Syst. Biol.* 7, 42. doi:10.1186/1752-0509-7-42

Mirschel, S., Steinmetz, K., Rempel, M., Ginkel, M., and Gilles, E. D. (2009). PROMOT: Modular Modeling for Systems Biology. *Bioinformatics* 25 (5), 687–689. doi:10.1093/bioinformatics/btp029

Mısırlı, G., Yang, B., James, K., and Wipat, A. (2021). Virtual Parts Repository 2: Model-Driven Design of Genetic Regulatory Circuits. *ACS Synth. Biol.* 10 (12), 3304–3315. doi:10.1021/acssynbio.1c00157

Ottoz, D. S. M., Rudolf, F., and Stelling, J. (2014). Inducible, Tightly Regulated and Growth Condition-independent Transcription Factor in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 42 (17), e130. doi:10.1093/nar/gku616

Sheff, M. A., and Thorn, K. S. (2004). Optimized Cassettes for Fluorescent Protein Tagging inSaccharomyces Cerevisiae. *Yeast* 21 (8), 661–670. doi:10.1002/yea.1130

Smith, L. P., Hucka, M., Hoops, S., Finney, A., Ginkel, M., Myers, C. J., et al. (2015). SBML Level 3 Package: Hierarchical Model Composition, Version 1 Release 3. *J. Integr. Bioinform* 12 (2), 268. doi:10.2390/biecoll-jib-2015-268

Song, W., Li, J., Liang, Q., and Marchisio, M. A. (2016). Can Terminators Be Used as Insulators into Yeast Synthetic Gene Circuits? *J. Biol. Eng.* 10, 19. doi:10.1186/s13036-016-0040-5

Watanabe, L., Nguyen, T., Zhang, M., Zundel, Z., Zhang, Z., Madsen, C., et al. (2019). iBioSim 3: A Tool for Model-Based Genetic Circuit Design. *ACS Synth. Biol.* 8 (7), 1560–1563. doi:10.1021/acssynbio.8b00078

Yeoh, J. W., Ng, K. B. I., Teh, A. Y., Zhang, J., Chee, W. K. D., and Poh, C. L. (2019). An Automated Biomodel Selection System (BMSS) for Gene Circuit Designs. *ACS Synth. Biol.* 8 (7), 1484–1497. doi:10.1021/acssynbio.8b00523

Yu, L., and Marchisio, M. A. (2021). *Saccharomyces cerevisiae* Synthetic Transcriptional Networks Harnessing dCas12a and Type V-A Anti-CRISPR Proteins. *ACS Synth. Biol.* 10 (4), 870–883. doi:10.1021/acssynbio.1c00006

Zhang, Y., and Marchisio, M. A. (2022). Interaction of Bare dSpCas9, Scaffold gRNA, and Type II Anti-CRISPR Proteins Highly Favors the Control of Gene Expression in the Yeast *S. cerevisiae*. *ACS Synth. Biol.* 11 (1), 176–190. doi:10.1021/acssynbio.1c00352

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Bioengineering and Biotechnology

REVIEW
published: 04 May 2022
doi: 10.3389/fbioe.2022.869206

Synthetic Biology Tools for Engineering Microbial Cells to Fight Superbugs

Angel León-Buitimea[1,2], *Francisco de Jesús Balderas-Cisneros*[1,2], *César Rodolfo Garza-Cárdenas*[1,2], *Javier Alberto Garza-Cervantes*[1,2] and *José Rubén Morones-Ramírez*[1,2]*

[1]Facultad de Ciencias Químicas, Universidad Autónoma de Nuevo León (UANL), San Nicolás de los Garza, Mexico, [2]Centro de Investigación en Biotecnología y Nanotecnología, Facultad de Ciencias Químicas, Parque de Investigación e Innovación Tecnológica, Universidad Autónoma de Nuevo León, Apodaca, Mexico

With the increase in clinical cases of bacterial infections with multiple antibiotic resistance, the world has entered a health crisis. Overuse, inappropriate prescribing, and lack of innovation of antibiotics have contributed to the surge of microorganisms that can overcome traditional antimicrobial treatments. In 2017, the World Health Organization published a list of pathogenic bacteria, including *Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii*, *Pseudomonas aeruginosa,* and *Escherichia coli* (ESKAPE). These bacteria can adapt to multiple antibiotics and transfer their resistance to other organisms; therefore, studies to find new therapeutic strategies are needed. One of these strategies is synthetic biology geared toward developing new antimicrobial therapies. Synthetic biology is founded on a solid and well-established theoretical framework that provides tools for conceptualizing, designing, and constructing synthetic biological systems. Recent developments in synthetic biology provide tools for engineering synthetic control systems in microbial cells. Applying protein engineering, DNA synthesis, and *in silico* design allows building metabolic pathways and biological circuits to control cellular behavior. Thus, synthetic biology advances have permitted the construction of communication systems between microorganisms where exogenous molecules can control specific population behaviors, induce intracellular signaling, and establish co-dependent networks of microorganisms.

Keywords: synthetic biology, antimicrobial resistance, genetic circuits, antibiotics, phages, whole-cell engineering

OPEN ACCESS

Edited by:
Jean Marie François, Institut Biotechnologique de Toulouse (INSA), France

Reviewed by:
Ho Chun Loong, Southern University of Science and Technology, China
Jingyu Chen, China Agricultural University, China

***Correspondence:**
José Rubén Morones-Ramírez
jose.moronesrmr@uanl.edu.mx

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *04 February 2022*
Accepted: *18 April 2022*
Published: *04 May 2022*

Citation:
León-Buitimea A, Balderas-Cisneros FdJ, Garza-Cárdenas CR, Garza-Cervantes JA and Morones-Ramírez JR (2022) Synthetic Biology Tools for Engineering Microbial Cells to Fight Superbugs. Front. Bioeng. Biotechnol. 10:869206. doi: 10.3389/fbioe.2022.869206

INTRODUCTION

With the increase in clinical cases of bacterial infections with multiple antibiotic resistance, the world has entered a health crisis. Overuse, inappropriate prescribing, and lack of innovation of antibiotics have contributed to the surge of microorganisms that can overcome traditional antimicrobial treatments (Ventola, 2015). In 2017, the World Health Organization published a list of pathogenic bacteria, including *Enterococcus faecium* (*E. faecium*), *Staphylococcus aureus* (*S. aureus*), *Klebsiella pneumoniae* (*K. pneumoniae*), *Acinetobacter baumannii* (*A. baumannii*), *Pseudomonas aeruginosa* (*P. aeruginosa*), and *Escherichia coli* (*E. coli*) (ESKAPE) (World Health Organization, 2017). These bacteria have the ability to adapt to multiple antibiotics and transfer their resistance to other organisms; therefore, studies to find new

FIGURE 1 | Recent works have demonstrated the impact of synthetic biology and the tools and techniques that the field has developed to allow genetically modifying antibiotics, enhancing antibiotic production, developing engineered phages, and designing microbial control systems.

therapeutic strategies are needed. One of these strategies is synthetic biology geared toward developing new antimicrobial therapies.

Synthetic biology is founded on a solid and well-established theoretical framework that provides tools for conceptualizing, designing, and constructing synthetic biological systems (Perrino et al., 2021). Recent developments in synthetic biology provide tools for engineering synthetic control systems in microbial cells. Applying protein engineering, DNA synthesis, and *in silico* design allows building metabolic pathways and biological circuits to control cellular behavior (McCarty and Ledesma-Amaro, 2019). Thus, synthetic biology advances have permitted the construction of communication systems between organisms where exogenous molecules can control specific population behaviors, induce intracellular signaling, and establish co-dependent networks of microorganisms (Hennig et al., 2015; McCarty and Ledesma-Amaro, 2019).

The applications of synthetic biology systems (artificial circuits and functions within biological systems) include producing biologically based products in agriculture, industry, environmental, and healthcare studies. In this review, we highlight recent works of the impact of synthetic biology on genetic engineering to modify antibiotics and enhance antibiotic production, engineered phages, microbial control systems as an alternative to fight antibiotic resistance, and some other synthetic biology tools to engineer microbial communities (**Figure 1**).

GENETIC ENGINEERING TO MODIFY ANTIBIOTICS

Organisms are naturally capable of producing metabolites with antimicrobial characteristics. Unfortunately, "wild-type" antibiotic producers have poor production and low titers; however, with the unravel of novel biosynthesis mechanisms of that produce antibiotics and the tools provided by synthetic biology, it is possible to engineer the capacities of organisms to over-produce and diversify these metabolites (Breitling and Takano, 2015; El Karoui et al., 2019; Zhang et al., 2020).

One of the most interesting antibiotic synthesis mechanism is the multimodal enzymatic complex presented in biosynthetic cluster genes (BCG). It is composed of non-ribosomal peptide synthetases (NRPS), polyketide synthases (PKS), and combinations. NRPS and PKS are multimodal enzymatic complexes that are responsible for assembling non-ribosomal peptides (NRP) and polyketides, respectively, and giving them an active chemical structure (Martínez-Núñez and López, 2016; Nivina et al., 2019; Hwang et al., 2020). With the tools provided by synthetic biology and applying some engineering to these enzyme complexes, it is possible to attack the problem of low production and limited diversity of antibiotics.

As a first approach to tackle low production and diversity, antimicrobial compounds members of the NRP group have been addressed especially the family of antibiotics known as glycopeptides (Yim et al., 2014). Some authors had focused their efforts on diversifying these antibiotics, such as Yim et al., who used combinations of 13 scaffold-modifying enzymes from 7 GPA BCGs to be introduced into *Streptomyces coelicolor* (*S. coelicolor*). Among these combinations, nine new compounds were reported. Interestingly, eight of those compounds had antimicrobial activity against vancomycin-resistant *Enterococcus faecalis* (*E. faecalis*), exhibiting a MIC between 0.5–4 μg/ml (Yim et al., 2016). One of the promising newcomers in the glycopeptide family is corbomycin. Despite its outstanding clinical performance, natural production with *Streptomyces sp.* has poor performance (Culp et al., 2020). Xu et al. used a glycopeptide antibiotic heterologous expression system (GPAHex) to enhance the expression of genes for corbomicyn synthesis in *S. coelicolor*, obtaining a 19-fold increase in titers using this platform (Xu et al., 2020).

It is also possible to find that combinations between NRPS and PKS modules can give rise to more classes of antibiotics, such as lipopeptides. Daptomycin is a clinically significant lipopeptide antibiotic used primarily against methicillin-resistant *Staphylococcus aureus* (MRSA), naturally produced in *Streptomyces roseosporus* (*S. roseosporus*). Unfortunately, similar to other antibiotic natural producers, "wild-type" *S. roseosporus* has low daptomycin titers (Ye et al., 2014). Recently, Ji et al., using a "top-down" synthetic biology approach, achieved an increase in total lipopeptide production up to ~2,300%, where up to 40% was daptomycin (Ji et al., 2022).

Not only do combinations between NRPS and PKS give rise to the production of antibiotics, in some microorganisms, these enzyme complexes are joined by other families of enzymes such as fatty acid synthases (FAS). Initially discovered in *Dickeya zeae*, the new family of zeamines is one example of naturally occurring antibiotics resulting from the interaction between NRPS, PKS, and FAS (Wu et al., 2010). Zeamines drew the scientific community's attention for their potent microbicidal activity

against Gram-positive and Gram-negative bacteria. Unfortunately, *Dickeya zeae* has poor production of these antibiotics (Liao et al., 2014). Therefore, different authors, such as Masschelein et al., have studied other natural producers of zeamines, such as *Serratia plymuthica* (*S. plymuthica*) RVH1. Using in-frame deletion of biosynthetic genes, some of the mechanisms involved in synthesizing these antibiotics have been uncovered. This has been an outstanding contribution to synthetic biology for future combinatorial biosynthesis and bioengineering to produce new antibacterial compounds (Masschelein et al., 2013; Masschelein et al., 2015).

In addition to the approaches made by the combination of enzyme complexes, metabolic engineering offers another perspective to improve and diversify the production of antibiotics. Metabolic engineering aims to understand the networks of cellular metabolism and redesign them to improve productive capacities (Olano et al., 2008; Kim et al., 2016; Palazzotto et al., 2019). The efforts provided by metabolic engineering usually focus on redirecting metabolic fluxes, regulating BCGs and enzymes causing bottlenecks.

Discovering the relationships between intermediate metabolites and the biosynthesis of antibiotics have allowed the scientific community to improve yields in antibiotic production. An example of this is the regulation of the metabolite S-Adenosylmethionine (SAM). Although SAM is an essential precursor in methylation processes, this compound is involved in antibiotic biosynthesis in different species (Huh et al., 2004; Wang et al., 2007). Authors, such as Cai et al., manage to increase the yields in bacitracin production in *Bacillus licheniformis* (*B. licheniformis*), analyzing the SAM synthetic pathway. Authors, such as Cai et al., through analyzing the SAM synthetic pathway, have managed to increase the yields in bacitracin production up to 28.97% in *B. licheniformis* with a combination of different synthetic biology techniques such as heterologous expression, deletion, and overexpression of different genes (Cai et al., 2019; Cai et al., 2020).

Up to this point, the described studies have shown that the elucidation of the different mechanisms involved in the biosynthesis of antibiotics and the tools provided by synthetic biology have allowed the redesign of organisms by engineering them to improve their productive capacities and thus contributing to the fight against antimicrobial resistance. Although there are still some critical challenges for the continuous application of synthetic biology strategies to diversify antibiotics, the next few years promise to be rewarding for discovering new antibiotic compounds.

GENETIC ENGINEERING TO ENHANCE ANTIBIOTIC PRODUCTION

Different microorganisms are sources of other compounds used as antimicrobial agents. A variety of genetic material codes those metabolites and accessing this genetic diversity could increase the possibility of finding new or better ways to fight resistant microorganisms. Genetic techniques like gene mutation and the ability to control cellular functions at the genetic level could improve antimicrobial biosynthesis (Miao et al., 2006; Ma et al., 2017). Arafat et al. (2021) observed changes in antibiotic production, exposing *Streptomyces graminofaciens* (*S. graminofaciens*) to UV light. Studying the cellular function of various genes have allowed researchers to produce higher amounts and even better antibiotics. Studies of the gene cluster of the nucleoside antibiotic A201A in *Marinactinospora thermotolerans* (*M. thermotolerans*) SCSIO 00652 (Zhu et al., 2012) have led to improving its production; this has also been accomplished through genetic modification for the biosynthesis of amphotericin analogs in *Streptomyces nodosus* (*S. nodosus*) by disruption of *amphDIII* and *amphL* (Byrne et al., 2003). Makitrynskyy developed a manipulation of *AdpA* regions of *Streptomyces ghangensis* (*S. ghangensis*) (Makitrynskyy et al., 2021), giving researchers the understanding of a metabolic pathway to improve a specific antimicrobial biosynthesis. By understanding metabolic pathways and applying different gene engineering techniques, research can improve antimicrobial production through punctual modifications to regulate specific genes, demonstrating that genetic modifications to control biosynthesis pathways could improve antimicrobial production and its activities (Song et al., 2017; Li D. et al., 2021; Li Y.-P. et al., 2021).

Similarly, intended gene modifications can improve antimicrobial production, improve the metabolic flux to precursor availability, and enhance biosynthesis (Meng et al., 2017; Moosmann et al., 2020). These modifications could increase the concentration of antimicrobial precursors, such as (Shomar et al., 2018) expressing the gene cluster for carbapenem in *E. coli* producing antibiotics with a 60-fold increase. Also, modifying metabolic pathways, like carbon flux from the pentose phosphate pathway (PPP) to glycolysis by modifying *zwf1* and *zwf2* on *Streptomyces lividans* (*S. lividans*) (Butler et al., 2002) have led to increased glycolysis intermediates needed for antibiotic production.

A better bio-factory could be achieved by introducing genes in or from a different host taking advantage of its synthesis route (Han et al., 2012; Sakai et al., 2012; Makitrynskyy et al., 2021). Chen *et al.* (Chen et al., 2009) studied the expression of a gene cluster of *Streptomyces cacaoi* (*S. cacaoi*) in *S. lividans* TK24. However, polyoxin production by *S. lividans* TK24 was entirely the polyoxin H derivate due to the lack of genes for hydroxylation/carboxylation, leading to the formation of polyoxin A/F derivates. Eustáquio *et al.* (Eustáquio et al., 2004) expressed the gene cluster to synthesize novobiocin in *S. coelicolor*. Two *S. coelicolor* mutants were obtained by modifying the *novO* gene, an 8′-unsubstituted novobiocin by inactivation of *novO*, and a chlorine atom at C-8′ by expressing the *clo-hal* gene from the clorobiocin gene cluster. Wang *et al.* (Wang et al., 2021) expressed CYP genes of the mushroom *Ganoderma lucidum* in *Saccharomyces cerevisiae* (*S. cerevisiae*) to produce a derivate of 3,28-dihydroxy-lanosta-8,24-dien-26-oic acid. By improving the copy number of two resistance plasmids, the production of this new antibiotic was increased 8.2 folds compared to the control strain. The expression of non-ribosomal peptide synthetases in *Bacillus subtilis* (*B. subtilis*) was studied by Eppelmann *et al.* (Eppelmann et al., 2001) by introducing the bacitracin

biosynthesis gene cluster of *B. licheniformis*. *B. subtilis* showed comparable self-resistance to bacitracin due to the gene replacement and increased bacitracin A production due to the high-level expression of the bacitracin synthetase and the higher growth rate compared to *B. licheniformis*. Similarly, Wu *et al.* (Wu et al., 2015) expressed the gene *ram29* of the ramoplanin producer *Actinoplanes* sp. into *Streptomyces fungicidius* (*S. fungicidius*)to produce monomannosylated enduracin derivates.

The main goal of creating an engineered microorganism is the increased production of a specific antimicrobial. Nevertheless, when genetic engineering is focused on manipulating foreign genes in an organism, various reasons could affect its biosynthesis. Modifying genes of a particular biosynthesis or metabolic pathway may be enhanced in production but could decrease its antimicrobial activity. Even in the same genre, the differences between strains give the possibility of obtaining derivates of the intended antimicrobial, which could also modify its action. Thus, knowing how a microorganism's genetic material helps it survive in the microbiome can help us acquire the tools to manipulate it and use it to our advantage, increasing the possibility of obtaining new and better antibiotics.

ENGINEERED PHAGES TO FIGHT SUPERBUGS

Phages are viruses that infect specific bacteria and are the most abundant organisms on Earth (up to 2.5×10^8 phages per mL in water) (Bergh et al., 1989). Phages need bacteria to replicate and survive; thus, naturally, they work as bacteria controllers. Phages were discovered at the beginning of the XX century by Frederick Tort and Felix d'Herelle (Sharma et al., 2017), but antibiotics rapidly replaced phages in treating microorganism infections. However, due to the emergence of multi-antibiotic resistant bacterial pathogens, phages are considered an alternative way to treat multi-resistant bacteria.

Nowadays, researchers are studying single phages, discovering new phages, and designing phage cocktails as therapies to treat multi-resistant bacterial infections. Natural phages could be enough to fight superbugs, just like the works reporting on the administration of phage cocktails to patients infected by a multi-drug resistant *A. baumannii* strain, with results showing complete recovery after phage therapy (Schooley et al., 2017). Nevertheless, synthetic biology can improve or extend phage's abilities to generate variants with unique properties (Yacoby et al., 2007; Edgar et al., 2012). One of the strategies is to design and build phages with a broad host range (Lin et al., 2012). Phages provide a highly specific target of bacterial strains; thus, cocktails are required to fight infections and reduce the possibility that bacterium acquires resistance to any phage. Yehl et al. (2019) have developed a high-throughput strategy to engineer host-range-determining regions (HRDRs) in T3 phage by site-directed mutagenesis. Inspired by antibody specify engineering, this approach reduces disruptions in tail structure, and they call it "phage bodies".

Following this strategy could reduce the number of phages in cocktails. Using the monophages approach reduces the preparation and purification efforts and there is less potential for complications derived from using phages cocktails, such as phage purification, phage compatibility, making it easier to use as treatment. In addition, cocktails could make bacteria develop "broad-spectrum" mechanisms of phage resistance, such as capsules that avoid phage binding (Schooley et al., 2017).

Furthermore, Ando et al. (2015) swapped tail fiber genes to allow a genetically *E. coli* to target *Klebsiella* and vice versa. The authors also demonstrated that synthetic phage cocktails with the same scaffold, but different tails selectively remove bacteria from multi-specie communities. Bacterial biofilms are difficult to eradicate since the physical properties of matrix, the physicochemical properties of the exopolysaccharides, and the heterogeneity of the bacterial cells within the biofilm confers them a strong resistance against chemicals (antibiotics, immunological approaches, and phages). Nevertheless, phages genetically engineered could be used to overcome this problem. Chen et al. (2021) constructed an engineered T4 called T4 Rn11 that exhibits antibiofilm activity against *Streptococcus mutants* (*S. mutants*). A reduction in biofilm biomass and formation of microcolonies was achieved using this phage. In a similar work, Pei and Lamas-Samanamud (2014) constructed a T7 phage that produces a lactonase enzyme with broad-range activity for quenching quorum sensing. The modified phage effectively degrades acyl-homoserine lactones (AHLs) from many bacteria. In addition, it inhibited *P. aeruginosa* and *E. coli* biofilm formation.

Reducing biofilm is crucial in the treatment of infection caused by microorganisms in cystic fibrosis disease (Dedrick et al., 2019) or in wound infections, but also it is useful in reducing biofilm in medical devices such as catheters (Fu et al., 2010). These devices are responsible for substantial morbidity and mortality among patients.

As we described above, combining phage's natural abilities with synthetic biology tools can improve the potential for phages by expanding their range of infection and mediating bacterial signaling or expressing enzymes (Bikard et al., 2014) that help eliminate multi-resistant bacteria.

MICROBIAL CONTROL SYSTEMS

In the last decade, synthetic biology tools have allowed the construction of microbial control systems by engineering whole living cells to act as biosensors and detect and respond to internal and external signals [quorum sensing (QS)] secreted by pathogens (Benítez-Chao et al., 2021; Perrino et al., 2021). Some QS-based phenotypical behaviors in bacteria are sporulation formation, virulence factors related to invasion, bioluminescence, and population control (Benítez-Chao et al., 2021).

In numerous studies, synthetic genetic circuits have been developed to analyze biological systems and guide their design based on QS. A group of researchers genetically modified *E. coli* to detect wild-type *P. aeruginosa* (PAO1), specifically via its QS molecule. Their results demonstrated that engineered *E. coli* sentinels successfully inhibit PAO1 growth by secreting a

TABLE 1 | Synthetic biology tools to fight antibiotic resistance using genetic engineering to modify antibiotics and enhance antibiotic production, engineered phages, and microbial control systems.

Genetic engineering to modify antibiotics		
Heterologous Expression	Combinations of 13 scaffold-modifying enzymes from 7 GPA BCGs in *Streptomyces coelicolor*	Yim et al. (2016)
	Corbomicyn improvement with glycopeptide antibiotic heterologous expression system GPAHex in *Streptomyces coelicolor*	Xu et al. (2020)
	Application of heterologous expression, deletion, and overexpression to achieve an increase of bacitraicn yield in *Bacillus licheniformis*	Cai et al. (2019); Cai et al. (2020)
Transcriptional optimization of genes	"Top-down" approach to increase lipopeptide production	Ji et al. (2022)
Deletion of genes	Elucidation of the mechanisms involved in synthesizing antibiotics using in-frame deletion of biosynthetic genes	Masschelein et al. (2013); Masschelein et al. (2015)
Genetic engineering to enhance antibiotic production		
Mutations	May increase or decrease antimicrobial synthesis depending of the mutation	Miao et al. (2006); Ma et al. (2017); Arafat et al. (2021)
Gene control (elimination, overexpression)	May enhance the biosynthesis of the antimicrobial compound. May require the modification of different genes to increase its production	Zhu et al. (2012) Byrne et al. (2003) Makitrynskyy et al. (2021) Song et al. (2017); Li et al. (2021a); Li et al. (2021b)
Metabolic pathway modifications	May increase the number of metabolic precursors for a certain biosynthesis pathway Metabolite's accumulation may overload the biosynthesis pathway	Meng et al. (2017); Moosmann et al. (2020) Shomar et al. (2018) Butler et al. (2002)
Gene introduction	Gives the possibility of using a different host for better biosynthesis The difference between genres and species may redirect the synthesis to different analogs	Han et al. (2012); Sakai et al. (2012); Makitrynskyy et al. (2021) Chen et al. (2009) Eustáquio et al. (2004) Wang et al. (2021) Eppelmann et al. (2001) Wu et al. (2015)
Engineered phages to fight superbugs		
Homologous recombination	Deliver genes to replace antibiotic resistance genes into bacteria Replace genes in phages to shift or broad host ranges	(Edgar et al., 2012) (Lin et al., 2012)
DNA sequence-specific antimicrobials	Deliver CRISPR-Cas9 system into cytoplasm to kill bacteria	(Bikard et al., 2014)
Phage-Display	Conjugate antibiotic with phages to enhance bactericidal activity	(Yacoby et al., 2007)
Microbial control systems		
Quorum sensing (QS)	*Escherichia coli* to detect wild-type *Pseudomonas aeruginosa* Lactococcus lactis to detect *Enterococcus faecalis* Biosensing *Vibrio cholerae* with *Escherichia coli* *Lactobacillus* reuteri to detect *Staphylococcus* sp *Escherichia coli biosensors to detect Pseudomonas aeruginosa and Burkholderia pseudomallei*	Gupta et al. (2013) Borrero et al. (2015) Holowko et al. (2016) Lubkowicz et al. (2018) Wu et al. (2021)

novel pathogen-specific engineered chimeric bacteriocin (Gupta et al., 2013). Later, a *Lactococcus lactis* (*L. lactis*) to detect *E. faecalis* was designed in another study. *L. lactis* was able to produce and secrete peptides that inhibit enterococcal growth and reduce the viability of enterococci in the surrounding area of *L. lactis*. This engineered system was demonstrated to be effective against multidrug-resistant *E. faecium* strains (Borrero et al., 2015). Holowko et al. (2016) designed and created a synthetic genetic sensing system using nonpathogenic *E. coli* as the host based on CRISPRi technology. They moved proteins used by *Vibrio cholerae* (*V. cholerae*) for QS into *E. coli* and showed high sensitivity to the presence of *V. cholerae* supernatant with tight control of expression of output GFP protein (Holowko et al., 2016). Lubkowicz et al. (2018) developed a probiotic lactic acid

bacteria (*Lactobacillus reuteri*) engineered to detect autoinducer peptide-I (AIP-I), a quorum-sensing molecule produced by *Staphylococcus* sp. during pathogenesis. Their results showed that the engineered biosensor could detect AIP-I levels under various strenuous conditions in the *S. aureus* (Lubkowicz et al., 2018). Recently, Wu *et al.* created a novel whole-cell biosensor to detect bacterial pathogens (*P. aeruginosa* and *Burkholderia pseudomallei*) by responding to the relevant QS signal molecules. Results showed that engineered whole-cell biosensors provide rapid and cost-effective detection of waterborne pathogens (Wu et al., 2021).

As described above, quorum sensing is a process in which bacteria communicate with their own species or across species to coordinate cellular behavior. In the last decade, synthetic biology has used the properties of quorum sensing as a tool to build and develop genetic circuits for population control. Thus, scientists seek to design biological systems with predictable behavior. **Table 1** displays a summary of the synthetic biology tools to fight antibiotic resistance using genetic engineering to modify antibiotics and enhance antibiotic production, engineered phages, and microbial control systems.

OTHER SYNTHETIC BIOLOGY TOOLS TO ENGINEER MICROBIAL COMMUNITIES

Synthetic biology uses engineering principles in biological systems in a predictable, controllable, and standardized manner to get new biological insights and create cells with improved abilities. The potential uses of emerging technologies such as the Design-Build-Test-Learn (DBTL) cycle, protein engineering against multidrug-resistant bacteria, and logic gates are described below.

Design-Build-Test-Learn Cycle

As we previously described, synthetic biology is a recently emerging discipline used to design or redesign biological systems and give them improved or new qualities. One of the SynBio engineering principles used to create new biological systems is the Design-Build-Test-Learn (DBTL) cycle (Whitford et al., 2021). DBTL cycle is an increasingly adopted metabolic engineering framework that helps systematize cellular activities and increase their efficacy and generalizability (Opgenorth et al., 2019). Thus, the design of biological circuits and high-throughput screening (HTS) technologies have begun to speed up modern drug discovery cycles and produce new medicines. Each phase of the DBTL cycle is a fundamental component of the cycle. The Design component identifies the problem and selects the desired pathway and host. The Build component selects, synthesizes, and assembles parts for the incorporation into the host. The Test component validates the engineered constructs and strains for target molecule production, transcripts, proteins, and metabolites. This phase generates a significant amount of data. The Learn component analyzes the Test data, and the learnings are used to create a novel testable hypothesis and incorporate them into the next cycle (Petzold et al., 2015; Carbonell et al., 2018). With the DBTL cycle, researchers can rapidly construct new biological systems. Coupling the DBTL cycle with engineering technologies can deliver solutions in drug discovery and solve global problems such as antimicrobial resistance.

Protein Engineering Against Multidrug-Resistant Bacteria

Like endolysins (encoded by bacterial viruses), proteins act by disrupting the bacterial cell wall or other proteins that inhibit the expression of genes related to antibiotic resistance (gene regulators) and leading to cell death. Some proteins alter microbial behavior, the bacterial cell wall and interrupt signals and genes confering multi-drug resistance. Based on these kinds of proteins, synthetic biology techniques are helpful for engineering proteins into making them more stable or efficient. Moreover, synthetic proteins can be created and manufactured such is the case of the Briers et al. (2014) group that engineered endolysins to act as artilysins (outer membrane-penetrating endolysins Briers et al., 2014). One drawback is that endolysins do not have activity against Gram-negative bacteria because of the impermeable lipopolysaccharide layer surrounding their cell wall. Still, artilysins are highly bactericidal against Gram-negative pathogens, including *P. aeruginosa* and *A. baumannii*.

Another example is dCas9 which is a "dead" catalytical protein. The system is the same used in CRISPR-Cas9 technology, with the difference that dCas9 can bind to the DNA but not cut the strands. In this way, dCas9 is targeted to suppress gene transcription. Wang and Nicholaou (2017) designed two CRISPR-dCas9 systems to target the *mecA* promoter in MRSA to repress the gene's transcription (Wang and Nicholaou, 2017). Although these alternatives are not applied to clinical, they are promising approaches to use alone or combined with antibiotics or another technology.

Logic Gates

An interesting line of research in synthetic biology is that of making microorganisms work as control elements, sending and receiving signals through logic gates. Logic gates are widely used in electronic devices. These devices have electronic circuits that transform an input signal into another output signal or convert two input signals into one output signal (Wang et al., 2011). In synthetic biology, living cells act as control elements; therefore, living cells can be programmed to produce the precisely desired behavior in response to intra or extracellular signals (Morris et al., 2010). Gene regulatory network that cells use to interact and respond to the environmental signals, can be used to program logic circuits to link cellular sensors and actuators. Genetic logic gates rely on direct control of the transcription, activating a promoter. One example is using these genetic elements to program a cell to recognize quorum-sensing signals. Benítez-Chao et al. (2021) designed a whole-cell biosensor to sense and kill MRSA. Using genetic logic gates, they design a genetic circuit where *E. coli* can recognize quorum-sensing molecules from MRSA and triggers the expression of a bacteriocin that kills MRSA.

Genetic circuits can also be designed to disrupt a metabolic pathway in multi-drug resistant bacteria or to create whole-cell biosensors, responding to a molecule produced by pathogen organisms and producing an antimicrobial compound. Whole-cell biosensors are still in research but have a promising future in controlling pathogen bacteria.

CONCLUSION

This review presents recent approaches based on synthetic biology as an emerging tool to produce new therapeutic compounds. Synthetic biology provides a wide variety of techniques that aid in the design and construction of microorganisms capable of creating, producing, and enhancing biological functions to use them to tackle antibiotic resistance. Still, there are significant challenges to overcome, like 1) many biological components lack clear description, 2) the construction of biological systems is complicated and can be unpredictable, and 3) bio-circuit test is complicated and time-consuming. Therefore, it is necessary to keep studying and innovating processes before being translated into human medicine.

AUTHOR CONTRIBUTIONS

Conceptualization, AL-B and JM-R. Writing-original draft preparation AL-B, FB-C, CG-C, JG-C, and JM-R. Writing-review and editing, AL-B, FB-C, CG-C, JG-C, and JM-R. Supervision, AL-B and JM-R. All authors contributed to the article and approved the submitted version.

FUNDING

We would like to acknowledge Paicyt 2019–2020 and 2020–2021 Science Grant from the Universidad Autónoma de Nuevo León; CONACyT Grants for Basic Science grant 221332; Fronteras de la Ciencia grant 1502, Infraestructura Grant 279957 and Apoyos a la Ciencia de Frontera Grant 316869.

ACKNOWLEDGMENTS

CG-C would like to thank the support from Beca Nacional de Maestría 2020–2022 (Scholarship number 1080929).

REFERENCES

Ando, H., Lemire, S., Pires, D. P., and Lu, T. K. (2015). Engineering Modular Viral Scaffolds for Targeted Bacterial Population Editing. *Cell Syst.* 1, 187–196. doi:10.1016/j.cels.2015.08.013

Arafat, H. H., Abu-Tahon, M. A., and Isaac, G. S. (2021). Production of Antibiotic Carbomycin from Streptomyces Graminofaciens with High Lipid Content Mutation. *Arch. Microbiol.* 203, 901–911. doi:10.1007/s00203-020-02085-6

Benítez-Chao, D. F., Balderas-Cisneros, F. d. J., León-Buitimea, A., and Morones-Ramírez, J. R. (2021). Design and In Silico Analysis of a Whole-cell Biosensor Able to Kill Methicillin-resistant *Staphylococcus aureus*. *Biotech App Biochem.* 2021, 2210. doi:10.1002/bab.2210

Bergh, Ø., BØrsheim, K. Y., Bratbak, G., and Heldal, M. (1989). High Abundance of Viruses Found in Aquatic Environments. *Nature* 340, 467–468. doi:10.1038/340467a0

Bikard, D., Euler, C. W., Jiang, W., Nussenzweig, P. M., Goldberg, G. W., Duportet, X., et al. (2014). Exploiting CRISPR-Cas Nucleases to Produce Sequence-Specific Antimicrobials. *Nature Biotechnol.* 32, 1146–1150. doi:10.1038/nbt.3043

Borrero, J., Chen, Y., Dunny, G. M., and Kaznessis, Y. N. (2015). Modified Lactic Acid Bacteria Detect and Inhibit Multiresistant Enterococci. *ACS Synth. Biol.* 4, 299–306. doi:10.1021/sb500090b

Breitling, R., and Takano, E. (2015). Synthetic Biology Advances for Pharmaceutical Production. *Curr. Opin. Biotechnol.* 35, 46–51. doi:10.1016/j.copbio.2015.02.004

Briers, Y., Walmagh, M., van Puyenbroeck, V., Cornelissen, A., Cenens, W., Aertsen, A., et al. (2014). Engineered Endolysin-Based "Artilysins" to Combat Multidrug-Resistant Gram-Negative Pathogens. *mBio* 5. doi:10.1128/mBio.01379-14

Butler, M. J., Bruheim, P., Jovetic, S., Marinelli, F., Postma, P. W., and Bibb, M. J. (2002). Engineering of Primary Carbon Metabolism for Improved Antibiotic Production in Streptomyces Lividans. *Appl. Environ. Microbiol.* 68, 4731–4739. doi:10.1128/AEM.68.10.4731-4739.2002

Byrne, B., Carmody, M., Gibson, E., Rawlings, B., and Caffrey, P. (2003). Biosynthesis of Deoxyamphotericins and Deoxyamphoteronolides by Engineered Strains of Streptomyces Nodosus. *Chem. Biol.* 10, 1215–1224. doi:10.1016/j.chembiol.2003.12.001

Cai, D., Zhang, B., Zhu, J., Xu, H., Liu, P., Wang, Z., et al. (2020). Enhanced Bacitracin Production by Systematically Engineering S-Adenosylmethionine Supply Modules in Bacillus Licheniformis. *Front. Bioeng. Biotechnol.* 8, 305. doi:10.3389/fbioe.2020.00305

Cai, D., Zhu, J., Zhu, S., Lu, Y., Zhang, B., Lu, K., et al. (2019). Metabolic Engineering of Main Transcription Factors in Carbon, Nitrogen, and Phosphorus Metabolisms for Enhanced Production of Bacitracin in Bacillus Licheniformis. *ACS Synth. Biol.* 8, 866–875. doi:10.1021/acssynbio.9b00005

Carbonell, P., Jervis, A. J., Robinson, C. J., Yan, C., Dunstan, M., Swainston, N., et al. (2018). An Automated Design-Build-Test-Learn Pipeline for Enhanced Microbial Production of Fine Chemicals. *Commun. Biol.* 1, 66. doi:10.1038/s42003-018-0076-9

Chen, J., Chen, Z., Yuan, K., Huang, Z., and Mao, M. (2021). Recombinant Bacteriophage T4 Rnl1 Impacts Streptococcus Mutans Biofilm Formation. *J. Oral Microbiol.* 13, 1–8. doi:10.1080/20002297.2020.1860398

Chen, W., Huang, T., He, X., Meng, Q., You, D., Bai, L., et al. (2009). Characterization of the Polyoxin Biosynthetic Gene Cluster from Streptomyces Cacaoi and Engineered Production of Polyoxin H. *J. Biol. Chem.* 284, 10627–10638. doi:10.1074/jbc.M807534200

Culp, E. J., Waglechner, N., Wang, W., Fiebig-Comyn, A. A., Hsu, Y. P., Koteva, K., et al. (2020). Evolution-Guided Discovery of Antibiotics That Inhibit Peptidoglycan Remodelling. *Nature* 578 (7796), 582–587.

Dedrick, R. M., Guerrero-Bustamante, C. A., Garlena, R. A., Russell, D. A., Ford, K., Harris, K., et al. (2019). Engineered Bacteriophages for Treatment of a Patient With a Disseminated Drug-Resistant Mycobacterium Abscessus. *Nature Med.* 25, 730–733. doi:10.1038/s41591-019-0437-z

Edgar, R., Friedman, N., Molshanski-Mor, S., and Qimron, U. (2012). Reversing Bacterial Resistance to Antibiotics by Phage-Mediated Delivery of Dominant Sensitive Genes. *Appl. Environ. Microbiol.* 78, 744–751. doi:10.1128/AEM.05741-11

El Karoui, M., Hoyos-Flight, M., and Fletcher, L. (2019). Future Trends in Synthetic Biology-A Report. *Front. Bioeng. Biotechnol.* 7, 175. doi:10.3389/fbioe.2019.00175

Eppelmann, K., Doekel, S., and Marahiel, M. A. (2001). Engineered Biosynthesis of the Peptide Antibiotic Bacitracin in the Surrogate Host Bacillus Subtilis. *J. Biol. Chem.* 276, 34824–34831. doi:10.1074/jbc.M104456200

Eustáquio, A. S., Gust, B., Li, S.-M., Pelzer, S., Wohlleben, W., Chater, K. F., et al. (2004). Production of 8′-Halogenated and 8′-Unsubstituted Novobiocin Derivatives in Genetically Engineered Streptomyces Coelicolor Strains. *Chem. Biol.* 11, 1561–1572. doi:10.1016/j.chembiol.2004.09.009

Fu, W., Forster, T., Mayer, O., Curtin, J. J., Lehman, S. M., Donlan, R. M., et al. (2010). Bacteriophage Cocktail for the Prevention of Biofilm Formation by

Pseudomonas Aeruginosa on Catheters in an in vitro Model System. *Antimicrob. Agents Chemother.* 54, 397–404. doi:10.1128/AAC.00669-09

Gupta, S., Bram, E. E., and Weiss, R. (2013). Genetically Programmable Pathogen Sense and Destroy. *ACS Synth. Biol.* 2, 715–723. doi:10.1021/sb4000417

Han, A. R., Shinde, P. B., Park, J. W., Cho, J., Lee, S. R., Ban, Y. H., et al. (2012). Engineered Biosynthesis of Glycosylated Derivatives of Narbomycin and Evaluation of Their Antibacterial Activities. *Appl. Microbiol. Biotechnol.* 93, 1147–1156. doi:10.1007/s00253-011-3592-9

Hennig, S., Rödel, G., and Ostermann, K. (2015). Artificial Cell-Cell Communication as an Emerging Tool in Synthetic Biology Applications. *J. Biol. Eng.* 9, 13–12. doi:10.1186/S13036-015-0011-2/FIGURES/2

Holowko, M. B., Wang, H., Jayaraman, P., and Poh, C. L. (2016). Biosensing *Vibrio cholerae* with Genetically Engineered *Escherichia coli*. *ACS Synth. Biol.* 5, 1275–1283. doi:10.1021/acssynbio.6b00079

Huh, J.-H., Kim, D.-j., Zhao, X.-Q., Li, M., Jo, Y.-Y., Yoon, T.-M., et al. (2004). Widespread Activation of Antibiotic Biosynthesis byS-Adenosylmethionine in Streptomycetes. *FEMS Microbiol. Lett.* 238, 439–447. doi:10.1111/j.1574-6968.2004.tb09787.x

Hwang, S., Lee, N., Cho, S., Palsson, B., and Cho, B. K. (2020). Repurposing Modular Polyketide Synthases and Non-ribosomal Peptide Synthetases for Novel Chemical Biosynthesis. *Front. Mol. Biosci.* 7, 87. doi:10.3389/fmolb.2020.00087

Ji, C.-H., Kim, H., Je, H.-W., Kwon, H., Lee, D., and Kang, H.-S. (2022). Top-down Synthetic Biology Approach for Titer Improvement of Clinically Important Antibiotic Daptomycin in Streptomyces Roseosporus. *Metab. Eng.* 69, 40–49. doi:10.1016/j.ymben.2021.10.013

Kim, H. U., Charusanti, P., Lee, S. Y., and Weber, T. (2016). Metabolic Engineering with Systems Biology Tools to Optimize Production of Prokaryotic Secondary Metabolites. *Nat. Prod. Rep.* 33, 933–941. doi:10.1039/C6NP00019C

Li, D., Tian, Y., Liu, X., Wang, W., Li, Y., Tan, H., et al. (2021a). Reconstitution of a Mini-gene Cluster Combined with Ribosome Engineering Led to Effective Enhancement of Salinomycin Production in Streptomyces Albus. *Microb. Biotechnol.* 14, 2356–2368. doi:10.1111/1751-7915.13686

Li, Y.-P., Bu, Q.-T., Li, J.-F., Xie, H., Su, Y.-T., Du, Y.-L., et al. (2021b). Genome-based Rational Engineering of Actinoplanes Deccanensis for Improving Fidaxomicin Production and Genetic Stability. *Bioresour. Technol.* 330, 124982. doi:10.1016/j.biortech.2021.124982

Liao, L., Cheng, Y., Liu, S., Zhou, J., An, S., Lv, M., et al. (2014). Production of Novel Antibiotics Zeamines through Optimizing Dickeya Zeae Fermentation Conditions. *PLoS One* 9, e116047. doi:10.1371/journal.pone.0116047

Lin, T.-Y. Y., Lo, Y.-H. H., Tseng, P.-W. W., Chang, S.-F. F., Lin, Y.-T. T., and Chen, T.-S. S. (2012). A T3 and T7 Recombinant Phage Acquires Efficient Adsorption and a Broader Host Range. *PLoS ONE* 7, 1–10. doi:10.1371/journal.pone.0030954

Lubkowicz, D., Ho, C. L., Hwang, I. Y., Yew, W. S., Lee, Y. S., and Chang, M. W. (2018). Reprogramming Probiotic Lactobacillus Reuteri as a Biosensor for *Staphylococcus aureus* Derived AIP-I Detection. *ACS Synth. Biol.* 7, 1229–1237. doi:10.1021/acssynbio.8b00063

Ma, J., Huang, H., Xie, Y., Liu, Z., Zhao, J., Zhang, C., et al. (2017). Biosynthesis of Ilamycins Featuring Unusual Building Blocks and Engineered Production of Enhanced Anti-tuberculosis Agents. *Nat. Commun.* 8, 1–10. doi:10.1038/s41467-017-00419-5

Makitrynskyy, R., Tsypik, O., and Bechthold, A. (2021). Genetic Engineering of Streptomyces Ghanaensis ATCC14672 for Improved Production of Moenomycins. *Microorganisms* 10, 30. doi:10.3390/microorganisms10010030

Martínez-Núñez, M. A., and López, V. E. L. y. (2016). Nonribosomal Peptides Synthetases and Their Applications in Industry. *Sustain. Chem. Process* 4, 13. doi:10.1186/s40508-016-0057-6

Masschelein, J., Clauwers, C., Awodi, U. R., Stalmans, K., Vermaelen, W., Lescrinier, E., et al. (2015). A Combination of Polyunsaturated Fatty Acid, Nonribosomal Peptide and Polyketide Biosynthetic Machinery Is Used to Assemble the Zeamine Antibiotics. *Chem. Sci.* 6, 923–929. doi:10.1039/C4SC01927J

Masschelein, J., Mattheus, W., Gao, L.-J., Moons, P., Van Houdt, R., Uyterhoeven, B., et al. (2013). A PKS/NRPS/FAS Hybrid Gene Cluster from Serratia Plymuthica RVH1 Encoding the Biosynthesis of Three Broad Spectrum, Zeamine-Related Antibiotics. *PLoS One* 8, e54143. doi:10.1371/journal.pone.0054143

McCarty, N. S., and Ledesma-Amaro, R. (2019). Synthetic Biology Tools to Engineer Microbial Communities for Biotechnology. *Trends Biotechnol.* 37, 181–197. doi:10.1016/j.tibtech.2018.11.002

Meng, S., Wu, H., Wang, L., Zhang, B., and Bai, L. (2017). Enhancement of Antibiotic Productions by Engineered Nitrate Utilization in Actinomycetes. *Appl. Microbiol. Biotechnol.* 101, 5341–5352. doi:10.1007/s00253-017-8292-7

Miao, V., Coëffet-Le Gal, M.-F., Nguyen, K., Brian, P., Penn, J., Whiting, A., et al. (2006). Genetic Engineering in Streptomyces Roseosporus to Produce Hybrid Lipopeptide Antibiotics. *Chem. Biol.* 13, 269–276. doi:10.1016/j.chembiol.2005.12.012

Moosmann, D., Mokeev, V., Kulik, A., Osipenkov, N., Kocadinc, S., Ort-Winklbauer, R., et al. (2020). Genetic Engineering Approaches for the Fermentative Production of Phenylglycines. *Appl. Microbiol. Biotechnol.* 104, 3433–3444. doi:10.1007/s00253-020-10447-9

Morris, M. K., Saez-Rodriguez, J., Sorger, P. K., and Lauffenburger, D. A. (2010). Logic-Based Models for the Analysis of Cell Signaling Networks. *Biochem.* 49, 3216–3224. doi:10.1021/bi902202q

Nivina, A., Yuet, K. P., Hsu, J., and Khosla, C. (2019). Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chem. Rev.* 119, 12524–12547. doi:10.1021/acs.chemrev.9b00525

Olano, C., Lombó, F., Méndez, C., and Salas, J. A. (2008). Improving Production of Bioactive Secondary Metabolites in Actinomycetes by Metabolic Engineering. *Metab. Eng.* 10, 281–292. doi:10.1016/j.ymben.2008.07.001

Opgenorth, P., Costello, Z., Okada, T., Goyal, G., Chen, Y., Gin, J., et al. (2019). Lessons from Two Design-Build-Test-Learn Cycles of Dodecanol Production in *Escherichia coli* Aided by Machine Learning. *ACS Synth. Biol.* 8, 1337–1351. doi:10.1021/acssynbio.9b00020

Palazzotto, E., Tong, Y., Lee, S. Y., and Weber, T. (2019). Synthetic Biology and Metabolic Engineering of Actinomycetes for Natural Product Discovery. *Biotechnol. Adv.* 37, 107366. doi:10.1016/j.biotechadv.2019.03.005

Pei, R., and Lamas-Samanamud, G. R. (2014). Inhibition of Biofilm Formation by T7 Bacteriophages Producing Quorum-quenching Enzymes. *Appl. Environ. Microbiol.* 80, 5340–5348. doi:10.1128/AEM.01434-14

Perrino, G., Hadjimitsis, A., Ledesma-Amaro, R., and Stan, G.-B. (2021). Control Engineering and Synthetic Biology: Working in Synergy for the Analysis and Control of Microbial Systems. *Curr. Opin. Microbiol.* 62, 68–75. doi:10.1016/j.mib.2021.05.004

Petzold, C. J., Chan, L. J. G., Nhan, M., and Adams, P. D. (2015). Analytics for Metabolic Engineering. *Front. Bioeng. Biotechnol.* 3, 135. doi:10.3389/fbioe.2015.00135

Sakai, A., Mitsumori, A., Furukawa, M., Kinoshita, K., Anzai, Y., and Kato, F. (2012). Production of a Hybrid 16-membered Macrolide Antibiotic by Genetic Engineering of Micromonospora Sp. TPMA0041. *J. Ind. Microbiol. Biotechnol.* 39, 1693–1701. doi:10.1007/s10295-012-1173-2

Schooley, R. T., Biswas, B., Gill, J. J., Hernandez-Morales, A., Lancaster, J., Lessor, L., et al. (2017). Development and Use of Personalized Bacteriophage-Based Therapeutic Cocktails to Treat a Patient with a Disseminated Resistant Acinetobacter Baumannii Infection. *Antimicrob. Agents Chemother.* 61, e00954–17. doi:10.1128/AAC.00954-17

Sharma, S., Chatterjee, S., Datta, S., Prasad, R., Dubey, D., Prasad, R. K., et al. (2017). Bacteriophages and its Applications: an Overview. *Folia Microbiol.* 62, 17–55. doi:10.1007/s12223-016-0471-x

Shomar, H., Gontier, S., Van Den Broek, N. J. F., Tejeda Mora, H., Noga, M. J., Hagedoorn, P.-L., et al. (2018). Metabolic Engineering of a Carbapenem Antibiotic Synthesis Pathway in *Escherichia coli*. *Nat. Chem. Biol.* 14, 794–800. doi:10.1038/s41589-018-0084-6

Song, K., Wei, L., Liu, J., Wang, J., Qi, H., and Wen, J. (2017). Engineering of the LysR Family Transcriptional Regulator FkbR1 and its Target Gene to Improve Ascomycin Production. *Appl. Microbiol. Biotechnol.* 101, 4581–4592. doi:10.1007/S00253-017-8242-4

Ventola, C. L. (2015). The Antibiotic Resistance Crisis: Part 1: Causes and Threats. *Pharm. Ther.* 40 (4), 277.

Wang, B., Kitney, R. I., Joly, N., and Buck, M. (2011). Engineering Modular and Orthogonal Genetic Logic Gates for Robust Digital-Like Synthetic Biology. *Nat. Commun.* 2, 508. doi:10.1038/ncomms1516

Wang, K., and Nicholaou, M. (2017). Suppression of Antimicrobial Resistance in MRSA Using CRISPR-dCas9.. *Am. Soc. Clin. Lab. Sci.* 30, 207–213. doi:10.29074/ascls.30.4.207

Wang, W.-F., Xiao, H., and Zhong, J.-J. (2021). Biosynthesis of a Novel Ganoderic Acid by Expressing CYP Genes from Ganoderma Lucidum in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 106, 523–534. doi:10.1007/s00253-021-11717-w

Wang, Y., Boghigian, B. A., and Pfeifer, B. A. (2007). Improving Heterologous Polyketide Production in *Escherichia coli* by Overexpression of an S-Adenosylmethionine Synthetase Gene. *Appl. Microbiol. Biotechnol.* 77, 367–373. doi:10.1007/s00253-007-1172-9

Whitford, C. M., Cruz-Morales, P., Keasling, J. D., and Weber, T. (2021). The Design-Build-Test-Learn Cycle for Metabolic Engineering of Streptomycetes. *Essays Biochem.* 65, 261–275. doi:10.1042/EBC20200132

World Health Organization, [WHO] (2017). Who Publishes List of Bacteria for Which new Antibiotics are Urgently Needed. *Saudi Med. J.*Available at: https://www.who.int/news-room/detail/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed.

Wu, J., Zhang, H.-B., Xu, J.-L., Cox, R. J., Simpson, T. J., and Zhang, L.-H. (2010). 13CLabeling Reveals Multiple Amination Reactions in the Biosynthesis of a Novel Polyketide Polyamine Antibioticzeamine from Dickeya Zeae. *Chem. Commun.* 46, 333–335. doi:10.1039/B916307G

Wu, M.-C., Struck, A.-W., Nunns, L., Styles, M. Q., Micklefield, J., and Law, B. J. C. (2015). Engineered Biosynthesis of Enduracidin Lipoglycopeptide Antibiotics Using the Ramoplanin Mannosyltransferase Ram29. *Microbiol. (United Kingdom)* 161, 1338–1347. doi:10.1099/mic.0.000095

Wu, Y., Wang, C.-W., Wang, D., and Wei, N. (2021). A Whole-Cell Biosensor for Point-of-Care Detection of Waterborne Bacterial Pathogens. *ACS Synth. Biol.* 10, 333–344. doi:10.1021/acssynbio.0c00491

Xu, M., Wang, W., Waglechner, N., Culp, E. J., Guitor, A. K., and Wright, G. D. (2020). GPAHex-A Synthetic Biology Platform for Type IV-V Glycopeptide Antibiotic Production and Discovery. *Nat. Commun.* 11, 5232. doi:10.1038/s41467-020-19138-5

Yacoby, I., Bar, H., and Benhar, I. (2007). Targeted Drug-Carrying Bacteriophages as Antibacterial Nanomedicines. *Antimicrob. Agents Chemother.* 51, 2156–2163. doi:10.1128/AAC.00163-07

Ye, C., Ng, I.-S., Jing, K., and Lu, Y. (2014). Direct Proteomic Mapping of Streptomyces Roseosporus NRRL 11379 with Precursor and Insights into Daptomycin Biosynthesis. *J. Biosci. Bioeng.* 117, 591–597. doi:10.1016/j.jbiosc.2013.10.021

Yehl, K., Lemire, S., Yang, A. C., Ando, H., Mimee, M., Torres, M. D. T., et al. (2019). Engineering Phage Host-Range and Suppressing Bacterial Resistance through Phage Tail Fiber Mutagenesis. *Cell* 179, 459–469.e9. doi:10.1016/j.cell.2019.09.015

Yim, G., Thaker, M. N., Koteva, K., and Wright, G. (2014). Glycopeptide Antibiotic Biosynthesis. *J. Antibiot.* 67, 31–41. doi:10.1038/ja.2013.117

Yim, G., Wang, W., Thaker, M. N., Tan, S., and Wright, G. D. (2016). How to Make a Glycopeptide: A Synthetic Biology Approach to Expand Antibiotic Chemical Diversity. *ACS Infect. Dis.* 2, 642–650. doi:10.1021/acsinfecdis.6b00105

Zhang, F., Gao, D., Lin, J., Zhu, M., Zhuang, Z., Duan, Y., et al. (2020). Construction of Inducible Genetic Switch for the Global Regulator WblA to Sustain Both Overproduction of Tiancimycins and On-Demand Sporulation in Streptomyces Sp. CB03234. *ACS Synth. Biol.* 9, 1460–1467. doi:10.1021/acssynbio.0c00114

Zhu, Q., Li, J., Ma, J., Luo, M., Wang, B., Huang, H., et al. (2012). Discovery and Engineered Overproduction of Antimicrobial Nucleoside Antibiotic A201A from the Deep-Sea Marine Actinomycete Marinactinospora Thermotolerans SCSIO 00652. *Antimicrob. Agents Chemother.* 56, 110–114. doi:10.1128/AAC.05278-11

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Bioengineering and Biotechnology

PERSPECTIVE
published: 12 May 2022
doi: 10.3389/fbioe.2022.888732

System Principles Governing the Organization, Architecture, Dynamics, and Evolution of Gene Regulatory Networks

Julio A. Freyre-González[1]*, *Juan M. Escorcia-Rodríguez*[1], *Luis F. Gutiérrez-Mondragón*[1,2], *Jerónimo Martí-Vértiz*[1], *Camila N. Torres-Franco*[1] and *Andrea Zorro-Aranda*[1,3]

[1]Regulatory Systems Biology Research Group, Program of Systems Biology, Center for Genomic Sciences, Universidad Nacional Autónoma de México, Cuernavaca, México, [2]Undergraduate Program in Genomic Sciences, Center for Genomic Sciences, Universidad Nacional Autónoma de México, Cuernavaca, México, [3]Department of Chemical Engineering, Universidad de Antioquia, Medellín, Colombia

OPEN ACCESS

Edited by:
Jean Marie François,
Institut Biotechnologique de Toulouse (INSA), France

Reviewed by:
Grant Bowman,
University of Wyoming, United States
Hyungyu Lim,
University of California, San Diego, United States

***Correspondence:**
Julio A. Freyre-González
jfreyre@ccg.unam.mx

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: 03 March 2022
Accepted: 27 April 2022
Published: 12 May 2022

Citation:
Freyre-González JA, Escorcia-Rodríguez JM, Gutiérrez-Mondragón LF, Martí-Vértiz J, Torres-Franco CN and Zorro-Aranda A (2022) System Principles Governing the Organization, Architecture, Dynamics, and Evolution of Gene Regulatory Networks. Front. Bioeng. Biotechnol. 10:888732. doi: 10.3389/fbioe.2022.888732

Synthetic biology aims to apply engineering principles for the rational, systematical design and construction of biological systems displaying functions that do not exist in nature or even building a cell from scratch. Understanding how molecular entities interconnect, work, and evolve in an organism is pivotal to this aim. Here, we summarize and discuss some historical organizing principles identified in bacterial gene regulatory networks. We propose a new layer, the concilion, which is the group of structural genes and their local regulators responsible for a single function that, organized hierarchically, coordinate a response in a way reminiscent of the deliberation and negotiation that take place in a council. We then highlight the importance that the network structure has, and discuss that the natural decomposition approach has unveiled the system-level elements shaping a common functional architecture governing bacterial regulatory networks. We discuss the incompleteness of gene regulatory networks and the need for network inference and benchmarking standardization. We point out the importance that using the network structural properties showed to improve network inference. We discuss the advances and controversies regarding the consistency between reconstructions of regulatory networks and expression data. We then discuss some perspectives on the necessity of studying regulatory networks, considering the interactions' strength distribution, the challenges to studying these interactions' strength, and the corresponding effects on network structure and dynamics. Finally, we explore the ability of evolutionary systems biology studies to provide insights into how evolution shapes functional architecture despite the high evolutionary plasticity of regulatory networks.

Keywords: gene regulatory networks, organization, functional architecture, system principles, hierarchy, consistency, incompleteness, evolution

INTRODUCTION

Synthetic biology aims to apply engineering principles for the rational, systematical design and construction of biological systems displaying functions that do not exist in nature or even building a cell from scratch (Abil and Danelon, 2020). To fulfill these ambitious goals, we not only need to understand how the various entities within a cell interact but also to identify the principles governing how the cellular systems interconnect, work, and evolve, as these are design cornerstones underpinning a successful rational design.

Whereas studying the whole set of molecular interactions across the different layers (e.g., transport, gene regulation, protein-protein interactions, metabolism, etc.) in a cell is necessary, it is not fully possible nowadays as current knowledge of the networks integrating the different layers is limited, and the integration of heterogeneous networks poses problems not yet solved. We thus focus on gene regulation as it is the key process that controls and integrates signals from all the other layers to cope with the environment.

Advances in understanding the inner workings of small regulatory circuits (i.e., network motifs) have provided good foundations to develop small synthetic circuits, but an understanding of the system principles governing the large-scale organization of complex biological networks is still elusive. However, these principles are pivotal to understanding how the organization of gene regulatory networks (GRNs) governs its possible dynamic outcomes (Ruklisa et al., 2019) and to enabling the successful integration of newly designed systems into the preexisting circuitry of molecular interactions in a chassis.

THE BASIC ORGANIZATIONAL LAYER, COUPLED GENES: THE OPERON

In 1960, Jacob *et al.* proposed the first genetic organizational level in the cell as a "unit of coordinated expression", the operon (Jacob et al., 1960). This functional unit plays a key role in the hypothesis of the operator, explaining the polar effect occurring because of some mutations affecting the induction of enzymes needed to metabolize lactose in *Escherichia coli*. An operon comprises a set of adjacent genes that are regulated as a unit and co-transcribed into a single polycistronic mRNA (Jacob and Monod, 1961) (**Figure 1A**, top left). Genes composing an operon are usually functionally related (de Daruvar et al., 2002; Osbourn and Field, 2009) as they collaborate to attain a specific physiological function, although they commonly possess different biochemical activities. However, there are also cases of operons comprising genes without any apparent functional relation. In these cases, genes may be required in the same environmental conditions despite being involved in different pathways (Osbourn and Field, 2009), as if a special element, responsible for integrating, at the promoter level, disparate physiological responses, was possibly lurking there. While the operon solves the problem of co-regulating functionally related genes diminishing gene expression noise and ensuring more precise stoichiometry (Osbourn and Field, 2009), it has some limitations. First, some cellular processes involve too many genes. For example, anaerobic respiration in *E. coli* comprises more than 150 genes. An operon containing all these genes would encode a huge transcript whose transcription and processing, if possible, would be inefficient. Besides, these dozens of genetic products must be, not only expressed, but also precisely coordinated in time and quantity, something that an operon is unable to achieve.

COORDINATING TIMING AND STOICHIOMETRY OF UNCOUPLED GENES: THE REGULON

A single regulatory protein may affect various promoters shaping what is defined as a regulon as was defined by Maas in 1964 (Maas, 1964). This organization enables the coordination of operons that are physically scattered throughout the genome. There are two types of regulons: simple and complex. Simple regulons are the set of genes, operons, or both regulated by a specific regulatory protein (Maas, 1964), whereas complex regulons are defined as the set of genes, operons, or both regulated by the same set of (two or more) regulatory proteins (Gutierrez-Rios et al., 2003) (**Figure 1A**, top right). As genes composing an operon are usually functionally related, the same holds for the operons controlled by a simple regulon. Besides, the expression of genes composing a regulon is not strictly coordinated, thus allowing variations in quantity and timing of synthesized products. These variations depend upon the concerted action of the respective promoters for each gene or operon in the regulon and the corresponding binding sites for their regulatory proteins. While regulons solve the organizational problems posed by operons, they open a new problem. How to control a single complex function that requires the coordinated expression of different regulons?

THE POWER OF DECENTRALIZED GLOBAL COORDINATION: THE MODULON

The integration of single regulatory circuits into complex networks led Susan Gottesman to propose the existence of global regulatory proteins controlling these global networks in 1984 (Gottesman, 1984). In her seminal paper, she also provided a set of diagnostic criteria to identify this kind of regulator: 1) global regulators control a large number of genes, 2) the regulated genes are involved in more than one metabolic pathway, and 3) global regulators coordinate gene expression in response to a common need. Four years later, Iuchi and Lin defined the modulon as the set of operons, regulons, or both modulated—hence the word modulon, which has no relation

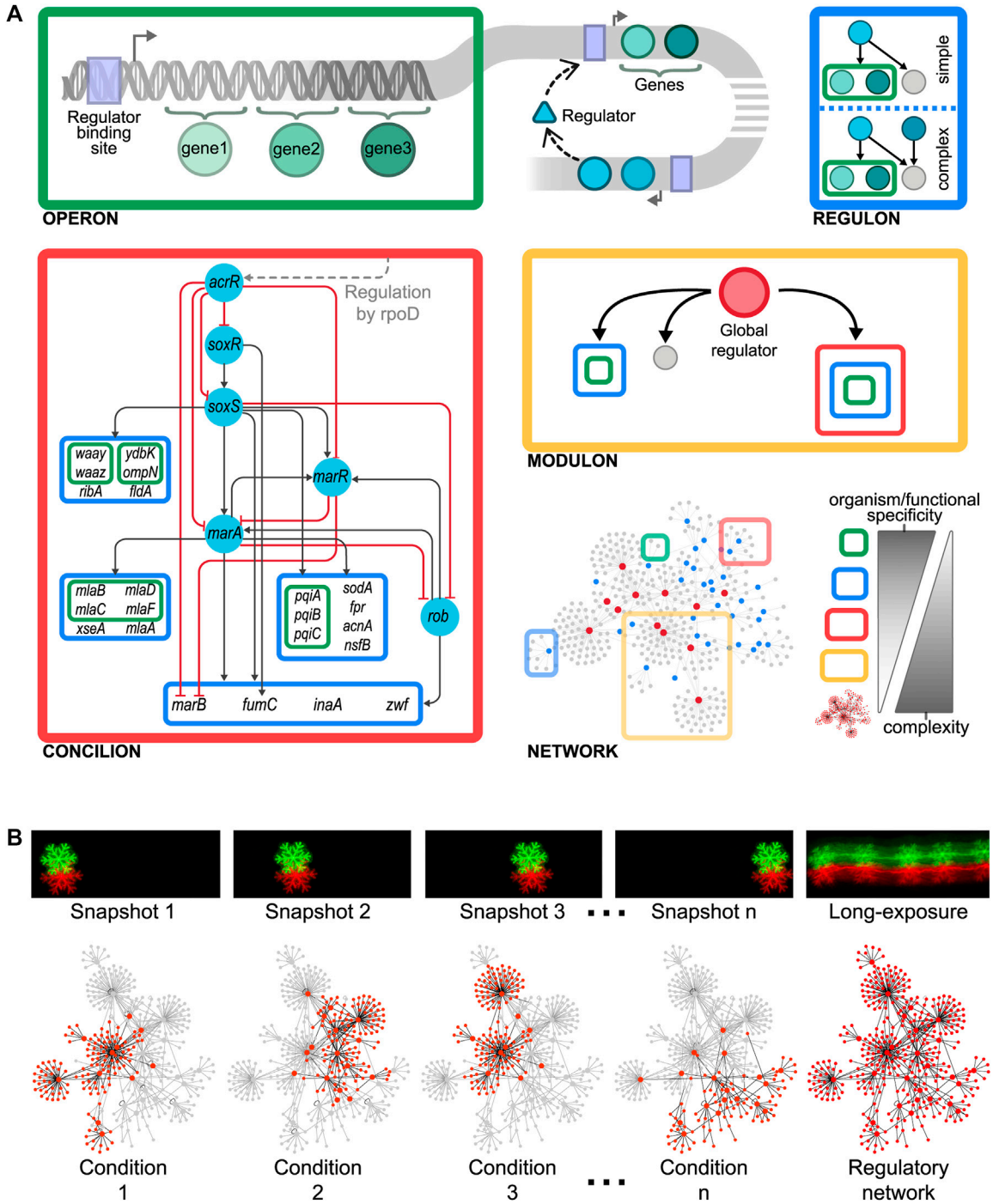


FIGURE 1 | (A) Organizational layers shaping the modular hierarchy of the gene regulatory organization as gene < operon < regulon < concilion < modulon. A biological example of the here-proposed concilion is the "response to multiple stresses" module found in *E. coli* (Escorcia-Rodriguez et al., 2020). The grey dashed line shows that *acrR* is globally controlled by *rpoD*, which also controls other concilions and regulons (**Figure 2A**). The master regulators in this hierarchy are SoxR and SoxS, which respond to oxidative stress through sensing superoxide and nitric oxide. SoxS, MarA, and Rob bind as monomers to the same DNA site, a 20-bp degenerated sequence known as Mar/Sox/Rob box. The differential regulation of these genes could be archived by the degeneracy of their DNA binding sites or by the regulators' concentration and the different affinities for the Mar/Sox/Rob box (Martin et al., 1999; Chubiz et al., 2012). The presence of several paralogous regulators (members of the AraC/XylS family) recognizing the same DNA binding site allows to archive a differential response by activating the same genes in response to different environmental cues (Martin et al., 2008). This phenomenon, known as commensurate regulon activation, enables bacteria to mount a proportionate response of the *marA/soxS/rob* regulon to the stress signal, keeping the number of activated genes to the minimum necessary to cope with prolonged stress (Martin et al., 2008; Wall et al., 2009). This balances the energetic cost of gene expression against the intensity of the stress. **(B)** Curated reconstructed regulatory networks merge many individual condition-specific subnetworks (such as picture snapshots) into a single network model thus capturing all the possible dynamic trajectories (such as a long-exposure photo does). Consequently, curated regulatory networks are not static representations of regulation, as they embed all the potential regulations that can occur thus constraining the large number of organizations a regulatory network could potentially have.

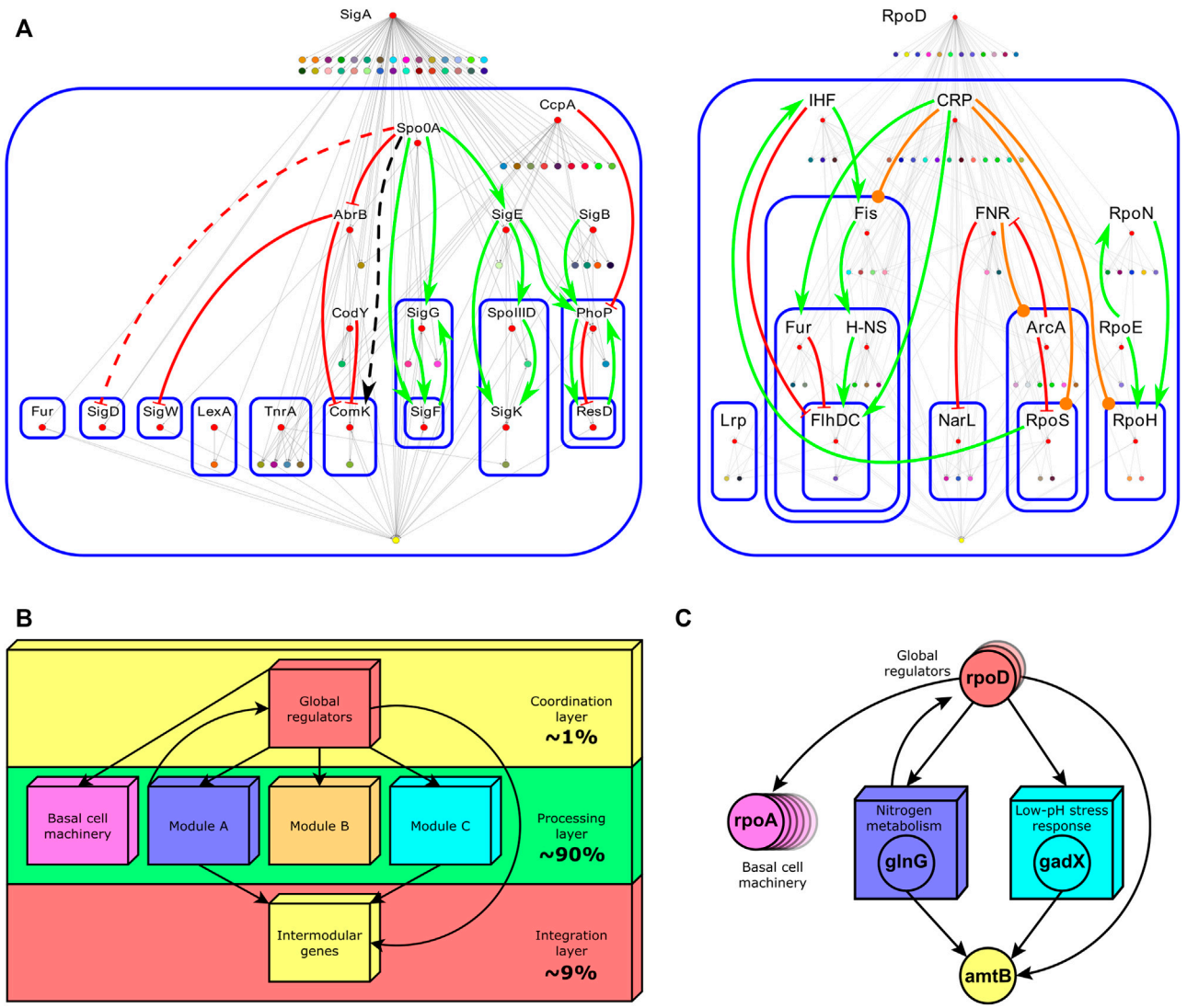


FIGURE 2 | (A) Hierarchies identified by the theoretical pleiotropy approach for *B. subtilis* (left) and *E. coli* (right). Labeled red nodes are global regulators. Nodes composing modules were shrunk into a single colored node. At the bottom of each figure, the yellow node contains the set of intermodular genes. Continuous arrows (red for negative interactions, green for positive ones, orange for duals, and black for interactions whose sign is unknown) indicate regulatory interactions between global regulators. Blue rounded-corner rectangles bound hierarchical layers. For a detailed description of this figure, the reader is referred to the original caption (Freyre-Gonzalez et al., 2012)[1]. **(B)** The common functional architecture found across bacteria by the NDA. Percentages indicate the fraction of genes in the GRN composing that layer. **(C)** A biological example of each layer composing the functional architecture of the *E. coli* GRN. The global regulator *rpoD* is one of the several global regulators controlling genes in many modules (concilions and regulons). Global regulators also control many single genes or operons not regulated by local regulators (basal machinery). Two examples of modules, 'Nitrogen metabolism' and 'Low-pH stress response', are shown. They jointly control the intermodular gene *amtB via* the local regulators *glnG* (NtrC) and *gadX* (GadX). NtrC is the general regulator of the nitrogen assimilation pathway. GadX is one of the central regulators of the glutamate-dependent acid resistance system (GAD system). The *amtB* gene encodes an NH_4^+ antiporter. Disruption of this gene impaired the growth on ammonium only under acidic conditions. Ammonium is also a precursor of glutamate, which plays a central role in the GAD system. This shows that intermodular genes integrate disparate physiological responses coming from different modules.

to the term module—by a common pleiotropic regulatory protein (Iuchi and Lin, 1988) (**Figure 1A**, center right). Here, pleiotropy implies that operons and regulons under control are no longer functionally related. Therefore, mutations in the pleiotropic regulatory protein controlling the modulon give rise to alterations in multiple phenotypic traits in a cell, confirming that global regulators are involved in disparate physiological functions. A pleiotropic or global regulator is responsible for

sensing and responding to signals of general interest for the cell such as DNA[1] damage, stresses, or energy levels (Freyre-Gonzalez et al., 2008). Each global regulator shapes only one modulon and these could overlap by the co-regulation of some genes. Global regulators also shape a hierarchy comprising chains of command having each a specific physiology as has been previously reported for *E. coli* (Freyre-Gonzalez et al., 2008) and *Bacillus subtilis* (Freyre-Gonzalez et al., 2012) (**Figure 2A**). These chains of command modulate the local responses carried on, at the regulon level, by local regulators according to general interest environmental cues (e.g., low glucose, heat, high oxidizing power). Hence, modulons mostly shape a top-down hierarchy that could be seen as the global control device of the cell responsible for the coordination of lower functionally related structures. An interesting biological example of this global control device and its chains of command was outlined for the global regulator CtrA of Caulobacter crescentus (Laub et al., 2000).

THE MISSING PIECE: COORDINATING A SINGLE FUNCTION USING A HIERARCHY OF LOCAL REGULATORS, THE CONCILION

Modularity is an organizing principle in the cell (Hartwell et al., 1999). Genomic islands (e.g., pathogenicity islands, secretion islands, antimicrobial resistance islands, and metabolic islands) and compartmentalization are clear examples of this. As we previously discussed, genes are grouped into operons, regulons, and modulons. All these are kinds of modules shaping the levels of the genetic organization. Indeed, regulons have been considered by far the ultimate level of genetic organization for functionally-related genes (Gutierrez-Rios et al., 2003). However, some complex processes, involving operons, regulons, or both devoted to closely related functions, require the coordinated expression, controlled in both time and quantity, of different regulons. Besides, processing genetic and environmental information may require both 1) dividing tasks into specialized processing units and 2) integrating the resulting information. For example, an antibiotics resistance module may comprise operons or regulons each responsible for providing resistance to different antibiotics. Hence, operons and regulons must be embedded into a complex structure that cannot be reduced into a simple regulon of regulons but that still is responsible for a unique, well-defined physiological function.

We defined this novel structure, previously only loosely named module, as the concilion [kon'si.li.on]. The term is derived from the Latin noun *concilium*, council or meeting, and the verb *conciliō*, to unite, to bring together. This refers to the group of structural genes and their local regulators responsible for a single function that, organized hierarchically, coordinates a response in a way reminiscent of the deliberation and negotiation that take place in a council (**Figure 1A**, bottom left). Concilions may be differentiated from regulons because the former exhibits interactions between their regulators resembling a hierarchical circuit that could even include some feedback and cross-regulation. Moreover, concilions do not contain any global regulator, they are local regulation devices devoted to a unique, well-defined function, contrary to modulons that include a global regulator by definition and control a diversity of functions. By analyzing a non-redundant set containing the most recent GRN for each of the 42 bacteria in Abasy Atlas, we found that, on average, roughly 17% of the modules identified by the natural decomposition approach (NDA, see next section) in a GRN are concilions. Furthermore, in the most recent reconstruction of the *E. coli* GRN (Abasy Atlas regnetid: 511145_v2020_sRDB18-13), we found that about 25% of the modules are concilions whereas the remainder modules are simple or complex regulons, highlighting the important role of the concilion in the functional architecture.

A biological example of this novel genetic organizational level is provided by the "response to multiple stresses" module found in *E. coli* (Escorcia-Rodriguez et al., 2020). This concilion comprises several regulons organized into a regulatory cascade mainly controlled by SoxR, SoxS, Rob, MarR, and MarA, which shapes a hierarchy regulating 22 structural genes, many of them regulated by two or more regulators, involved in the response of *E. coli* to stress from antibiotics, organic compounds, mechanical, oxidative, and xenobiotics (https://abasy.ccg.unam.mx/module?regnetid=511145_v2020_sRDB18-13_eStrong&class=39.2).
Therefore, the different organizational layers shape the modular hierarchy of the gene regulatory organization as gene < operon < regulon < concilion < modulon. As we ascend in this hierarchy network complexity increases whereas functional and organism specificity decrease (**Figure 1A**, bottom right). This introduces at least two new problems for the study of genetic organization: 1) how a concilion can be identified, and 2) how the hierarchy governing these different genetic levels can be inferred.

UNRAVELING THE COMMON FUNCTIONAL ARCHITECTURE AND SYSTEM ELEMENTS OF GLOBAL GRNS

Studying the system dynamics of large-scale regulatory networks is challenging. Using a standard differential equations model to study the evolution in time of a system having thousands of interactions renders the model prohibitively complex. Moreover, despite the large availability of genomic data, incomplete knowledge of the system also hinders this goal (e.g., the poor availability of kinetic parameters). Therefore, as system complexity increases less detail must be included in the model (Bornholdt, 2005). On the other hand, the study of the system organization is fundamental as it constrains the possible dynamic outcomes (Ruklisa et al., 2019). Traditionally, one is interested in those genes responding to a particular condition, while this is interesting to study a specific response it is just an instantaneous

[1]Reprinted from Journal of Biotechnology, 161:3, Julio A. Freyre-González, Luis G. Treviño-Quintanilla, Ilse A. Valtierra-Gutiérrez, Rosa María Gutiérrez-Ríos, José A. Alonso-Pavón, Prokaryotic regulatory systems biology: Common principles governing the functional architectures of *Bacillus subtilis* and *Escherichia coli* unveiled by the natural decomposition approach, 278-286, Copyright (2012), with permission from Elsevier.

snapshot of the system. The combinatorial nature of gene expression requires many individual condition-specific subnetworks (akin to picture snapshots) merged into a single network model to capture all the possible dynamic trajectories, in the way a long-exposure photo does (**Figure 1B**). This global network model is not a static representation of regulation, contrary to the specific-condition network. Instead, it embeds all the potential regulations that can occur, forming a regulatory landscape by constraining the large number of organizations a regulatory network could potentially have.

Curation efforts have allowed the reconstruction of updated regulatory networks for many organisms, alleviating the large-scale study of the architecture of regulatory networks. Further curation can help to improve the current reconstructions and even increase the number of organisms with an available reconstructed regulatory network. However, the massive curation of regulatory networks is limited by competitive funding with short grant cycles, which renders long-term funding, if available, uncertain, although alternative subscription-based funding models have been proposed (Reiser et al., 2016). Recently, Abasy Atlas (https://abasy.ccg.unam.mx) has gathered the largest collection of disambiguated and homogenized regulatory networks with experimentally validated interactions (Ibarra-Arellano et al., 2016). Such networks cover 42 bacteria distributed in nine species, including historical snapshots of the regulatory network reconstruction of some organisms at different stages of curation (Escorcia-Rodriguez et al., 2020). The construction of Abasy Atlas has exposed the poor knowledge we have about regulation in bacteria as only roughly 10% of the organisms in Abasy Atlas have a reconstructed regulatory network with interaction completeness > 65%. This statistic is based on a recent model developed to quantify the total number of interactions that the regulatory network of an organism will have according to its genome size (Campos and Freyre-Gonzalez, 2019; Escorcia-Rodriguez et al., 2020). This interaction completeness model is implemented and available in Abasy Atlas since version 2.2. Regulatory networks deposited in Abasy Atlas include different types of regulations (e.g., protein-DNA, small RNAs, and protein-protein interactions). Abasy Atlas also provides the system elements identified by the natural decomposition approach (NDA) that compose the functional architecture of a regulatory network.

The NDA leverages the global structure of a regulatory network to define mathematical diagnostic criteria and an algorithm to identify these system elements by the controlled decomposition of a network (Freyre-Gonzalez et al., 2008; Freyre-Gonzalez and Trevino-Quintanilla, 2010; Freyre-Gonzalez et al., 2012; Ibarra-Arellano et al., 2016; Freyre-Gonzalez and Tauch, 2017; Escorcia-Rodriguez et al., 2020; 2021). First, the κ-value is computed as the solution $\left(\sqrt[\alpha+1]{\alpha\gamma}\cdot k_{\mathrm{out}_{\max}}\right)$ to the equation $\mathrm{d}C(k_{\mathrm{out}})/\mathrm{d}k_{\mathrm{out}} = -1$, where $C(k_{\mathrm{out}}) = \gamma k_{\mathrm{out}}^{-\alpha}$ is the clustering coefficient distribution of a GRN as a function of the out-connectivity (k_{out}) and is obtained by robust least-squares fitting. The global regulators are identified as those having out-connectivity > κ-value. The global regulators and their interactions are removed from the network to naturally reveal the modules (remaining connected subgraphs) and the basal machinery (disconnected nodes). Intermodular genes are identified as structural genes (nodes having zero out-connectivity ($k_{\mathrm{out}} = 0$) and therefore no coding for regulators) being controlled by different modules and then integrating disparate physiological responses. For further details on the NDA methodology, please see **Figures 1**, **2** in both (Ibarra-Arellano et al., 2016; Freyre-Gonzalez and Tauch, 2017). Sensitivity analyses have shown that the global regulators are the most robust to network incompleteness, whereas the intermodular genes are the most labile. By focusing on the modular and basal machinery genes, it has been observed that the NDA is highly robust to incompleteness in the set of interactions and more labile to incompleteness in the set of genes. This suggests that NDA predictions from GRNs having high network genomic coverage are quite reliable (Freyre-Gonzalez and Tauch, 2017). These observations have been supported by analyzing historical snapshots of the *E. coli* GRN (Escorcia-Rodriguez et al., 2020). Additionally, an assessment of the NDA predictions obtained by using three network models of the *C. glutamicum* GRN with different confidence degrees, including the addition of small RNAs, and an analysis of historical snapshots, have also confirmed these observations (Escorcia-Rodriguez et al., 2021).

All together global regulators, modules comprising modular genes, basal machinery genes, and intermodular genes compose a non-pyramidal, three-tier, diamond-like hierarchy common to all the organisms in Abasy Atlas (**Figure 2B**). The diamond-like nature of the functional architecture follows from the asymmetry in the number of genes composing each layer. The coordination layer comprises roughly 1% of the genes in the network, whereas the processing layer, composed of modular and basal machinery genes, accounts for about 90%, and the integration layer comprises roughly 9%. The global regulators (coordination layer) modulate the expression of genes belonging to the two lower layers (processing and integration), whereas some feedback could occasionally occur between the processing and coordination layers. Modules identified by the NDA can be concilions or regulons, but neither modulons nor single operons. Nevertheless, modulons globally coordinate modules. Basal machinery genes account for the cell's housekeeping functions and are controlled only by global regulators (Freyre-Gonzalez et al., 2012). Each module is responsible for a specific different function, whose combinatorial expression allows the cell to cope with a variety of environments. Remarkably, the NDA revealed that modules are locally independent meaning that there is no cross-regulation between them (Freyre-Gonzalez et al., 2012). Global regulators only coordinate the modules, and intermodular genes integrate some of their responses. Intermodular genes compose the integration layer. They were first identified by the NDA, they integrate, at the promoter level, the response of functionally disparate modules, and they thus enable the cell to cope with complex environments such as the assimilation of nitrogen under acidic conditions (**Figure 2C**) (Freyre-Gonzalez and Trevino-Quintanilla, 2010).

An alternative approach is to study expression data to elucidate the underlying network structures governing gene

expression (Saelens et al., 2018). Recent works have applied independent component analysis (ICA) to transcription datasets to unravel the signals that govern gene expression in *E. coli*, *S. aureus*, and *B. subtilis* (Sastry et al., 2019; Rychel et al., 2021). The analysis produces a series of so-called iModulons (unrelated to the traditional term modulon, see above), a group of genes that are governed by a certain signal. This signal in many cases can be assigned to a certain regulator, based on biological knowledge. A gene can be included in more than one iModulon and some iModulons are assigned to more than one regulator, which is consistent with the existence of complex regulons. This analysis partially reconstructs some of the known regulons of the network and even aids in predicting new regulatory interactions.

DEALING WITH GRNS INCOMPLETENESS

From the perspective of rational synthetic biology, the top-down approach can be applied to identify disposable components in an organism using a global GRN (Lastiri-Pancardo et al., 2020). So far, not even the model organisms in gene regulation have a complete experimentally supported GRN (Escorcia-Rodriguez et al., 2020) because of the time and resource consumption needed for experimental validation and curation. Therefore, network inference is currently one of the best alternatives to reconstructing complete GRNs. However, it is a still-going challenge that, on one hand, has been approached through a plethora of transcriptomics-based strategies ranging from mechanistic models to machine learning, all of them with modest to poor results (Marbach et al., 2012). Network inference based on the identification of regulatory binding sites has performed significantly better (Zorro-Aranda et al., 2022), but it requires a prior network for its application. One way to deal with this limitation is to transfer regulatory information from one organism to another (Alkema et al., 2004). Nevertheless, this approach requires the organisms to be similar enough so the interactions are conserved (McCue et al., 2002; Escorcia-Rodriguez et al., 2021), and prior regulatory information for the source organism is still required. Inferences based on gene expression data have also benefited from the integration of biological information. For instance, the pre-selection of transcription factors (TFs) from experimental data constrains the number of potential inferences, and the application of structural properties of biological GRNs improves the assessment of the predictions (Zorro-Aranda et al., 2022). Other works have also shown improvements in the inference of regulatory networks integrating multiple omics data (Cheng et al., 2011; Banf and Rhee, 2017) and network structure (Castro et al., 2019). This suggests that the integration of biological data and network structure might be the approach to pursue in the inference of GRNs.

There is no straightforward nor standard way to infer a global regulatory network. A few precomputed inferences based on sequence or transcriptomics are scattered across the literature and organism-specific databases (Galan-Vasquez et al., 2020; Parise et al., 2020). Most of these inferences come from different approaches making it difficult to assess them. Besides, for the organisms with transcriptomic data, we need to gather and normalize the data to apply one of the top-ranking tools in previous assessments (Marbach et al., 2012). There exist databases hosting inferences of regulatory networks based on regulatory binding sites [e.g., PRECISE (Novichkov et al., 2013)]. However, these predictions have not been systematically assessed. We need to standardize the benchmarking of network inference tools with biological datasets and GRN gold standards used as reference. This way, we could keep pace with the rate of emerging methodologies. Moreover, the incompleteness of the GRN gold standards hinders proper assessment of inferred networks as actual true positive interactions are incorrectly labeled as false positive if they are not part of the current gold standard. We can leverage the constrained space for structural properties found in biological GRNs (Campos and Freyre-Gonzalez, 2019; Escorcia-Rodriguez et al., 2021) to verify if the inferred networks have similar properties.

Once we know the inferred networks behave as the biological ones, we can study their functional architecture and system-level components (Freyre-Gonzalez et al., 2012). Although the diamond-like structure has been found across all the organisms in Abasy Atlas, the system-level conservation has been quantitatively evaluated only between *E. coli* and *B. subtilis* (Freyre-Gonzalez et al., 2012), and *Corynebacterium glutamicum* and *Streptomyces coelicolor* (Zorro-Aranda et al., 2022). Future work assessing the conservation across all the available organisms and the robustness of the node classification to network incompleteness would shed light on the missing interactions for incomplete networks and their hierarchical role in the global network.

CONSISTENCY OF GRNS: CORRELATION DOES NOT IMPLY CAUSATION

The consistency between GRNs and expression data has been previously studied by assuming a causal effect between the expression of the TFs and their target genes (TGs). Recent studies using expression data in *E. coli* and *C. glutamicum* have assessed this causal effect by using correlations to show a weak correlation of the known regulatory TF-TG pairs compared to all the possible random pairs as background (Larsen et al., 2019; Parise et al., 2021). Moreover, repressor interactions were associated with a positive correlation, rather than the expected negative correlation. This apparent inconsistency between GRNs and expression data may be explained by some molecular factors that cause known TF-TG pairs not to correlate well, e.g., the time delay between the stimulus and the regulatory response or TFs not being in their allosteric active configuration (Yu et al., 2003; Maier et al., 2009; Ghazalpour et al., 2011). Thus, we should not attempt to invalidate, through correlations of high-throughput expression data, reconstructed GRNs that are the result of experiments showing the physical binding of a TF to a DNA binding site. Further, expression data might capture false positive interactions and lead to an inherently noisy reconstruction of GRNs because found interactions are based on correlations and not necessarily causal.

Instead, an alternative approach, not yet reported, is to assess consistency within expression data considering the GRN architecture and organization. The functional architecture found in bacterial regulatory networks by the NDA (Escorcia-Rodriguez et al., 2020) (**Figure 2B**) proposes a robust partitioning of the network into physiologically correlated gene clusters and specific interaction roles for each regulatory interaction. This partitioning of the network may allow finding pairs of expressed genes that are significantly co-expressed across conditions by removing the noise in the previously unstructured set of interactions using a properly structured background. As mentioned above, expression data have been analyzed using ICA yielding significant biological results and partially reconstructing known regulons (Sastry et al., 2019; Rychel et al., 2021). This would be entirely impossible if expression data were completely inconsistent with the known structure of GRNs.

INTEGRATING QUANTITATIVE INFORMATION INTO NETWORK REPRESENTATIONS OF GENE REGULATION

Weighted gene co-expression networks have been widely and successfully used to identify biologically relevant subgraphs, outperforming approaches based solely on network structure (Li et al., 2011; Niu et al., 2019; Farhadian et al., 2021). Perhaps including quantitative information in the network could aid structure-based approaches, such as the NDA, in discovering relevant modules. Optimally partitioning the network into subgraphs comprising strong interactions could also help identify sections of the network that can be modeled independently.

Research on GRNs has focused mainly on structural aspects, leaving out any quantitative information about how a certain regulator affects the expression of its targets. Although the modeling of gene expression dynamics based on Hill kinetics and differential equations becomes prohibitively complex as network size increases, simpler models could perhaps yield interesting information about how GRNs are globally organized. A first approach could be representing the network as a weighted graph, i.e., having each edge on the network assigned a certain weight that represents the strength of the interaction. The sole definition of what this strength would be (the affinity of the TF to its binding site, the TF-TG correlation, or some other measure) is itself a challenge as it is inherently related to the data used to quantify this information.

Integrating quantitative information into the network could yield valuable insights into the dynamical stability of the system as a whole or provide parameters with which to model small circuits within the network. Gene regulatory networks seem to be constrained in their density, tending towards lower values as network nodes increase, following a power law (Campos and Freyre-Gonzalez, 2019). In that study, the authors discuss that this restriction may stem from the necessity of dynamic stability, as predicted by the May-Wigner theorem (Gardner and Ashby, 1970; May, 1972). A 2018 study on the dynamics of phage λ demonstrated that, although some of its behavior can be solely explained by the structure of its network, the relative ordering of transcription factor binding site affinities determined modified behaviors of the attractors of the system (i.e., the set of the stable states the system arrives after perturbation) (Ruklisa et al., 2019). Advancing global network models from the purely qualitative to the quantitative are surely one of the ongoing challenges of biological network science, and essential to furthering our understanding of dynamic living systems.

EVOLUTION OF GRNS FROM A SYSTEM-LEVEL PERSPECTIVE

In a seminal paper in 1962, Herbert Simon proposed the idea that the evolution of a complex structure from simple elements must proceed through a hierarchy of potential stable subassemblies (Simon, 1962). In his parable of the two watchmakers, Simon argues that these hierarchical structures will evolve faster than non-hierarchical counterparts of similar size. Consequently, in the study of the evolution of complex structures such as complex biological networks, it is imperative to adopt an approach that considers how these potential stable subassemblies have played a role in their evolution. These subassemblies could be operons, regulons, concilions, or modulons in GRNs, all of them collectively referred to as systems hereafter. Therefore, for the study of the evolution of GRNs, we need a system-level approach.

Previous evolutionary studies have focused on the effect of gene duplication and horizontal gene transfer in the evolution of GRNs but without taking into account the network organization and how these mechanisms have given rise to its functional architecture (Madan Babu and Teichmann, 2003; Teichmann and Babu, 2004; Price et al., 2008). Further studies have assessed the conservation and evolution of GRNs by using networks inferred through orthology (Madan Babu et al., 2006) or biding sites prediction (Gonzalez Perez et al., 2008). All these advances have been properly summarized (Janga and Collado-Vides, 2007; Babu, 2010). Recently, some studies on eukaryotes have aimed to study how modularity evolves in developmental GRNs by using gene co-expression data (Peter and Davidson, 2011; Verd et al., 2019) or completely theoretical approaches (Espinosa-Soto and Wagner, 2010; Espinosa-Soto, 2018). An interesting study focuses on exploring the evolution of non-developmental GRNs (Defoort et al., 2018). By using genomic phylostratigraphy (Domazet-Loso et al., 2007), the authors explore the evolution at the level of small regulatory subgraphs (i.e., network motifs) of a mix of different types of GRNs in two eukaryotic organisms. Whereas this is an interesting study, the question of how evolution shapes the systems composing a GRN and its functional architecture is still an open question.

The lack of reliable methodologies to identify the system components integrating a GRN and the low completeness and standardization of GRNs have limited the study of its evolution from a systems perspective. Previous analyses have shown that the system elements proposed by the NDA are poorly conserved and that their evolution is possibly driven by evolutionary convergence (Freyre-Gonzalez et al., 2012). The recent

availability of databases providing homogenized and standardized GRNs (Escorcia-Rodriguez et al., 2020), including the modules and system-level elements composing each GRN, provides the basis to explore how these systems have been shaped by evolution and whether stable subassemblies have arisen during the evolution of the currently known systems.

DISCUSSION

Without the study of the basic principles governing cell systems, it will be impossible for synthetic biology to become a true biological engineering discipline as has been defined by a European NEST (New and Emerging Science and Technology) High-Level Expert Group (European Commission, 2005; Pei et al., 2012) and repeatedly elsewhere (Serrano, 2007; Cheng and Lu, 2012; Bartley et al., 2017; Hanczyc, 2020). Even if the aim of being a true biological engineering discipline becomes elusive, the study of these fundamental principles is necessary to improve our basic understanding of biological complex systems (Schwille, 2011). All the themes presented in this paper are interconnected. Therefore, advance in one area affects the others. For example, having a model that describes the global organization of GRNs helps to delimit and guide their study in dynamics and evolution, as well as improve the understanding of their consistency with expression data. In turn, improvements in these subjects help to refine this model of the global network organization. Furthermore, all these topics together help to infer more and better GRNs incrementally improving our understanding of genomic regulation. Overall, during the last decade, some basic principles governing the still incomplete GRNs of a few organisms have been found. It is time to continue the research of these basic principles of biological complex networks to contribute to achieving rational synthetic biology.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in https://abasy.ccg.unam.mx and the article/Supplementary Materials, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JF-G Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Visualization, Writing—original draft, Writing—review and editing. JE-R, LG-M, JM-V Investigation, Visualization, Writing—original draft, Writing—review and editing. CT-F, AZ-A Investigation, Writing—original draft.

FUNDING

This work was supported by the Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT-UNAM) [IN202421 to JF-G].

ACKNOWLEDGMENTS

We thank Felipe B. and Victor M. Barrera Rivera for their support with long-exposure photography used in **Figure 1B**. JE-R is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM). He received fellowship 959406 from CONACYT. We apologize to all those colleagues whose research was not cited due to space constraints.

REFERENCES

Abil, Z., and Danelon, C. (2020). Roadmap to Building a Cell: An Evolutionary Approach. *Front. Bioeng. Biotechnol.* 8, 927. doi:10.3389/fbioe.2020.00927

Alkema, W. B. L., Lenhard, B., and Wasserman, W. W. (2004). Regulog Analysis: Detection of Conserved Regulatory Networks across Bacteria: Application to *Staphylococcus aureus*. *Genome Res.* 14 (7), 1362–1373. doi:10.1101/gr.2242604

Babu, M. M. (2010). Structure, Evolution and Dynamics of Transcriptional Regulatory Networks. *Biochem. Soc. Trans.* 38 (5), 1155–1178. doi:10.1042/BST0381155

Banf, M., and Rhee, S. Y. (2017). Enhancing Gene Regulatory Network Inference through Data Integration with Markov Random Fields. *Sci. Rep.* 7, 41174. doi:10.1038/srep41174

Bartley, B. A., Kim, K., Medley, J. K., and Sauro, H. M. (2017). Synthetic Biology: Engineering Living Systems from Biophysical Principles. *Biophysical J.* 112 (6), 1050–1058. doi:10.1016/j.bpj.2017.02.013

Bornholdt, S. (2005). Less Is More in Modeling Large Genetic Networks. *Science* 310 (5747), 449–451. doi:10.1126/science.1119959

Campos, A. I., and Freyre-González, J. A. (2019). Evolutionary Constraints on the Complexity of Genetic Regulatory Networks Allow Predictions of the Total Number of Genetic Interactions. *Sci. Rep.* 9 (1), 3618. doi:10.1038/s41598-019-39866-z

Castro, D. M., de Veaux, N. R., Miraldi, E. R., and Bonneau, R. (2019). Multi-study Inference of Regulatory Networks for More Accurate Models of Gene Regulation. *PLoS Comput. Biol.* 15 (1), e1006591. doi:10.1371/journal.pcbi.1006591

Cheng, A. A., and Lu, T. K. (2012). Synthetic Biology: an Emerging Engineering Discipline. *Annu. Rev. Biomed. Eng.* 14, 155–178. doi:10.1146/annurev-bioeng-071811-150118

Cheng, C., Yan, K.-K., Hwang, W., Qian, J., Bhardwaj, N., Rozowsky, J., et al. (2011). Construction and Analysis of an Integrated Regulatory Network Derived from High-Throughput Sequencing Data. *PLoS Comput. Biol.* 7 (11), e1002190. doi:10.1371/journal.pcbi.1002190

Chubiz, L. M., Glekas, G. D., and Rao, C. V. (2012). Transcriptional Cross Talk within Themar-Sox-robRegulon in *Escherichia coli* Is Limited to therobandmarRABOperons. *J. Bacteriol.* 194 (18), 4867–4875. doi:10.1128/JB.00680-12

de Daruvar, A., Collado-Vides, J., and Valencia, A. (2002). Analysis of the Cellular Functions of *Escherichia coli* Operons and Their Conservation in Bacillus Subtilis. *J. Mol. Evol.* 55 (2), 211–221. doi:10.1007/s00239-002-2317-1

Defoort, J., Van de Peer, Y., and Vermeirssen, V. (2018). Function, Dynamics and Evolution of Network Motif Modules in Integrated Gene Regulatory Networks of Worm and Plant. *Nucleic Acids Res.* 46 (13), 6480–6503. doi:10.1093/nar/gky468

Domazet-Lošo, T., Brajković, J., and Tautz, D. (2007). A Phylostratigraphy Approach to Uncover the Genomic History of Major Adaptations in Metazoan Lineages. *Trends Genet.* 23 (11), 533–539. doi:10.1016/j.tig.2007.08.014

Escorcia-Rodríguez, J. M., Tauch, A., and Freyre-González, J. A. (2020). Abasy Atlas v2.2: The Most Comprehensive and Up-To-Date Inventory of Meta-

Curated, Historical, Bacterial Regulatory Networks, Their Completeness and System-Level Characterization. *Comput. Struct. Biotechnol. J.* 18, 1228–1237. doi:10.1016/j.csbj.2020.05.015

Escorcia-Rodríguez, J. M., Tauch, A., and Freyre-González, J. A. (2021). Corynebacterium Glutamicum Regulation beyond Transcription: Organizing Principles and Reconstruction of an Extended Regulatory Network Incorporating Regulations Mediated by Small RNA and Protein-Protein Interactions. *Microorganisms* 9 (7), 1395. doi:10.3390/microorganisms9071395

Espinosa-Soto, C. (2018). On the Role of Sparseness in the Evolution of Modularity in Gene Regulatory Networks. *PLoS Comput. Biol.* 14 (5), e1006172. doi:10.1371/journal.pcbi.1006172

Espinosa-Soto, C., and Wagner, A. (2010). Specialization Can Drive the Evolution of Modularity. *PLoS Comput. Biol.* 6 (3), e1000719. doi:10.1371/journal.pcbi.1000719

European Commission (2005). *Synthetic Biology: Applying Engineering to Biology : Report of a NEST High-Level Expert Group*. Office for Official Publications of the European Communities. London, UK: Luxembourg.

Farhadian, M., Rafat, S. A., Panahi, B., and Mayack, C. (2021). Weighted Gene Co-expression Network Analysis Identifies Modules and Functionally Enriched Pathways in the Lactation Process. *Sci. Rep.* 11 (1), 2367. doi:10.1038/s41598-021-81888-z

Freyre-Gonzalez, J. A., Alonso-Pavon, J. A., Trevino-Quintanilla, L. G., and Collado-Vides, J. (2008). Functional Architecture of *Escherichia coli*: New Insights provided by a Natural Decomposition Approach. *Genome Biol.* 9 (10), R154. doi:10.1186/gb-2008-9-10-r154

Freyre-González, J. A., and Tauch, A. (2017). Functional Architecture and Global Properties of the Corynebacterium Glutamicum Regulatory Network: Novel Insights from a Dataset with a High Genomic Coverage. *J. Biotechnol.* 257, 199–210. doi:10.1016/j.jbiotec.2016.10.025

Freyre-Gonzalez, J. A., and Trevino-Quintanilla, L. G. (2010). Analyzing Regulatory Networks in Bacteria. *Nat. Educ.* 3 (9). 1.

Freyre-González, J. A., Treviño-Quintanilla, L. G., Valtierra-Gutiérrez, I. A., Gutiérrez-Ríos, R. M., and Alonso-Pavón, J. A. (2012). Prokaryotic Regulatory Systems Biology: Common Principles Governing the Functional Architectures of Bacillus Subtilis and *Escherichia coli* Unveiled by the Natural Decomposition Approach. *J. Biotechnol.* 161 (3), 278–286. doi:10.1016/j.jbiotec.2012.03.028

Galán-Vásquez, E., Luna-Olivera, B. C., Ramírez-Ibáñez, M., and Martínez-Antonio, A. (2020). RegulomePA: A Database of Transcriptional Regulatory Interactions in Pseudomonas aeruginosa PAO1, *Database (Oxford)*. 2020: baaa106. doi:10.1093/database/baaa106

Gardner, M. R., and Ashby, W. R. (1970). Connectance of Large Dynamic (Cybernetic) Systems: Critical Values for Stability. *Nature* 228 (5273), 784. doi:10.1038/228784a0

Ghazalpour, A., Bennett, B., Petyuk, V. A., Orozco, L., Hagopian, R., Mungrue, I. N., et al. (2011). Comparative Analysis of Proteome and Transcriptome Variation in Mouse. *PLoS Genet.* 7 (6), e1001393. doi:10.1371/journal.pgen.1001393

González Pérez, A. D., González González, E., Espinosa Angarica, V., Vasconcelos, A. T. R., and Collado-Vides, J. (2008). Impact of Transcription Units Rearrangement on the Evolution of the Regulatory Network of Gamma-Proteobacteria. *BMC Genomics* 9, 128. doi:10.1186/1471-2164-9-128

Gottesman, S. (1984). Bacterial Regulation: Global Regulatory Networks. *Annu. Rev. Genet.* 18, 415–441. doi:10.1146/annurev.ge.18.120184.002215

Gutiérrez-Ríos, R. M., Rosenblueth, D. A., Loza, J. A., Huerta, A. M., Glasner, J. D., Blattner, F. R., et al. (2003). Regulatory Network of *Escherichia coli*: Consistency between Literature Knowledge and Microarray Profiles. *Genome Res.* 13 (11), 2435–2443. doi:10.1101/gr.1387003

Hanczyc, M. M. (2020). Engineering Life: A Review of Synthetic Biology. *Artif. Life* 26 (2), 260–273. doi:10.1162/artl_a_00318

Hartwell, L. H., Hopfield, J. J., Leibler, S., and Murray, A. W. (1999). From Molecular to Modular Cell Biology. *Nature* 402 (6761 Suppl. l), C47–C52. doi:10.1038/35011540

Ibarra-Arellano, M. A., Campos-González, A. I., Treviño-Quintanilla, L. G., Tauch, A., and Freyre-González, J. A. (2016). Abasy Atlas: A Comprehensive Inventory of Systems, Global Network Properties and Systems-Level Elements across Bacteria,*Database* 2016. baw089. doi:10.1093/database/baw089

Iuchi, S., and Lin, E. C. (1988). arcA (Dye), a Global Regulatory Gene in *Escherichia coli* Mediating Repression of Enzymes in Aerobic Pathways. *Proc. Natl. Acad. Sci. U.S.A.* 85 (6), 1888–1892. doi:10.1073/pnas.85.6.1888

Jacob, F., Perrin, D., Sanchez, C., and Monod, J. (1960). Operon: a Group of Genes with the Expression Coordinated by an Operator. *C R. Hebd. Seances Acad. Sci.* 250, 1727–1729.

Jacob, F., and Monod, J. (1961). Genetic Regulatory Mechanisms in the Synthesis of Proteins. *J. Mol. Biol.* 3, 318–356. doi:10.1016/s0022-2836(61)80072-7

Janga, S. C., and Collado-Vides, J. (2007). Structure and Evolution of Gene Regulatory Networks in Microbial Genomes. *Res. Microbiol.* 158 (10), 787–794. doi:10.1016/j.resmic.2007.09.001

Larsen, S. J., Röttger, R., Schmidt, H. H. H. W., and Baumbach, J. (2019). E. Coligene Regulatory Networks Are Inconsistent with Gene Expression Data. *Nucleic Acids Res.* 47 (1), 85–92. doi:10.1093/nar/gky1176

Lastiri-Pancardo, G., Mercado-Hernández, J. S., Kim, J., Jiménez, J. I., and Utrilla, J. (2020). A Quantitative Method for Proteome Reallocation Using Minimal Regulatory Interventions. *Nat. Chem. Biol.* 16 (9), 1026–1033. doi:10.1038/s41589-020-0593-y

Laub, M. T., McAdams, H. H., Feldblyum, T., Fraser, C. M., and Shapiro, L. (2000). Global Analysis of the Genetic Network Controlling a Bacterial Cell Cycle. *Science* 290 (5499), 2144–2148. doi:10.1126/science.290.5499.2144

Li, W., Liu, C.-C., Zhang, T., Li, H., Waterman, M. S., and Zhou, X. J. (2011). Integrative Analysis of Many Weighted Co-expression Networks Using Tensor Computation. *PLoS Comput. Biol.* 7 (6), e1001106. doi:10.1371/journal.pcbi.1001106

Maas, W. K., and Clark, A. J. (1964). Studies on the Mechanism of Repression of Arginine Biosynthesis in *Escherichia coli*. *J. Mol. Biol.* 8, 365–370. doi:10.1016/s0022-2836(64)80200-x

Madan Babu, M., Teichmann, S. A., and Aravind, L. (2006). Evolutionary Dynamics of Prokaryotic Transcriptional Regulatory Networks. *J. Mol. Biol.* 358 (2), 614–633. doi:10.1016/j.jmb.2006.02.019

Madan Babu, M., and Teichmann, S. A. (2003). Evolution of Transcription Factors and the Gene Regulatory Network in *Escherichia coli*. *Nucleic Acids Res.* 31 (4), 1234–1244. doi:10.1093/nar/gkg210

Maier, T., Güell, M., and Serrano, L. (2009). Correlation of mRNA and Protein in Complex Biological Samples. *FEBS Lett.* 583 (24), 3966–3973. doi:10.1016/j.febslet.2009.10.036

Marbach, D., Costello, J. C., Costello, J. C., Küffner, R., Vega, N. M., Prill, R. J., et al. (2012). Wisdom of Crowds for Robust Gene Network Inference. *Nat. Methods* 9 (8), 796–804. doi:10.1038/nmeth.2016

Martin, R. G., Bartlett, E. S., Rosner, J. L., and Wall, M. E. (2008). Activation of the *Escherichia coli* marA/soxS/rob Regulon in Response to Transcriptional Activator Concentration. *J. Mol. Biol.* 380 (2), 278–284. doi:10.1016/j.jmb.2008.05.015

Martin, R. G., Gillette, W. K., Rhee, S., and Rosner, J. L. (1999). Structural Requirements for Marbox Function in Transcriptional Activation of Mar/sox/rob Regulon Promoters in *Escherichia coli*: Sequence, Orientation and Spatial Relationship to the Core Promoter. *Mol. Microbiol.* 34 (3), 431–441. doi:10.1046/j.1365-2958.1999.01599.x

May, R. M. (1972). Will a Large Complex System Be Stable? *Nature* 238 (5364), 413–414. doi:10.1038/238413a0

McCue, L. A., Thompson, W., Carmack, C. S., and Lawrence, C. E. (2002). Factors Influencing the Identification of Transcription Factor Binding Sites by Cross-Species Comparison. *Genome Res.* 12 (10), 1523–1532. doi:10.1101/gr.323602

Niu, X., Zhang, J., Zhang, L., Hou, Y., Pu, S., Chu, A., et al. (2019). Weighted Gene Co-expression Network Analysis Identifies Critical Genes in the Development of Heart Failure after Acute Myocardial Infarction. *Front. Genet.* 10, 1214. doi:10.3389/fgene.2019.01214

Novichkov, P. S., Kazakov, A. E., Ravcheev, D. A., Leyn, S. A., Kovaleva, G. Y., Sutormin, R. A., et al. (2013). RegPrecise 3.0 - A Resource for Genome-Scale Exploration of Transcriptional Regulation in Bacteria. *BMC Genomics* 14, 745. doi:10.1186/1471-2164-14-745

Osbourn, A. E., and Field, B. (2009). Operons. *Cell. Mol. Life Sci.* 66 (23), 3755–3775. doi:10.1007/s00018-009-0114-3

Parise, D., Parise, M. T. D., Kataka, E., Kato, R. B., List, M., Tauch, A., et al. (2021). On the Consistency between Gene Expression and the Gene Regulatory

Network of Corynebacterium Glutamicum. *Netw. Syst. Med.* 4 (1), 51–59. doi:10.1089/nsm.2020.0014

Parise, M. T. D., Parise, D., Kato, R. B., Pauling, J. K., Tauch, A., Azevedo, V. A. d. C., et al. (2020). CoryneRegNet 7, the Reference Database and Analysis Platform for Corynebacterial Gene Regulatory Networks. *Sci. Data* 7 (1), 142. doi:10.1038/s41597-020-0484-9

Pei, L., Gaisser, S., and Schmidt, M. (2012). Synthetic Biology in the View of European Public Funding Organisations. *Public Underst. Sci.* 21 (2), 149–162. doi:10.1177/0963662510393624

Peter, I. S., and Davidson, E. H. (2011). Evolution of Gene Regulatory Networks Controlling Body Plan Development. *Cell* 144 (6), 970–985. doi:10.1016/j.cell.2011.02.017

Price, M. N., Dehal, P. S., and Arkin, A. P. (2008). Horizontal Gene Transfer and the Evolution of Transcriptional Regulation in *Escherichia coli*. *Genome Biol.* 9 (1), R4. doi:10.1186/gb-2008-9-1-r4

Reiser, L., Berardini, T. Z., Li, D., Muller, R., Strait, E. M., Li, Q., et al. (2016). Sustainable Funding for Biocuration: The Arabidopsis Information Resource (TAIR) as a Case Study of a Subscription-Based Funding Model,*Database* 2016. baw018. doi:10.1093/database/baw018

Ruklisa, D., Brazma, A., Cerans, K., Schlitt, T., and Viksna, J. (2019). Dynamics of Gene Regulatory Networks and Their Dependence on Network Topology and Quantitative Parameters - the Case of Phage λ. *BMC Bioinforma.* 20 (1), 296. doi:10.1186/s12859-019-2909-z

Rychel, K., Decker, K., Sastry, A. V., Phaneuf, P. V., Poudel, S., and Palsson, B. O. (2021). iModulonDB: a Knowledgebase of Microbial Transcriptional Regulation Derived from Machine Learning. *Nucleic Acids Res.* 49 (D1), D112–D120. doi:10.1093/nar/gkaa810

Saelens, W., Cannoodt, R., and Saeys, Y. (2018). A Comprehensive Evaluation of Module Detection Methods for Gene Expression Data. *Nat. Commun.* 9 (1), 1090. doi:10.1038/s41467-018-03424-4

Sastry, A. V., Gao, Y., Szubin, R., Hefner, Y., Xu, S., Kim, D., et al. (2019). The *Escherichia coli* Transcriptome Mostly Consists of Independently Regulated Modules. *Nat. Commun.* 10 (1), 5536. doi:10.1038/s41467-019-13483-w

Schwille, P. (2011). Bottom-up Synthetic Biology: Engineering in a Tinkerer's World. *Science* 333 (6047), 1252–1254. doi:10.1126/science.1211701

Serrano, L. (2007). Synthetic Biology: Promises and Challenges. *Mol. Syst. Biol.* 3, 158. doi:10.1038/msb4100202

Simon, H. A. (1962). The Architecture of Complexity. *Proc. Am. Philosophical Soc.* 106 (6), 467–482.

Teichmann, S. A., and Babu, M. M. (2004). Gene Regulatory Network Growth by Duplication. *Nat. Genet.* 36 (5), 492–496. doi:10.1038/ng1340

Verd, B., Monk, N. A., and Jaeger, J. (2019). Modularity, Criticality, and Evolvability of a Developmental Gene Regulatory Network. *Elife* 8. e42832. doi:10.7554/eLife.42832

Wall, M. E., Markowitz, D. A., Rosner, J. L., and Martin, R. G. (2009). Model of Transcriptional Activation by MarA in *Escherichia coli*. *PLoS Comput. Biol.* 5 (12), e1000614. doi:10.1371/journal.pcbi.1000614

Yu, H., Luscombe, N. M., Qian, J., and Gerstein, M. (2003). Genomic Analysis of Gene Expression Relationships in Transcriptional Regulatory Networks. *Trends Genet.* 19 (8), 422–427. doi:10.1016/S0168-9525(03)00175-6

Zorro-Aranda, A., Escorcia-Rodríguez, J. M., González-Kise, J. K., and Freyre-González, J. A. (2022). Curation, Inference, and Assessment of a Globally Reconstructed Gene Regulatory Network for Streptomyces Coelicolor. *Sci. Rep.* 12 (1), 2840. doi:10.1038/s41598-022-06658-x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Bioengineering and Biotechnology

REVIEW
published: 17 May 2022
doi: 10.3389/fbioe.2022.890479

Bioengineered Probiotics: Synthetic Biology Can Provide Live Cell Therapeutics for the Treatment of Foodborne Diseases

Karla Cristina P. Cruz†, *Laura O. Enekegho*† *and David T. Stuart* *

Department of Biochemistry, University of Alberta, Edmonton, AB, Canada

The rising prevalence of antibiotic resistant microbial pathogens presents an ominous health and economic challenge to modern society. The discovery and large-scale development of antibiotic drugs in previous decades was transformational, providing cheap, effective treatment for what would previously have been a lethal infection. As microbial strains resistant to many or even all antibiotic drug treatments have evolved, there is an urgent need for new drugs or antimicrobial treatments to control these pathogens. The ability to sequence and mine the genomes of an increasing number of microbial strains from previously unexplored environments has the potential to identify new natural product antibiotic biosynthesis pathways. This coupled with the power of synthetic biology to generate new production chassis, biosensors and "weaponized" live cell therapeutics may provide new means to combat the rapidly evolving threat of drug resistant microbial pathogens. This review focuses on the application of synthetic biology to construct probiotic strains that have been endowed with functionalities allowing them to identify, compete with and in some cases kill microbial pathogens as well as stimulate host immunity. Weaponized probiotics may have the greatest potential for use against pathogens that infect the gastrointestinal tract: *Vibrio cholerae*, *Staphylococcus aureus*, *Clostridium perfringens* and *Clostridioides difficile*. The potential benefits of engineered probiotics are highlighted along with the challenges that must still be met before these intriguing and exciting new therapeutic tools can be widely deployed.

Keywords: probiotic, synthetic biology, biosensors, metabolic engineering, *Vibrio cholerae*, *Clostridium perfringens*, *Staphylococcus aureus*, *Clostridioides difficile*

OPEN ACCESS

Edited by:
Shota Atsumi,
University of California, Davis, United States

Reviewed by:
Jingyu Chen,
China Agricultural University, China
Corina-Diana Ceapa,
Universidad Nacional Autónoma de México, Mexico

***Correspondence:**
David T. Stuart
dtstuart@ualberta.ca

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *06 March 2022*
Accepted: *29 April 2022*
Published: *17 May 2022*

Citation:
Cruz KCP, Enekegho LO and Stuart DT (2022) Bioengineered Probiotics: Synthetic Biology Can Provide Live Cell Therapeutics for the Treatment of Foodborne Diseases. Front. Bioeng. Biotechnol. 10:890479. doi: 10.3389/fbioe.2022.890479

1 INTRODUCTION

The discovery and application of antibiotic drugs is among the most significant accomplishments of medical science. Alexander Fleming's discovery of penicillin (Fleming, 1929) and subsequent discovery and development of multiple classes of natural product antibiotics have been transformational to modern society. These compounds have yielded cheap and effective treatments for diseases caused by common bacterial infections that would previously have proven fatal. The advent of effective antibiotic drugs has made it possible to survive complex surgical procedures like open heart surgery and organ transplants and extended the average human life-span (Riley, 2005; Kaviani et al., 2020). The benefits of readily available antibiotic drugs have extended into agriculture and aquaculture, making it possible to increase productivity of farmed

animals (Park et al., 1994; Patel et al., 2020). Indeed, the application of prophylactic antibiotics was recognized as growth promoting for farm animals in the 1940s and this has had a significant economic benefit to the agriculture industry (Moore et al., 1946; Graham et al., 2007).

The first widely employed antibiotics introduced in the early 1900s were synthetic compounds identified by Paul Ehrlich. The compound arsphenamine was one of the first, employed as a treatment for syphilis (Ehrlich, 1913; Gelpi et al., 2015). This was followed by the development of sulfonamides (Domagk, 1935; Bickel, 1988). A variety of other synthetic compounds were subsequently identified through systematic screening efforts and these have been previously reviewed (Hutchings et al., 2019). The synthetic antibiotic drugs were largely displaced by penicillin and the natural product antibiotics that were subsequently identified and developed (Jones et al., 1944; Bryer et al., 1948; Duggar, 1948; Walksman and Lechevalier, 1949; Newton and Abraham, 1955). Most of the major classes of antibiotic drugs were discovered during the period from the 1950s to the 1970s which was considered a "golden era" for antibiotic drug development (Nicolaou and Rigol, 2018). Many of these including streptomycin, neomycin and tetracycline remain in use today and virtually all widely used antibiotics either are or are derived from natural products produced by soil bacteria and fungi, a testament to the chemical diversity explored by these organisms in their natural environments (Waksman et al., 2010). The full realm of natural product and synthetic antibiotic drugs has been extensively reviewed (Durand et al., 2019; Hutchings et al., 2019). Since that time there has been a decline in the discovery of natural product antibiotics and an increased reliance on the modification of existing drug compounds to respond to the rising antibiotic resistance of pathogens (Olsufyeva and Yankovskaya, 2020; Acharya et al., 2022). High throughput screening of chemical libraries has met with limited success in identification and development of new antibiotic drugs, reinforcing how powerful a force natural selection has been in generating antibiotic natural products in microbial cells (Chopra et al., 2002; Durand et al., 2019).

The decline in discovery of new antibiotic drugs has become an issue of increasing concern reflecting a rise in the frequency of infections caused by pathogens resistant to most or even all of the currently used antimicrobial drugs. It has been estimated that antibiotic resistant bacterial infections were responsible for greater than one million deaths in 2019 (Murray et al., 2022). It has further been projected, that deaths directly attributable to antibiotic resistant infections could reach 10 million per year by 2050 (Murray et al., 2022). The growing occurrence of antibiotic resistance poses not only a global health threat but a potential economic crisis that could cost up to $1 trillion per year by 2050 (O'Neil, 2016).

Bacterial strains with resistance to virtually all currently employed antibiotic drugs have been identified (Chang et al., 2003). Widespread use and misuse of antibiotic drugs for both humans and animals has been touted as a key factor in the development of antibiotic resistance (Ventola, 2015; Pulingam et al., 2022). However, given that most antibiotic drugs are produced by bacteria and fungi resident in soil, it is not surprising that resistant strains exist in the environment independent of clinical or agricultural exposure to antibiotics (D'Costa et al., 2006; Larsen et al., 2022). The answer to this ongoing problem lies in developing processes to identify and produce new antibiotic drugs and therapies. These processes need to keep pace with the ability of pathogens to acquire resistance to these measures.

2 SYNTHETIC BIOLOGY APPROACHES TO ANTIMICROBIAL DRUG RESISTANCE

Advances in genomics technology have allowed for a near exponential increase in the number of new microbial genome sequences (Medema and Fischbach, 2015). This wealth of genome sequence data, coupled with powerful computational search capabilities, has revealed new biosynthetic gene clusters that may potentially encode natural product biosynthetic pathways (**Figure 1A**), (Sekurova et al., 2019). However, a significant portion of newly discovered gene clusters are silent under laboratory conditions (Blin et al., 2019). This has led to extensive efforts to identify conditions that can activate silent natural product pathways in the native organisms in the hopes of discovering novel antibiotic molecules (Gehrke et al., 2019; Tomm et al., 2019). While it is likely that new antibiotic compounds will continue to be revealed by this strategy, it is an untargeted process whose outcomes are difficult to predict.

Synthetic biology can be described as the application of engineering principles to biological systems. The key power of synthetic biology lies in the ability to program cells to produce desired compounds or perform specific tasks. Therefore, synthetic biology and metabolic engineering can offer tools and means to reveal or to generate new antibiotic compounds. An early example of this strategy was the generation of hybrid gene clusters from distinct *Streptomyces* species to yield the production of novel natural product compounds (Hopwood et al., 1985). More contemporary applications of metabolic engineering to the production of antibiotics include the use of biosensors combined with mutagenesis to induce and detect activation of a silent gene cluster for the production of the antibiotic coelimycin in *Streptomyces lividans* (Sekurova et al., 2021). Further directed strategies have taken advantage of genetic and genomic data to express known or predicted transcriptional regulators to induce the expression of silent gene clusters encoding mayamycin, warkmycin and chartreusin-like compounds or engineer promoter elements to into a phenazine biosynthesis gene cluster to yield new derivatives of phenazine-1-carboxilic acid (**Figure 1B**) (Saleh et al., 2012; Mingyar et al., 2021). The ability to manipulate the individual gene components of biosynthetic gene clusters or generate hybrid proteins from closely related enzymes with distinct specificities also adds to the potential for engineering and mixing genetic sequence elements to generate novel small molecules with antibiotic potential, in effect accelerating evolution that might happen through naturally occurring gene transfer and/or genetic recombination events (**Figure 1C**) (Mao D. et al., 2018; Awakawa et al., 2018; Zhang et al., 2019).

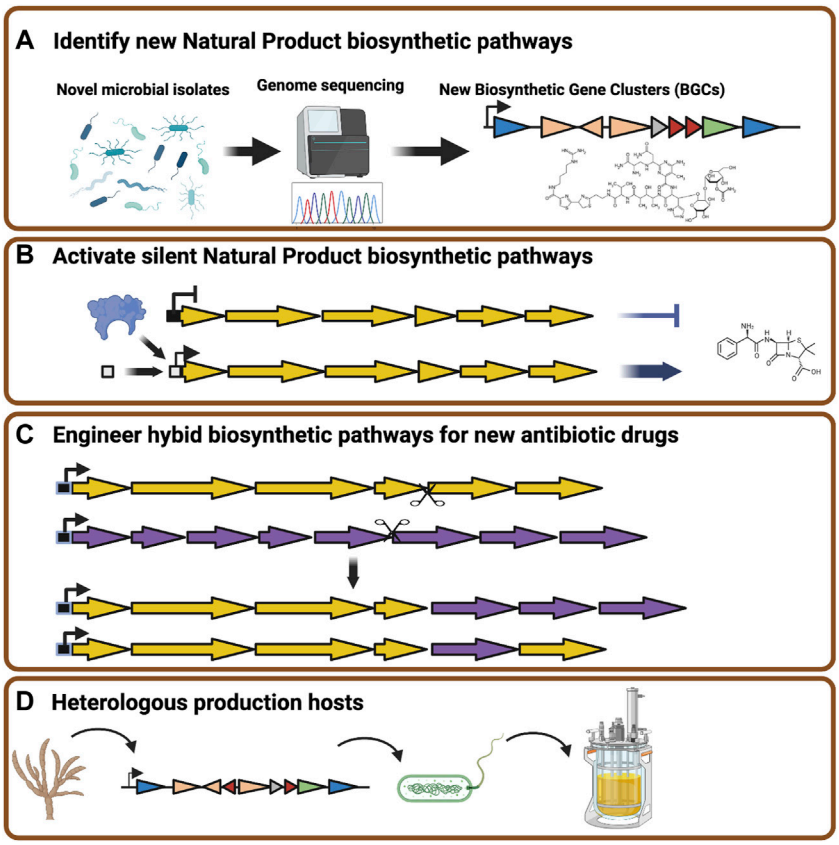


FIGURE 1 | Strategies to expand the diversity of natural product antibiotics. **(A)** Rapid genome sequencing combined with bioinformatic tools can be applied to a broad range of bacteria and fungi isolated from unusual environments to identify natural product biosynthetic gene clusters. **(B)** Genome engineering applied to native natural product producing organisms can activate silent biosynthetic gene clusters through modification of promoter elements or the introduction of transcriptional activators. **(C)** Synthetic biology and DNA assembly technologies allow for the assembly of hybrid gene clusters to synthesize novel or modified natural product antibiotic drugs. **(D)** Natural product biosynthetic gene clusters can be transplanted from native producers into heterologous host organisms modified to maximize product biosynthesis. Figure constructed using Biorender.

An alternative to engineering the native producer of antimicrobial natural products is to transplant the biosynthetic gene clusters into heterologous production hosts (**Figure 1D**). These approaches have become possible through the application of recombination mediated engineering (recombineering) and transformation-associated recombination (TAR) strategies (Yamanaka et al., 2014; Li et al., 2015). Expression of heterologous biosynthetic gene clusters in favorable production hosts has been further enhanced by our ability to rapidly synthesize codon optimized open reading frame sequences. This has allowed for the assembly and optimal expression of complex natural product pathways in manipulable microbial host organisms like *Saccharomyces cerevisiae* (Ro et al., 2006). The recreation of complex pathways in heterologous production hosts bypasses the challenge of culturing and engineering poorly characterized microbes, and even allows pathway synthesis based upon metagenomic data (Wu et al., 2019).

2.1 Probiotics as Live Cell Therapeutics

The gut is a common site of infection by pathogenic bacteria and colonization can lead to serious disorders (Lamps, 2009). Infection and illnesses occur when a pathogen or toxins produced by a pathogen are ingested and the pathogen establishes itself within the host (Bintsis, 2017; Mousavi Khaneghah et al., 2020). Some notable bacterial pathogens include *Vibrio cholera, Salmonella*, *Clostridium* species, *Staphylococcus aureus*, *Shigella and Listeria*. Upon ingestion these pathogens can adhere to the mucosal layer and colonize the intestine (Martens et al., 2018). In some cases aggressive infectious microbes secrete toxins and enzymes that allow them to invade further into the epithelial layer where they can induce extensive tissue damage, inflammation and potentially gain access to other organs and the bloodstream (Fasano, 2002; Lamps, 2009).

Food-borne pathogens typically enter into a gut environment that has a fully established resident microbial community (Khan et al., 2021). Competition with the endogenous community, coupled with the host's immune system, can often control infections. Indeed, some opportunistic pathogens are commonly resident in the gut of most mammals but do not trigger disease unless the endogenous microbial community is disrupted by stress or more often in response to antibiotic drugs

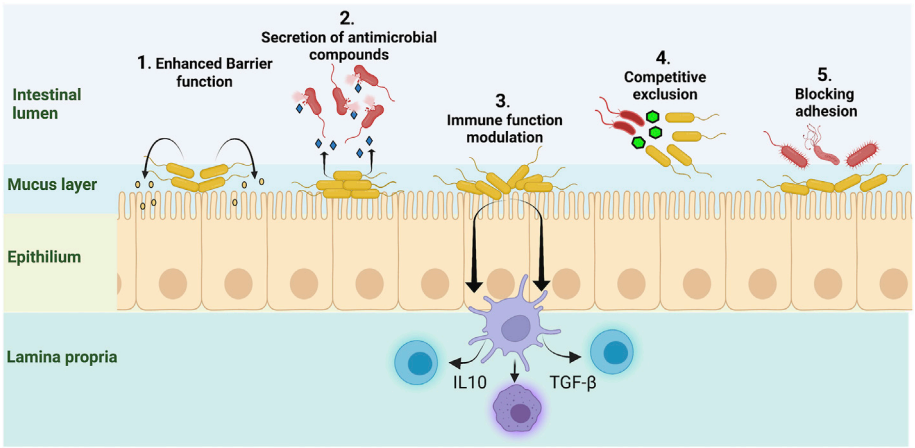


FIGURE 2 | Native probiotic microbes provide beneficial activities. 1) Secretion of short chain fatty acids (SCFA) and other small molecules can enhance epithelial barrier function. 2) Probiotic strains can secrete anti-microbial peptides and other small molecules that inhibit pathogenic bacteria. 3) Probiotic strains can modulate inflammatory responses through influence on dendritic cells and T-cells to improve immune function and dampen secretion of inflammatory cytokines. 4) Probiotics compete with pathogens for space and nutrients thus limiting their proliferation in a crowded gut environment. 5) Resident microbiota and probiotics inhibit pathogen adhesion and colonization of the gut walls. Figure constructed using Biorender.

that decimate the resident microbes creating an opportunity for pathogenic strains to rapidly expand (Raplee et al., 2021; Abd El-Hack et al., 2022).

Despite development in the areas of medicine, nutrition and food science, food-borne intestinal infections remain a global challenge. The World Health Organization (WHO) estimated that 22 foodborne pathogens caused about two billion illnesses, resulting in over one million deaths in 2015 alone (Kirk et al., 2015). The impacts of these illnesses are diverse and pose a significant threat to public health and socio-economic development globally. Since foodborne pathogens cause deaths and illnesses of millions of people worldwide, developing strategies to control and kill these microorganisms is a high priority. Antibiotic drugs have been our most powerful tool to prevent and treat pathogenic bacterial infection. However, antimicrobial resistant pathogenic bacteria have been reported with increased frequently and pose a challenge in the food industry. Therefore, there is an urgent need for alternative strategies to control foodborne pathogens and illnesses.

The gut environment is populated with commensal microbial species adapted with capabilities to thrive in a competitive environment. One strategy to combat invading pathogenic organisms is to harness and exploit beneficial microbial species. Probiotics are microbial organisms that safely reside in the gut and confer a health benefit to a host with potential for treatment of infectious diseases and inflammation of the gastrointestinal (GI) tract. Some of the most widely employed probiotic strains include *Lactobacillus*, *Bifidobacterium* and *Saccharomyces* species. These can survive the acidic environment of the stomach and colonize the gut to confer health benefits making them a staple throughout the food industry and offer the potential to act as live cell therapeutic agents (Mazzantini et al., 2021).

Probiotic microbes exhibit a diverse array of mechanisms that have potential to modulate the gut environment and immune system. Some of the benefits supplied by probiotic organisms are: 1) improvement of the intestinal epithelial layer barrier through secretion of metabolites and small molecules, 2) secretion of antimicrobial and anti-inflammatory factors, 3) ability to enhance innate immunity and modulate pathogen-induced inflammation. 4) competition with opportunistic pathogens for space and nutrients, and 5) inhibition of pathogen adhesion and translocation (Zhang et al., 2017; Wan et al., 2019; Flaugnatti et al., 2021) (**Figure 2**). However, these native abilities can be overwhelmed in response to dietary changes, stress, broad spectrum antibiotic treatments or other causes of dysbiosis (Gagliardi et al., 2018). Advances in genomics, DNA sequencing, DNA synthesis, DNA assembly techniques and probiotic research have provided the tools to engineer live cells therapeutics capable of addressing current challenges posed by disease causing bacteria.

2.2 Engineered Probiotics for Use in the Control of Foodborne Pathogens

The power of synthetic biology lies in its ability to program living cells with functionalities that may not currently exist, or not exist in a context that can be applied to solve the desired problem. Synthetic biology in conjunction with adaptive laboratory evolution (ALE) strategies can be used to accelerate evolution and sample an increased chemical space that may be necessary to generate effective therapeutics to control antibiotic resistant microbial pathogens (Bober et al., 2018; Pedrolli et al., 2019). Among the possible opportunities for engineering live cell therapeutics include: 1) Building biosensors for rapid specific detection of pathogens. 2) Engineering probiotic strains with improved ability to compete with or specifically kill selected pathogens. 3) Rewiring metabolism in probiotic strains to modulate the host immune response to improve recognition of

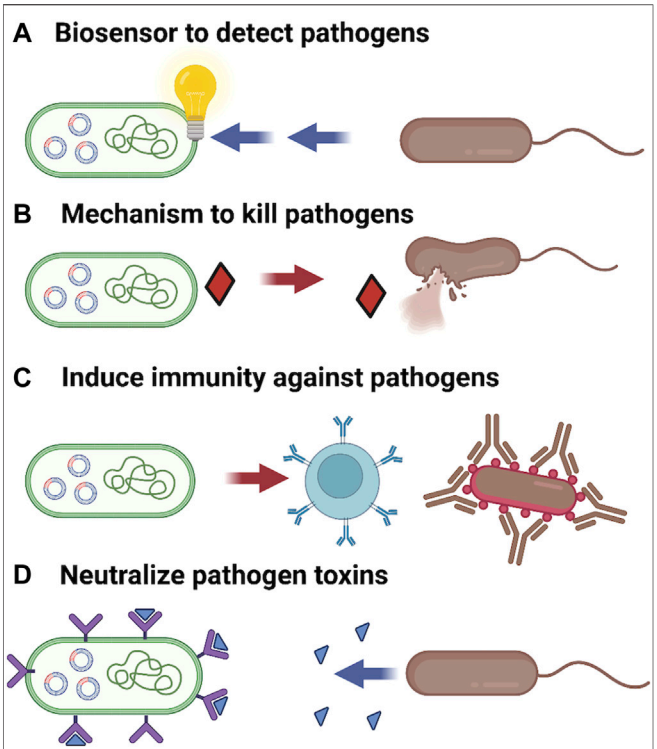


FIGURE 3 | Probiotics can be engineered with improved functionalities to inhibit pathogens. **(A)** Receptors wired to a genetic circuit capable of providing an easily measurable response can act as biosensors to detect and even quantitate pathogens presence. **(B)** Secretion of bacteriocins and lytic enzymes can be programed into favorable strains to degrade biofilms and allowing specific killing of pathogenic microbes. **(C)** Engineering expression of pathogens surface markers on harmless probiotic strains can provide for vaccination to increase immunity against pathogens. **(D)** Expression of receptors for bacterial toxins can allow a probiotic to act as a biological vacuum to neutralize toxins and reduce tissue damage from infection. Additionally, probiotic strains can be programmed to secrete quorum sensor antagonists to reduce expression of virulence factors. Figure constructed using Biorender.

pathogens or reduce harmful inflammatory reactions. 4) Engineering probiotic strains to neutralize toxins secreted by pathogens or disrupt communication and virulence (**Figure 3**). It is notable that these are not mutually exclusive as strains can be engineered to include more than one of these functionalities. Here we will consider some recent advances that have been made in engineering probiotics for the treatment and prevention of foodborne illnesses induced by *Vibrio cholera*, *Staphylococcus aureus*, *Clostridium perfringens* and *Clostridioides difficile*.

2.2.1 *Vibrio cholerae*

Cholera, a waterborne, life-threatening, gastrointestinal infection, is caused by the Gram-negative bacterial pathogen *Vibrio cholerae* and affects millions of individuals annually with an estimate of 21,000 to 143,000 deaths recorded worldwide (Zuckerman et al., 2007; Ali et al., 2015; Cho J. Y. et al., 2021). It is characterized by watery diarrhea and vomiting which leads to rapid dehydration, hypovolemic shock, acidosis and death (Ali et al., 2015; Hsiao and Zhu, 2020). *V. cholerae* colonizes the epithelium of the small intestine where it can form a resistant biofilm, produces cholera toxin and toxin-coregulated pili and employs a type 6 secretion system to kill competing native gut bacteria (Faruque et al., 1998; Zhu and Mekalanos, 2003; Logan et al., 2018).

Cholera toxins are responsible for altering the hosts cellular signaling pathways causing cellular damage and the watery diarrhea characteristic of cholera; while the toxin-coregulated pili aids in colonization of the gut epithelium (Herrington et al., 1988; Cho J. Y. et al., 2021). Current treatments for cholera are centered around antibiotic therapy and while these have dramatically reduced case fatality, resistance to antibiotics has been demonstrated in cholera endemic and epidemic countries creating a need for new antimicrobial strategies (Clemens et al., 2017). Some natural probiotic microbes can aid in *V. cholerae* infections. For example, oral administration of *Lactobacillus acidophilus* and *Bfidobacterium* spp. in mouse models that were initially infected with cholera has shown to have enhanced mucosal and systemic immune response to the cholera toxin (Tejada-Simon et al., 1999).

Aside from antibiotic treatment, vaccination is considered key to controlling the spread of cholera (World Health Organization, 2018). Beyond the potential of their native therapeutic properties, probiotic microbial cells offer a potential vehicle for vaccine delivery to the gastric mucosa (Gardlik et al., 2012; Iqbal et al., 2021). The budding yeast *Saccharomyces boulardii* has properties including acid tolerance and good growth at 37°C that allow it to be employed as a probiotic (Sazawal et al., 2006; Czerucka et al., 2007; Kelesidis and Pothoulakis, 2012). This yeast has been engineered to express the *V. cholerae* toxin coregulated pilin gene tcpA, with the objective of inducing a pre-emptive immune response to inhibit gut colonization by *V. cholerae* (Awad et al., 2020). Although no data are available on whether ingestion of the probiotic can induce a strong mucosal immune response to *V. cholerae,* it is relevant that *S. boulardii* can be engineered to either secrete foreign antigens or display them on the cell surface (Wang et al., 2014; Li et al., 2021). Oral vaccination using recombinant probiotics as the delivery system could thus be an effective means to activate both the innate and adaptive immune responses of the host to reduce the risk of transmission of foodborne infections (Mathipa and Thantsha, 2017).

Many bacterial pathogens including *V. cholerae* employ a quorum sensing system (QS) to induce colonization, biofilm formation, virulence and broad dissemination in the host's gastrointestinal tract (Hsiao and Zhu, 2020). QS is a signaling mechanism that bacteria use to respond to chemical hormone-like molecules called autoinducers, secreted by bacteria of the same species (Reading and Sperandio, 2006). The soluble QS molecule is of necessity unique to the pathogen and thus offers a target to allow detection of the pathogen. The *V. cholerae* quorum sensing system has been transplanted into *E. coli* and linked to a pathway to induce expression of a guide RNA that targeted dCAS9 to inhibit expression of a GFP reporter gene. In the absence of the *V. cholerae* quorum sensing molecule CA1, the GFP reporter was silenced. Upon binding of CA1 the expression of the gRNA was repressed leading to activation of the GFP reporter signal. This biosensor was extremely sensitive and

allowed detection of *V. cholerae* (Holowko et al., 2016). The same group further engineered the *E. coli* biosensor and constructed a sense-and-kill suicide microbe (Jayaraman et al., 2017). Upon detection of the *V. cholerae* quorum sensing molecule a synthetic circuit was activated that induced expression of a synthetic endolysin that was exported to the periplasm leading to lysis of the biosensor and releasing the endolysin, leading to killing of the *V. cholerae* cells. A similar synthetic biosensor circuit was constructed in *Lactococcus lactis* which permitted effective detection of *V. cholerae* in infected mice (Mao N. et al., 2018). *L. lactis* has a natural capability to inhibit *V. cholerae* through secretion of lactic acid to acidify the local environment and consumption of a native strain is sufficient to reduce cholera-induced morbidity in mice (Mao N. et al., 2018). In this case the biosensor circuit was not linked to any mechanisms to kill or inhibit *V. cholerae* but this is a direction that would likely yield a powerful therapeutic tool (Mao N. et al., 2018). The power of such a synthetic circuit lies in the specificity of response, allowing the sense-and-kill strain to reside in the gut and only respond upon detection of the pathogen.

Signalling molecules of the quorum sensing pathway are critical to regulating the expression of toxins and other virulence genes by *V. cholerae*. These signaling molecules provide not only a biomarker to allow detection of the pathogen but provide a potential means to disrupt induction of virulence factors. An engineered *E. coli* strain expressing and secreting the *V. cholerae* autoinducer CA1 was able to disrupt toxin expression in mice inoculated with *V. cholerae,* leading to a significant increase in survival (Duan and March, 2010). An additional strategy to reduce the harm done by cholerae toxin has been to engineer a probiotic *E. coli* to express a surface receptor for cholera toxin. This probiotic could significantly reduce the amount of cholera toxin binding to the gut wall and improve recovery in a mouse model of *V. cholerae* infection (Focareta et al., 2006).

The concept of recombinant probiotics improving the inherent functions of natural probiotic strains, adding a surveillance attribute, being used as a drug or vaccine delivery system, or being "weaponized" to neutralize pathogen virulence factors, is a promising direction for the development of new therapies intervening with the pathogenic action of *V. cholerae*.

2.2.2 *Staphylococcus aureus*

Staphylococcus aureus are a species of potentially pathogenic bacteria that are widespread in the environment. They commonly infect the blood, respiratory and most notably, the gastrointestinal tract of affected hosts. *Staphylococcus* infection is also one of the primary causes of community and hospital acquired infection, often resulting in severe and sometimes life-threatening conditions (Díaz et al., 2021). It is estimated that staphylococcal food poisoning causes approximately 241,188 illnesses and 1,064 hospitalizations in the United States alone (Scallan et al., 2011). *S. aureus* poses a serious public health hazard owing to the ability to secrete heat resistant enterotoxins that trigger inflammatory responses and the ability to form highly resistant biofilms that resist the host defenses, decrease the efficiency of antimicrobial factors, and provide resistance to antibiotics (Otto, 2013; Bintsis, 2017; Paik et al., 2019; Díaz et al., 2021). This pathogen infects a host and employs a quorum sensing signaling pathway to induce expression a collection of virulence factors including toxins, enzymes and secreted biofilm components that allow it to colonize and become established in the host.

The rapid appearance of antibiotic resistance in this microbe presents a challenge for developing antimicrobial therapies to treat and prevent infections. *S. aureus* acquires mutations at a rate that is about 1,000 times faster than for *E. coli* (Pipiya et al., 2020). This has allowed for the emergence of the notorious methicillin-resistant (MRSA) strains that are resistant to the entire β-lactam antibiotic class and created the need for improved therapies for detection and clearance of *S. aureus*. This is where bioengineered probiotics may have a pivotal role.

Several probiotic organisms including *lactobacillus* species, *Bacillus subtilis* and *S. boulardii* have demonstrated inhibitory activities toward *S. aurerus* though competitive exclusion, inhibiting growth by metabolite secretion, inhibition of biofilm formation and disruption of quorum sensing (Saidi et al., 2019; Nataraj and Mallappa, 2021). These antimicrobial activities coupled to the more general probiotic properties of reducing inflammation and improving epithelial barrier function make these organisms excellent candidates for engineering to improve upon their existing capabilities and to add more specific antimicrobial functionalities (Mathipa and Thantsha, 2017).

One way to improve the function of favorable probiotic strains is to make them more competitive within the gut environment. While these strains are all naturally tolerant of the acid conditions of the stomach and the routine 37°C temperature, one place that an advantage can be given is through bestowing resistance to the natural population of bacteriophage that populate the gut. Resistance to phage populations might permit a desirable probiotic strain to more effectively compete with pathogens and opportunistic pathogens that invade the human gut. Nagarajan *et al.* employed mutagenesis and a selective resistance procedure to isolate a probiotic bacterium, *Lactobacillus plantarum*, that was resistant to a phage population. *In vitro* competition tests between *S. aureus* and the selected *L. plantarum* strain demonstrated that in the presence of a phage population the probiotic was able to out compete *S. aureus* thereby reducing the ability of the pathogen to adhere to and colonize the epithelial cell layer (Nagarajan et al., 2019). This proof of principle experiment clearly has limited utility in that the probiotic was selected for resistance to the phage population from a specific environment and would likely be susceptible to phage from other environments. Additionally, given the high rate of mutagenesis in bacteriophage it is likely that phage capable of lysing the probiotic would readily emerge. None-the-less improving the competitiveness of favorable probiotic strains is a notion worthy of further investigation. Bacterial species with fully engineered and recoded genomes have been produced to eliminate redundant codons (Fredens et al., 2019). These are fully resistant to phage infection as they lack tRNA required for phage gene expression (Robertson et al., 2021). Application of this technology to a probiotic strain could

give that strain a competitive advantage, allowing it to easily outgrow opportunistic pathogens in the gut.

Relatively simple genetic circuits have been engineered into probiotic strains to detect and in some cases respond to *S. aureus* infections. The probiotic bacteria strain *Lactobacillus reuteri* is effective in reducing *S. aureus* viability in a host through the mechanism of competitive exclusion, biosurfactant secretion and production of lactic acid (Sikorska and Smoragiewicz, 2013; Wang et al., 2015). An *L. reuteri* strain that expressed the *S. aureus* AgrA and AgrC genes to provide a receptor and response pathway for detection of the AIP1 signaling molecule was constructed such that activation of AgrA. would repress expression of a β-glucuronidase gene to allow colorimetric detection (Lubkowicz et al., 2018). This mimic system can detect the presence of the autoinducer at very low concentrations and could be useful for high-throughput quantitative and qualitative detection of *S. aureus*, especially in hospital settings (Lubkowicz et al., 2018). This method is also cost-effective in comparison to current mass spectrophotometry and DNA sequencing approaches. The probiotic system can also be fine-tuned as a chassis for the eradication of *S. aureus*.

Several probiotic organisms described above are capable of competing with *S. aureus* in the gut environment. In an effort to further improve the utility of probiotics to impeded *S. aureus* infections several groups have sought to "weaponize" probiotic strains through engineering them with genes encoding antimicrobial factors with specificity for the target pathogen. Bacteriocins are proteinaceous compounds produced by bacteria to facilitate anti-microbial activity against other bacteria (Bastos et al., 2010). Lysostaphin is a metallo-enzyme that is relatively specific for the cell wall structures of *Staphylococcal* species (Kokai-Kun et al., 2007; Bastos et al., 2010). *P. pastoris* strains that have been engineered to produce lysostaphin displayed efficient eradication of *S. aureus* during *in vitro* co-culture experiments (Pipiya et al., 2020). Despite the effectiveness of lysostaphin to kill *S. aureus in vitro* infections in the human gut pose a greater challenge owing to biofilm formation by the colonizing *S. aureus*.

A probiotic strategy to deal with *S. aureus* infections that has emerged is the use of engineered *Mycoplasma pneumoniae*. This organism has a small genome, is genetically manipulable, has an unusual codon usage and limited recombination that limits horizontal gene transfer (Osawa et al., 1990; Himmelreich et al., 1996; Krishnakumar et al., 2010; Sluijter et al., 2010). Additionally, it lacks a cell wall and so has a limited ability to trigger an inflammatory response in mammalian hosts (Sukhithasri et al., 2013). An *M. pneumoniae* with reduced virulence owing to deletion of the *mpn133* and *mpn372* genes was engineered to secrete dispersin B, a glycosyl hydrolase that can efficiently dissolve biofilms formed by *S. aureus*, and the bacteriocin lysostaphin (Garrido et al., 2021). The engineered strain was effective in disrupting and reducing *S. aureus* biofilm development in a catheterized mouse model as well as eliminating biofilms formed *ex-vivo* (Garrido et al., 2021).

Staphylococcus aureus remains a challenging pathogen for medicine to control owing to its ability to rapidly acquire drug resistance, the tenacious nature of its protective biofilms, ability to evade the immune system and ability to invade a variety of tissues. Clever genetic engineering of favorable host organisms to act as biosensors and "hunter-killers" that can disrupt biofilms may become a key tool that can be combined with other therapies including the use of engineered bacteriophage to combat this common microbial pathogen.

2.2.3 *Clostridium perfringens*

Clostridia species are anaerobic organisms prevalent in soil. Several of these are well known human pathogens including *C. tetani, C. difficile and C. perfringens* and can induce disease in humans and agricultural animals (Popoff and Bouvet, 2009; Onderdonk and Garrett, 2020). The ability of these organisms to form spores that are resistant to heat and many common cleaning agents makes them challenging to control (Augustyn et al., 2022). *C. perfringens* is common within ecosystems and frequents the gastrointestinal tracts of livestock and poultry (Abudabos et al., 2019; Bai et al., 2020). It is responsible for a range of severe conditions, including food poisoning, gas gangrene, respiratory infections and necrotizing enteritis (NE) in both humans and animals (Lindström et al., 2011; Gervasi et al., 2014a). *C. perfringens* ranks second among the causes of foodborne illness and is responsible for greater than one million illnesses each year in the United States alone (Bintsis, 2017).

C. perfringens infection has been associated with the development of necrotizing enterocolitis (NEC), a gastrointestinal disorder affecting pre-term infants with a mortality rate of about 20% (Duchon et al., 2021). This pathogenic species is also the leading cause of necrotic enteritis (NE) in farm animals, costing the global poultry industry over two billion dollars annually, mainly due to the high prices of antibiotic treatments and therapeutic feed supplementation (Van Immerseel et al., 2009; Paiva et al., 2014; Smith et al., 2014). *C. perfringens*, like *S. aureus*, employs a quorum sensing system and secretes a battery of adhesins, proteolytic enzymes and toxins allowing it to colonize and invade the host (Pruteanu and Shanahan, 2013; Kiu and Hall, 2018; Navarro et al., 2018; Bai et al., 2020). They are also able to secrete bacteriocins that inhibit some competing bacteria to increase availability of nutrients and clear space to colonize (Mora et al., 2020).

Antibiotic drugs are the most common means to treat and prevent *C. perfringens* infection. However, the growing number of antibiotic resistant pathogenic strains found in isolated livestock and agricultural sources suggest their use may not be beneficial in the long term (Slavić et al., 2011). Indeed, up to 33% of isolates from some meat samples harbor *C. perfringens* with resistance to the common antibiotics: ampicillin, tetracyclin, amoxicillin, ciprofloxiacin and chloramphenicol. Further multidrug resistance was found in 38% of those isolates and multidrug resistant *C. perfringens* has been isolated from livestock (Zhang et al., 2021; Hassani et al., 2022). Vaccination processes that have been successful against *C. tetani* have also been developed for *C. perfringens* but have yet to become widely and successfully established (Alizadeh et al., 2021). The growing health and economic threat posed by resistant forms of *C. perfringens* is creating an urgent need to develop new approaches to controlling

infection by this pathogen that are effective, affordable and can be widely disseminated.

The requirements for an effective chassis for a live cell therapeutic would include the ability to thrive in the gut environment, being genetically manipulable, controllable, exhibit limited horizontal gene transfer and have intrinsic properties that make it competitive to the pathogen. A variety of probiotic organisms including *Lactobacillus*, *Bifidobacterium*, *S. cerevisiae* and *S. boulardii*, have been demonstrated to compete with or inhibit the growth of *C. perfringens in vitro* and *in vivo* (Schoster et al., 2013; Guo et al., 2017; Chuang et al., 2019; Gong et al., 2020; Khalique et al., 2020; Wang et al., 2020). A limited number of probiotic species have been demonstrated to secrete bacteriocins that inhibit *C. perfringens* (O'Shea et al., 2009; Heo et al., 2018; Golneshin et al., 2020; Lee et al., 2021). These organisms and their bacteriocins may be exploited in the future as tools to inhibit opportunistic *C. perfringens* infections.

Extensive effort has been expended on engineering probiotic organisms to express *C. perfringens* virulence factors with the aim of generating effective live cell vaccines. Delivering toxins to humans or animals as a vaccine has inherent risks even if those toxins have been inactivated with heat or formaldehyde. In contrast, toxoids mimic toxins released by the pathogens but are not associated with virulence. A live cell expressing target antigens has the potential benefit of inducing immunity at the level of the gut mucosa and provides a mechanism that allows extended exposure to the toxoid antigen rather than single doses. Probiotic *Bacillus subtilis* has been engineered to produce, in cells, or display on spore walls, a fragment of *C. perfringens* toxin-α fused to gluthatione S-transferase. Mice immunized with toxoid displaying spores developed immunity and were protected against lethal doses of the toxin (Hoang et al., 2008). This approach provided immunity to only a single toxin but it could clearly be extended to cover other toxins. Additionally, the use of spores as a display and delivery vehicle provides a stable vaccine that could easily be produced in large volumes.

The probiotic bacteria *Lactobacilli casei* has been engineered to express a *C. perfringens* α-toxoid mimicking the α-toxin. Mice vaccinated with the probiotic developed anti-α toxin antibodies and survived challenge with a lethal dose α-toxin while the control animals rapidly succumbed to the effects of the toxin (Alimolaei et al., 2018; Song et al., 2018; Gao et al., 2019). Oral administration of an **ε**-toxoid expressing *L. casei* strain or a β-toxoid expressing *L. casei* strain was also able to induce an effective antibody response to the toxin (Alimolaei et al., 2016; Alimolaei et al., 2017). Similar results were found by Bai et al. and Zhao et al. who also engineered *L. casei* to express toxoids mimicking the *C. perfringens* α, ε, β1, and β2 toxins and vaccinated rabbits with the recombinant probiotic. Challenge experiments demonstrated that this strategy yielded an 80% protection rate, even more effective than vaccinating with inactivated, whole *C. perfringens* cells (Zhao et al., 2017; Bai et al., 2020). An attenuated *Salmonella typhimurium* expressing *C. perfringens* antigens NetB, α-toxoid and Fba administered to chickens induced both serum and mucosal antibody responses and provided partial protection from necrotizing enteritis intestinal lesions (Wilde et al., 2019). A similar investigation using a self-lysing *S. typhimurium* vector expressing α-toxoid and NetB administered to broiler chickens yielded induction of IgA, IgY, and IgM antibodies against the toxins and provided significant protection from a subsequent *C. perfringens* infection (Jiang et al., 2015). The yeast *S. boulardii* was engineered to secrete a fusion protein including the *C. perfringens* enterotoxin CPE. Oral administration of this live cell vaccine to mice induced both IgG and IgA antibody responses suggesting that *S. boulardii* may be an effective vector for application as a live cell vaccine (Bagherpour et al., 2018). Additionally, a "disarmed" *C. perfringens* strain lacking the NetB toxin, when orally administered to vaccinate broiler chickens, produced a significant protective effect and reduced the incidence of necrotizing enteritis lesions upon challenge with virulent *C. perfringens* (Mishra and Smyth, 2017). Recently, a native microbial strain isolated from poultry, *Ligilactobacillus agilis* La3, has been engineered to effectively produce *C. perfringens* NetB (Vezina et al., 2021). Although its effectiveness has yet to be characterized, it is likely that a native member of the poultry gut microbiota will be an effective delivery vehicle.

Metabolic engineering of probiotics has also been directed toward generating live cell therapeutics capable of not just competing with *C. perfringens* but able to attack and lyse *C. perfringens* to promote their clearance. This work has borrowed from the highly effective endolysin enzymes produced by bacteriophage to penetrate the cell walls of infected host cells to allow phage release (Schmelcher et al., 2012). Bacteriophage derived endolysins specific for binding to and degrading the cell walls of *C. perfringens* and have been identified and demonstrated to be effective for killing those organisms *in vitro* (Gervasi et al., 2014a). Additionally, protein engineering has been applied to improve activity and increase specificity of *C. perfringens* killing (Swift et al., 2015). *Clostridia* and other pathogenic bacteria can acquire resistance to antibiotic drugs through modification of or degradation of the drug or the use of efflux pumps. Endolysins specifically bind to and cleave highly conserved peptidoglycan structures that are essential to the integrity of the cell wall (Fischetti, 2010). Acquiring resistance to an endolysin would require the cells to develop changes in critical conserved cell wall structures that would likely reduce the integrity of that structure and so is expected to be a very rare event (Fischetti, 2006).

Synthetic biology techniques have been applied to engineer the probiotic bacteria, *Lactobacillus johnsonii,* secrete an endolysin against effective against *C. perfringens.* They found that this led to an improved ability of the probiotic bacteria to lyse *C. perfringens* cells and increased clearance of the pathogen in plate assays (Gervasi et al., 2014a; Gervasi et al., 2014b; Cho J.-H. et al., 2021). *S. cerevisiae* have been engineered to effectively display active *C. perfringens*-specific endolysins on their cell walls (Ritter and Hackel, 2019). These investigators used the yeast display system to engineer the endolysin for improved stability but they did not test them for *C. perfringens* killing. Such a system could be potentially employed with a probiotic yeast for application *in vivo*. While engineered *C. perfringens* killing probiotics have yet to be proven *in vivo* they display the potential to reduce pathogen load and aid in control of serious

infections. Overall, these models demonstrate that probiotics can be manipulated to produce anti-microbial factors and directly promote the clearance of the pathogenic bacteria *C. perfringens*.

2.2.4 *Clostridioides difficile*

Clostridioides difficile, like *C. perfringens*, is an anaerobic Gram positive bacteria that is widely disseminated in the environment. It is unclear how prevalent *C. difficile* is within the gut microbiome of human populations but it can be isolated from the gut of new born and healthy children (Tullus et al., 1989; Cui et al., 2021). *C. difficile* infection leads to diarrhea and colitis that can be fatal and is responsible for greater than 29,000 deaths each year in the United States (Lessa et al., 2015a; Lessa et al., 2015b). Indeed, *C. difficile* infection is among the top ten forms of health care associated infection and is a continuing problem in hospital environments (Suetens et al., 2018). *C. difficile*, like *C. perfringens*, secretes a set of toxins *tcdA* (toxin A), *tcdB* (toxin B), and *CDT* that are responsible for the tissue damage and pathology associated with infections (Kuehne et al., 2014; Chandrasekaran and Lacy, 2017). The pathology of *C. difficile* infection in humans and animals has been extensively reviewed (Abt et al., 2016; Moono et al., 2016).

Pathogenic infections with *C. difficile* are associated with conditions that yield dysbiosis including antibiotic treatments or severe stress that provide the opportunity for *C. difficile* to colonize the gut (Hooks and O'Malley, 2017; Vasilescu et al., 2021). The use of a subset of antibiotic drugs is associated with an increased risk of inducing *C. difficile* infections likely owing the reduction of competing microbial species in the gut caused by the treatment (Vardakas et al., 2016). Infections with *C. difficile* are notoriously persistent. Most antibiotic treatments are not successful and even following extensive antibiotic regimens infections are recurrent in 20%–40% of patients (Pepin et al., 2005; Vardakas et al., 2016). The high rate of recurrence may be due to the ability of the cells to sporulate and embed spores in the gut wall (Castro-Cordova et al., 2021). Owing to the persistent nature of *C. difficile* infections the antibiotics currently used in treatment are vancomycin, fidaxomicin and metronidazole (McDonald et al., 2018). An ominous trend is that these antibiotic treatments now display reduced effectiveness in treatment of *C. difficile*. The effectiveness of metronidazole has reduced from 95% to 75% since 2010. Even vancomycin has displayed a reduced effectiveness from 95% to 85% (McDonald et al., 2018). The increased incidence of recurrent *C. difficile* infections reflects a failure of the currently employed antibiotic treatments and reveals the need for more effective drugs or alternative therapeutic approaches (Marra et al., 2020; Sholeh et al., 2020).

There has long been evidence that the normal diversity and composition of the gut microbiome is able to suppress outgrowth of *C. difficile* and limit pathological colonization of the gut (Freter, 1955). Administration of probiotic microbes including *Bifidobacterium breve* and *Lacticaseibacillus casei*, *S. boulardii*, and *L. rhamnosus GG*, either alone or in combination with antibiotic regimens have seen some success in inhibiting or resolving *C. difficile* induced gut pathologies in animal models (Gorbach et al., 1987; McFarland et al., 1994; Esposito et al., 2021; Panpetch et al., 2021; Yang and Yang, 2021; Yang et al., 2022). Treatment with defined combinations of probiotic strains that compete with *C. difficile* for mucin as a nutrient source have shown some success (Pereira et al., 2020). More complex combinations of probiotic strains have reached stage III clinical trials for treatment of *C. difficile* induced pathology and the application of fecal transplantation is gaining acceptance as a treatment despite concerns over the need to characterize the donor microbiota (Khoruts et al., 2021; McGovern et al., 2021). The mechanism by which probiotic administration aids in resolving *C. difficile* infection is likely a combination of competition for resources coupled with the ability of a subset of microbiota to secrete secondary bile acids, bacteriocins, and other compounds inhibitory to *C. difficile* vegetative cells, germination of *C. difficile* spores and potentially disruption of *C. difficile* quorum sensing (Spinler et al., 2017; Thanissery et al., 2017; Lv et al., 2020; Gunaratnam et al., 2021).

The effective deployment of probiotic microbes to resolve and prevent *C. difficile* infection has motivated efforts to engineer probiotics for improved specific functions. The pathologies and morbidity associated with *C. difficile* infection are induced by the toxins *tcdA, tcdB*, and *CDT*. Monoclonal antibodies directed against the toxins have proven effective in limiting the pathology of *C. difficile* infection (Wilcox et al., 2017). But thus far vaccination with toxoids has had limited success at inducing immunity or treating *C. difficile* infection (Dieterle et al., 2019; de Bruyn et al., 2021) However, engineered microbial cells expressing *C. difficile* toxins can induce effective protective immune response when orally administered in animal models. An attenuated *V. cholerae* strain secreting a toxin A fragment fusion protein induced anti-toxin A antibody production and a protective immune response in mice that survived a subsequent challenge with *C. difficile* infection (Ryan et al., 1997). Engineered attenuated *Salmonella typhimurium* and *Bacillus subtilis* have both been used as vectors for secretion of toxin A fragments to demonstrate the induction of protective immune response through mucosal delivery of the vaccinating antigen (Ward et al., 1999; Permpoonpattana et al., 2011). Oral vaccination with a "disarmed" *C. difficile* engineered to express toxin fragments provided substantial protective immunity to *C. difficile* challenge in animal models (Donald et al., 2013; Wang et al., 2018). Similarly, administration of an engineered *Lactococcus lactis* expressing a fragment of toxin A, either secreted from the cells or displayed on the cell wall induced a strong immune response and provided substantial protection to *C. difficile* challenge in mice (Yang et al., 2013). Engineering *L. lactis* to express toxin A and toxin B fragments in combination provided no greater protection than toxin A alone (Guo et al., 2015). *Clostridia* employ surface proteins to assist with adherence to the intestinal cell wall and promote colonization. *L. casei* engineered to express and display the *C. difficile* SlpA on its cell wall were able to effectively engraft in mouse intestines following inoculation and triggered a robust immune response that protected hamsters from death following infection with *C. difficile* (Vedantam et al., 2018). Additionally, *C. difficile* surface antigen *CD0873* involved

in attachment to the epithelial cell wall has emerged as an effective antigen for induction of a protective immune response and may be a good candidate for surface display on probiotics to induce immunity (Karyal et al., 2021). The apparent effectiveness of the live cell vaccines in animal models is encouraging and the relative ease with which a variety of toxin and surface antigens can be tested for the ability to induce protective immunity means that this approach may achieve further uptake.

One of the limitations of vaccination efforts against *C. difficile* is the relatively slow sero-conversion time as well as the inherent limitation of systemic antibodies directed against toxins in the intestinal lumen (Greenberg et al., 2012; de Bruyn et al., 2016). Microbial cells have been engineered to deliver therapeutic antibodies to neutralize *C. difficile* toxins during infection. *L. paracasei* strains were developed to express and secrete or display on their cell walls antibody chains capable of binding toxin B. Oral delivery of the recombinant probiotics provided protection against infection with a toxin B expressing *C. difficile* strain in a hamster model (Andersen et al., 2016). The yeast *S. boulardii* has been employed to express and secrete a multidomain neutralizing antibody directed against the major toxins of *C. difficile* (Chen et al., 2020). The engineered live cell therapeutic was effective in limiting disease caused by acute and recurrent infections by *C. difficile* in a mouse model. An advantage of a *S. boulardii* as a chassis for delivering antimicrobial activity is that it can be used in combination with antibiotic drugs that would render a bacterial probiotic ineffective. Additionally, this strategy has potential for further enhancement through improved expression and secretion of the therapeutic antibody. It could also be more broadly applied through expression of antibodies against other virulence factors.

A final consideration for deployment of live cell therapeutics to treat *C. difficile* infection is the application of lytic bacteriophage. Phage that specifically target *C. difficile* have been identified but treatment with phage has had limited success (Mayer et al., 2008; Meader et al., 2013; Nale et al., 2016; Whittle et al., 2022). The effectiveness of a *C. difficile*-specific phage was improved through engineering the phage to provide guide RNA to induce the endogenous *C. difficile* CRISPR system to target its own genome (Selle et al., 2020). Despite the improved effectiveness in killing *C. difficile* vegetative cells this phage will likely suffer the limitation of all phage therapies in that the target cells will rapidly evolve resistance.

Bacteriophage encoded endolysins with specificity for *C. difficile* cell walls have been identified (Mayer et al., 2008; Mondal et al., 2020). Oral application of purified endolysin has proven effective in controlling *C. difficile* infection in mice but owing to limitations of protein stability the therapeutic effect is modest (Peng et al., 2018). This limitation might be overcome through expression of the endolysin by a probiotic microbial strain allowing prolonged secretion into the gut and extending the time allowed for effective killing of colonizing *C. difficile*. While this is an intriguing idea it would be necessary to ensure that the endolysin did not have "off target" effects by killing other microbial species which might add to any dysbiosis rather than simply eradicate the pathogen.

The increasing frequency of drug resistant *C. difficile* infections and the severity of their pathology coupled with the high incidence of infections occurring in health care facilities creates urgency around the development of new strategies to control this pathogen. The research described above highlight the potential to use bioengineered probiotics as live cell vaccines and vehicles for delivering therapeutics to effectively control the damaging effects of opportunistic *C. difficile* outgrowth.

3 EMERGING APPLICATIONS AND CHALLENGES OF BIOENGINEERED PROBIOTICS

Bioengineering can be used to enhance the overall efficacy and effectiveness of known probiotic strains, particularly in their application to antagonize and treat diseases. The previous sections discussed the ways by which genetic engineering could enhance the inherent therapeutic action of probiotics. Common approaches include: increasing the specificity of the probiotic microbes for their target molecules, improving the targeted therapeutic delivery action in infected hosts, and improving the direct antimicrobial activity of the probiotic against pathogenic infections of the gastrointestinal tract. Genetic modifications can also be applied to probiotic strains to improve tolerance of ecological or host environmental stress including high temperatures, acidification, oxygen and food processing.

3.1 Potential Benefits of Engineered Live Cell Therapeutics to Combat Food Borne Pathogens

Engineered probiotics have significant potential for deployment as tools to detect pathogenic bacteria and eliminate or control them within the gut environment. For this purpose, live cell therapeutics display a number of potential benefits.

1. Engineered probiotic strains that have been trained to detect and kill pathogens can have very high specificity for the target pathogen. These can be applied with precision to eliminate the pathogen without disruption of the gut microbiota, thus imposing less stress on the host and avoiding dysbiosis that would later need to be remediated (**Figure 4A**).
2. There is a low probability of engineered probiotics leading to "off target" effects. Unlike systemically delivered drug therapies, engineered probiotics are localized to the site of infection in the gut. Since these are known to be safe organisms they are unlikely to have adverse side effects (**Figure 4B**).
3. They do not generate or encourage the development of antibiotic resistance. Most antibiotic drugs are cytostatic, offering the opportunity for microbes to escape the drug. Probiotics can be engineered as cytolytic agents to kill the target pathogen, reducing the opportunity for developing resistance. Additionally, these will not spawn resistance to antibiotic drugs since they have distinct modes of action and in

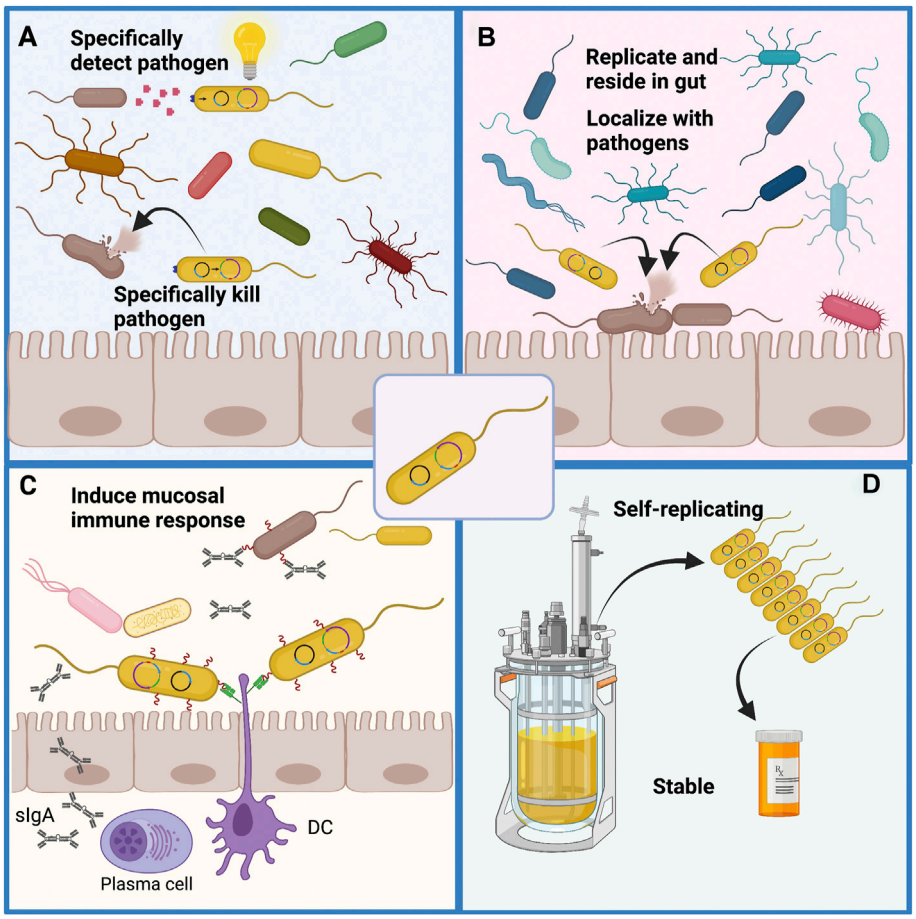


FIGURE 4 | Potential benefits of engineered live cell therapeutics. **(A)** Engineered probiotics have the potential to be trained to detect specific pathogens by way of signaling molecules, surface antigens, secreted metabolites or other small molecules unique to the pathogen. This high specificity may allow for selective killing of the pathogen without loss of the resident microbial population, thus avoiding the dysbiosis that results from administration of broad-spectrum antibiotic drugs. **(B)** An engineered probiotic would reside in the gut in close proximity to sites of infection to allow direct and specific killing of intestinal pathogens with limited systemic effects. As the probiotic would self-replicate it would be effective for a prolonged time without the need for multiple applications. **(C)** As a live-cell vaccine, probiotics can act as a delivery vehicle to secrete or display pathogen surface antigens, adhesion molecules or toxin fragments to induce mucosal immunity at the site of infection which may have benefits over systemic immunization. **(D)** Engineered live-cell therapeutics can be produced in large quantity relatively cheaply through bioreactor culture and the cells can be stored dry or frozen with high viability.

some cases such as engineered yeasts they could be deployed in conjunction with antibiotic drugs.

4. Live cell therapeutics with capability of killing pathogens will reside and replicate in the gut. Microbial cells can be engineered for increased stress resistance and improved ability to colonize the gut environment. Localized with the target pathogens, a live cell therapeutic is self-replicating and can remain in the gut to deliver its therapeutic activity for a longer period than any antibiotic drug would remain active, making it more effective with a lower number of administrations. In the absence of selection, the treatment is self-limiting as the therapeutic would be lost (**Figure 4B**).
5. As live cell vaccines, a probiotic expressing surface antigens or toxin fragments from a pathogen would have the advantage of generating mucosal immunity effective at the site of infection and colonization. Additionally, a live cell vaccine would allow for prolonged exposure to the target antigens rather than a single dose. An oral vaccine is simpler to provide and likely to have higher uptake than an injectable vaccine (**Figure 4C**).
6. Live cell therapeutics are self-replicating and so would be cheap to produce, transport and employ once developed. Although considerable time and money is necessary to generate highly effective live cell therapeutics, once generated, the production cost is low, it remains viable at room temperatures and can retain its stability in dried or frozen forms. This stability would permit their distribution in areas where prolonged freezing or refrigeration might not be practical (**Figure 4D**).

3.2 Challenges to Successful Application of Engineered Live Cell Therapeutics

Although bioengineered probiotics have been shown to be effective against pathogens that commonly infect the gut, their

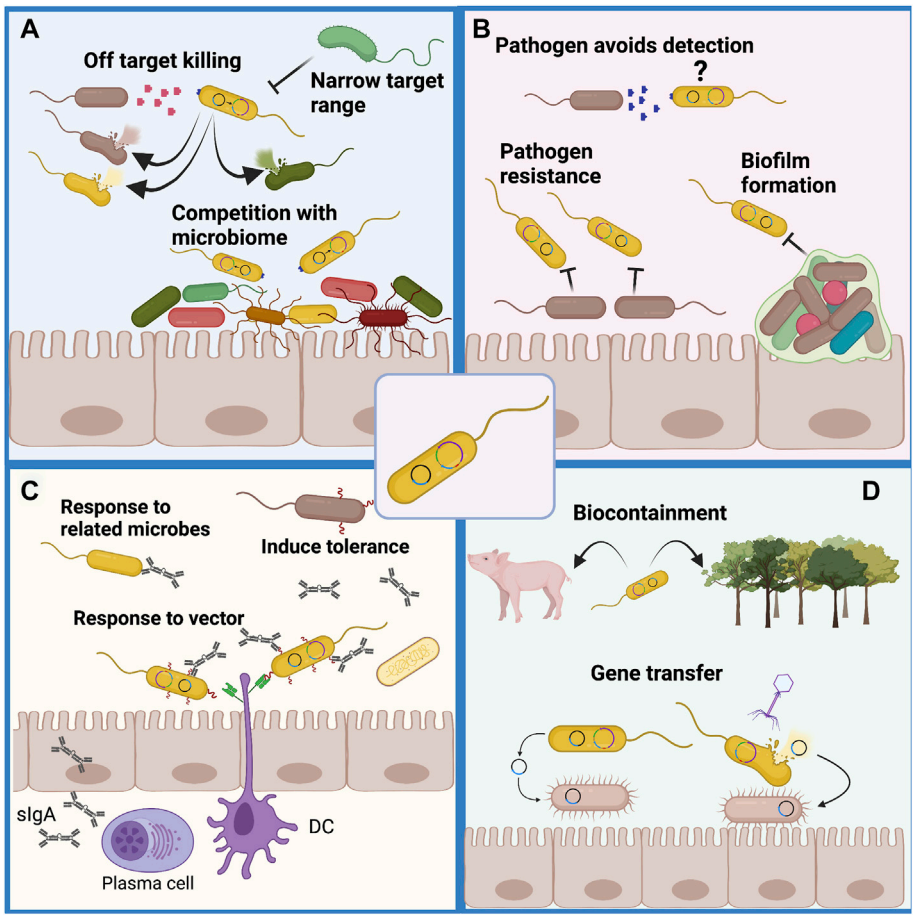


FIGURE 5 | Challenges remain to the wide-spread application of engineered probiotics. **(A)** Engineered probiotics must compete with the resident microbial species for nutrients to replicate and then act on pathogens. To compete effectively they must be as well adapted as the resident microbes but must not contribute to overwhelming the resident microbiota. The killing of pathogens must be precise to avoid off-target killing of resident microbiota which would increase dysbiosis. This imposes the limitation that the pathogen must be specifically identified before the application of the engineered probiotic. **(B)** It is likely that variants of any pathogen will arise that can avoid detection or develop resistance to the mechanism of killing used by the engineered probiotic. Biofilms that are formed by a variety of pathogens that colonize the gut pose a challenge for both detection and killing of the pathogen. **(C)** Administration of live-cell probiotic vaccines presents the concern of tolerant pathogens resistant to immune cell responses from vaccination. Live cell probiotics can also be recognized by immune cells as foreign microbes and targeted for cell death, thereby decreasing their viability and vaccine actions. **(D)** Containment of an engineered probiotic is important as it is difficult to predict how the organism will act in different environments and there could be unexpected effects on animals, insects or plants for which the probiotic was not engineered. Even within the planned host environment transfer of genetic material to other organisms is a concern as this can occur through conjugation and direct transfer to bacteria as well as through phage mediated transduction and natural DNA uptake occurring upon cell lysis for any reason.

activity can be limited by several factors and their usage can also come with several challenges.

1. Competitiveness and stability within the host gut will be a challenge. The physical robustness of a probiotic microbe is critical to its ability to function as an antagonist against pathogenic bacteria (Sola-Oladokun et al., 2017). The human gut is populated by resident microbes well adapted to life in that environment. Any engineered organism will need to be genetically stable, survive passage through the GI tract, multiply and be highly competitive to have any hope of colonizing or engrafting in the gut. Expression of heterologous genes is likely to impose a significant metabolic burden on any engineered microbe making it less competitive within the gut environment. An engineered microbe might benefit from a "stripped down" genome engineered to remove non-essential sequences. This is likely to prove challenging since even *E. coli* has many genes and loci with no clearly assigned function (Ghatak et al., 2019). However, increased fitness for colonizing the gut environment could be a two-edged sword. If the engineered microbe is too effective it might displace the natural microbiota leading to unpredictable effects of dysbiosis (**Figure 5A**). They must also tolerate commercial scale production and storage process without contamination or significant loss of viability. A variety of strategies around this have been reviewed (Wan et al., 2019).
2. Target specificity needs to be tuned such that only pathogenic microbes are targeted and killed. Failure to achieve specificity

for pathogens could lead to destruction of helpful microbial strains and increase dysbiosis (**Figure 5A**). Target specificity can be improved by means of sensors and switches to activate the killing mechanisms only when the target cells have been detected. Another approach might include expression of receptors on the surface of the engineered probiotic allowing it to adhere to or associate with the pathogen. However, while there is a benefit to the ability to target specific pathogens in the gut, this also presents the significant limitation that the target organism must be specifically identified (**Figure 5A**). In contrast, broad spectrum antibiotic drugs can be employed effectively even without specific characterization of the infectious organism.

3. Pathogen resistance will continue to pose a challenge. Even without specific selection, the diversity inherent in microbial populations means that variants that are resistant to the mechanism of killing programmed into a live cell therapeutic will arise (**Figure 5B**). It is also expected that pathogen variants with alterations in signaling molecules or molecules targeted by biosensors would eventually arise and avoid detection (**Figure 5B**). An additional means by which pathogens may develop resistance to any engineered probiotic is through biofilm formation (**Figure 5B**). While these can be overcome by production of degradative enzymes, they do pose a further challenge. Without selection these may not become prevalent but these will always limit the effectiveness of an engineered probiotic.
4. In the case of live cell vaccines, challenges will include achieving prolonged immunity following induction of mucosal immunity (**Figure 5C**). It is likely that this can be improved with suitable adjuvant strategies. Other concerns include the potential to induce tolerance to pathogens which might provide them increased opportunities. It is also possible that an immune response would be generated against the probiotic vehicle presenting the antigens. Such a response might reduce the effectiveness of the vaccine, induce a response to native gut microbiota leading to dysbiosis, and limit the effectiveness of readministering the live cell vaccine (**Figure 5C**). Any strains of the target pathogen with variation in the selected surface antigens may escape the immune response and induce the disease state.
5. Biocontainment: The importance of ensuring the safety of an engineered probiotic cannot be understated. While we may be able to demonstrate that a microbe has no negative effects within a lab environment or within test subjects, it is impossible to predict its behavior and interactions within other environments or in the guts of other organisms (**Figure 5D**). Auxotrophic requirements, self-limiting components and "kill-switches" can be engineered into live cell therapeutics to limit their dissemination but these may not be entirely sufficient for some purposes. Horizontal gene transfer can occur in most bacterial cell populations and this can be exacerbated through cell killing and by phage that can transduce genetic material (**Figure 5D**) (Sleator, 2010). In this regard, *S. cerevisiae* and *S. boulardii* have benefits as chassis organisms since they have no means for horizontal gene transfer, have no phage or virus that transmit cell to cell, and can be made sterile to avoid gene transfer to closely related wild yeasts.

4 CONCLUSION

The ominous rise in the frequency of infections caused by antibiotic resistant pathogens creates urgency around the development of new drugs and therapeutics to control infections. Synthetic biology and genomics offer powerful tools to discover, unmask and develop new natural product antibiotics and provide pipelines for their development. While this may prove to be a finite resource, it has yet to be fully exploited and may at least provide new compounds to keep us ahead of the wave of drug resistant pathogens. Engineered probiotics and live cell therapeutics are conceptually intriguing as they offer tools that can be applied with precision to detect, disarm and eliminate pathogens. While these solutions are provocative and exciting, a variety of challenges and concerns around safety and effectiveness must be overcome before this strategy can be considered a viable alternative to conventional therapies.

AUTHOR CONTRIBUTIONS

Author DS conceived the manuscript. All three authors, KC, LE, and DS contributed to writing the original manuscript and editing subsequent versions. Figures were prepared by DS with assistance from KC and LO.

FUNDING

This work was supported by grant NFRFE-2020-00786 from the New Frontiers in Research Fund and Discovery Grant RGPIN-2021-02898 from the Natural Sciences and Engineering Research Council of Canada to DS. KC is recipient of a fellowship from the Republic of the Philippines Department of Science and Technology, Science Education Institute.

REFERENCES

Abd El-Hack, M. E., El-Saadony, M. T., Elbestawy, A. R., El-Shall, N. A., Saad, A. M., Salem, H. M., et al. (2022). Necrotic Enteritis in Broiler Chickens: Disease Characteristics and Prevention Using Organic Antibiotic Alternatives - a Comprehensive Review. *Poult. Sci.* 101 (2), 101590. doi:10.1016/j.psj.2021.101590

Abt, M. C., McKenney, P. T., and Pamer, E. G. (2016). *Clostridium difficile* Colitis: Pathogenesis and Host Defence. *Nat. Rev. Microbiol.* 14 (10), 609–620. doi:10.1038/nrmicro.2016.108

Abudabos, A. M., Ali, M. H., Nassan, M. A., and Saleh, A. A. (2019). Ameliorative Effect of *Bacillus Subtilis* on Growth Performance and Intestinal Architecture in Broiler Infected with *Salmonella*. *Animals* 9 (4), 190. doi:10.3390/ani9040190

Acharya, Y., Bhattacharyya, S., Dhanda, G., and Haldar, J. (2022). Emerging Roles of Glycopeptide Antibiotics: Moving beyond Gram-Positive Bacteria. *ACS Infect. Dis.* 8 (1), 1–28. doi:10.1021/acsinfecdis.1c00367

Ali, M., Nelson, A. R., Lopez, A. L., and Sack, D. A. (2015). Updated Global Burden of Cholera in Endemic Countries. *PLoS Negl. Trop. Dis.* 9 (6), e0003832. doi:10.1371/journal.pntd.0003832

Alimolaei, M., Golchin, M., and Daneshvar, H. (2016). Oral Immunization of Mice against *Clostridium perfringens* Epsilon Toxin with a *Lactobacillus Casei* Vector Vaccine Expressing Epsilon Toxoid. *Infect. Genet. Evol.* 40, 282–287. doi:10.1016/j.meegid.2016.03.020

Alimolaei, M., Golchin, M., and Ezatkhah, M. (2017). Orally Administered Recombinant *Lactobacillus Casei* Vector Vaccine Expressing β-toxoid of *Clostridium perfringens* that Induced Protective Immunity Responses. *Res. Vet. Sci.* 115, 332–339. doi:10.1016/j.rvsc.2017.06.018

Alimolaei, M., Golchin, M., Abshenas, J., Ezatkhah, M., and Bafti, M. S. (2018). A Recombinant Probiotic, *Lactobacillus Casei*, Expressing the *Clostridium perfringens* Alpha-Toxoid, as an Orally Vaccine Candidate against Gas Gangrene and Necrotic Enteritis. *Probiotics Antimicro. Prot.* 10 (2), 251–257. doi:10.1007/s12602-017-9276-8

Alizadeh, M., Shojadoost, B., Boodhoo, N., Astill, J., Taha-Abdelaziz, K., Hodgins, D. C., et al. (2021). Necrotic Enteritis in Chickens: a Review of Pathogenesis, Immune Responses and Prevention, Focusing on Probiotics and Vaccination. *Anim. Health. Res. Rev.* 22 (2), 147–162. doi:10.1017/s146625232100013x

Andersen, K. K., Strokappe, N. M., Hultberg, A., Truusalu, K., Smidt, I., Mikelsaar, R.-H., et al. (2016). Neutralization of *Clostridium difficile* Toxin B Mediated by Engineered *Lactobacilli* that Produce Single-Domain Antibodies. *Infect. Immun.* 84 (2), 395–406. doi:10.1128/iai.00870-15

Augustyn, W., Chruściel, A., Hreczuch, W., Kalka, J., Tarka, P., and Kierat, W. (2022). Inactivation of Spores and Vegetative Forms of *Clostridioides Difficile* by Chemical Biocides: Mechanisms of Biocidal Activity, Methods of Evaluation, and Environmental Aspects. *Ijerph* 19 (2), 750. doi:10.3390/ijerph19020750

Awad, S., Al-Ahmer, S. D., and Salih, S. M. (2020). Expression in *Saccharomyces Boulardii* of Recombinant Toxin-Coregulated Pilus a Subunit (TcpA) of *Vibrio cholerae O1*. *VacciMonitor* 29 (3), 136–142.

Awakawa, T., Fujioka, T., Zhang, L., Hoshino, S., Hu, Z., Hashimoto, J., et al. (2018). Reprogramming of the Antimycin NRPS-PKS Assembly Lines Inspired by Gene Evolution. *Nat. Commun.* 9 (1), 3534. doi:10.1038/s41467-018-05877-z

Bagherpour, G., Ghasemi, H., Zand, B., Zarei, N., Roohvand, F., Ardakani, E. M., et al. (2018). Oral Administration of Recombinant *Saccharomyces Boulardii* Expressing Ovalbumin-CPE Fusion Protein Induces Antibody Response in Mice. *Front. Microbiol.* 9, 723. doi:10.3389/fmicb.2018.00723

Bai, J., Qiao, X., Ma, Y., Han, M., Jia, S., Huang, X., et al. (2020). Protection Efficacy of Oral Bait Probiotic Vaccine Constitutively Expressing Tetravalent Toxoids against *Clostridium perfringens* Exotoxins in Livestock (Rabbits). *Vaccines* 8 (1), 17. doi:10.3390/vaccines8010017

Bastos, M. d. C. d. F., Coutinho, B. G., and Coelho, M. L. V. (2010). Lysostaphin: A *Staphylococcal* Bacteriolysin with Potential Clinical Applications. *Pharmaceuticals* 3 (4), 1139–1161. doi:10.3390/ph3041139

Bickel, M. H. (1988). The Development of Sulfonamides (1932-1938) as a Focal Point in the History of Chemotherapy. *Ges* 45, 67–86. doi:10.1163/22977953-04501006

Bintsis, T. (2017). Foodborne Pathogens. *AIMS Microbiol.* 3 (3), 529–563. doi:10.3934/microbiol.2017.3.529

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic acids Res.* 47 (W1), W81–W87. doi:10.1093/nar/gkz310

Bober, J. R., Beisel, C. L., and Nair, N. U. (2018). Synthetic Biology Approaches to Engineer Probiotics and Members of the Human Microbiota for Biomedical Applications. *Annu. Rev. Biomed. Eng.* 20, 277–300. doi:10.1146/annurev-bioeng-062117-121019

Bryer, M. S., Schoenbach, E. B., Chandler, C. A., Bliss, E. A., and Long, P. H. (1948). Aureomycin; Experimental and Clinical Investigations. *Jama* 138 (2), 117–119. doi:10.1001/jama.1948.02900020013004

Castro-Córdova, P., Mora-Uribe, P., Reyes-Ramírez, R., Cofré-Araneda, G., Orozco-Aguilar, J., Brito-Silva, C., et al. (2021). Entry of Spores into Intestinal Epithelial Cells Contributes to Recurrence of *Clostridioides Difficile* Infection. *Nat. Commun.* 12 (1), 1140. doi:10.1038/s41467-021-21355-5

Chandrasekaran, R., and Lacy, D. B. (2017). The Role of Toxins in *Clostridium difficile* Infection. *FEMS Microbiol. Rev.* 41 (6), 723–750. doi:10.1093/femsre/fux048

Chang, S., Sievert, D. M., Hageman, J. C., Boulton, M. L., Tenover, F. C., Downes, F. P., et al. (2003). Infection with Vancomycin-Resistant *Staphylococcus aureus* Containing the vanA Resistance Gene. *N. Engl. J. Med.* 348 (14), 1342–1347. doi:10.1056/NEJMoa025025

Chen, K., Zhu, Y., Zhang, Y., Hamza, T., Yu, H., Saint Fleur, A., et al. (2020). A Probiotic Yeast-Based Immunotherapy against *Clostridioides Difficile* Infection. *Sci. Transl. Med.* 12 (567), eaax4905. doi:10.1126/scitranslmed.aax4905

Cho, J.-H., Kwon, J.-G., O'Sullivan, D. J., Ryu, S., and Lee, J.-H. (2021a). Development of an Endolysin Enzyme and its Cell Wall–Binding Domain Protein and Their Applications for Biocontrol and Rapid Detection of *Clostridium perfringens* in Food. *Food Chem.* 345, 128562. doi:10.1016/j.foodchem.2020.128562

Cho, J. Y., Liu, R., Macbeth, J. C., and Hsiao, A. (2021b). The Interface of *Vibrio cholerae* and the Gut Microbiome. *Gut Microbes* 13 (1), 1937015. doi:10.1080/19490976.2021.1937015

Chopra, I., Hesse, L., O'Neill, A., and Connor, P. M. (2002). "Discovery and Development of New Anti-bacterial Drugs," in *Pharmacochemistry Library*. Editor H. van der Goot (Amsterdam: Elsevier), 213–225. doi:10.1016/s0165-7208(02)80022-8

Chuang, W. Y., Lin, W. C., Hsieh, Y. C., Huang, C. M., Chang, S. C., and Lee, T.-T. (2019). Evaluation of the Combined Use of *Saccharomyces cerevisiae* and *Aspergillus oryzae* with Phytase Fermentation Products on Growth, Inflammatory, and Intestinal Morphology in Broilers. *Animals* 9 (12), 1051. doi:10.3390/ani9121051

Clemens, J. D., Nair, G. B., Ahmed, T., Qadri, F., and Holmgren, J. (2017). Cholera. *Lancet* 390 (10101), 1539–1549. doi:10.1016/S0140-6736(17)30559-7

Cui, Q.-q., Yang, J., Sun, S.-j., Li, Z.-r., Qiang, C.-x., Niu, Y.-n., et al. (2021). Carriage of *Clostridioides Difficile* in Healthy Infants in the Community of Handan, China: A 1-year Follow-Up Study. *Anaerobe* 67, 102295. doi:10.1016/j.anaerobe.2020.102295

Czerucka, D., Piche, T., and Rampal, P. (2007). Review Article: Yeast as Probiotics - *Saccharomyces Boulardii*. *Aliment. Pharmacol. Ther.* 26 (6), 767–778. doi:10.1111/j.1365-2036.2007.03442.x

D'Costa, V. M., McGrann, K. M., Hughes, D. W., and Wright, G. D. (2006). Sampling the Antibiotic Resistome. *Science* 311 (5759), 374–377. doi:10.1126/science.1120800

de Bruyn, G., Saleh, J., Workman, D., Pollak, R., Elinoff, V., Fraser, N. J., et al. (2016). Defining the Optimal Formulation and Schedule of a Candidate Toxoid Vaccine against *Clostridium difficile* Infection: A Randomized Phase 2 Clinical Trial. *Vaccine* 34 (19), 2170–2178. doi:10.1016/j.vaccine.2016.03.028

de Bruyn, G., Gordon, D. L., Steiner, T., Tambyah, P., Cosgrove, C., Martens, M., et al. (2021). Safety, Immunogenicity, and Efficacy of a *Clostridioides Difficile* Toxoid Vaccine Candidate: a Phase 3 Multicentre, Observer-Blind, Randomised, Controlled Trial. *Lancet Infect. Dis.* 21 (2), 252–262. doi:10.1016/s1473-3099(20)30331-5

Díaz, M. A., Alberto, M. R., Vega-Hissi, E. G., González, S. N., and Arena, M. E. (2021). Interference in *Staphylococcus aureus* Biofilm and Virulence Factors Production by Human Probiotic Bacteria with Antimutagenic Activity. *Arab. J. Sci. Eng.* 47, 241–253. doi:10.1007/s13369-021-05934-8

Dieterle, M. G., Rao, K., and Young, V. B. (2019). Novel Therapies and Preventative Strategies for Primary and recurrentClostridium Difficileinfections. *Ann. N.Y. Acad. Sci.* 1435 (1), 110–138. doi:10.1111/nyas.13958

Domagk, G. (1935). Ein beitrag zur chemotherapie der bakteriellen infektionen. *Dtsch. Med. Wochenschr.* 61, 250–253. doi:10.1055/s-0028-1129486

Donald, R. G. K., Flint, M., Kalyan, N., Johnson, E., Witko, S. E., Kotash, C., et al. (2013). A Novel Approach to Generate a Recombinant Toxoid Vaccine against *Clostridium difficile*. *Microbiol. Read.* 159 (Pt 7), 1254–1266. doi:10.1099/mic.0.066712-0

Duan, F., and March, J. C. (2010). Engineered Bacterial Communication Prevents *Vibrio cholerae* Virulence in an Infant Mouse Model. *Proc. Natl. Acad. Sci. U.S.A.* 107 (25), 11260–11264. doi:10.1073/pnas.1001294107

Duchon, J., Barbian, M. E., and Denning, P. W. (2021). Necrotizing Enterocolitis. *Clin. Perinatol.* 48 (2), 229–250. doi:10.1016/j.clp.2021.03.002

Duggar, B. M. (1948). Aureomycin: a Product of the Continuing Search for New Antibiotics. *Ann. N. Y. Acad. Sci.* 51 (Art. 2), 177–181. doi:10.1111/j.1749-6632.1948.tb27262.x

Durand, G. A., Raoult, D., and Dubourg, G. (2019). Antibiotic Discovery: History, Methods and Perspectives. *Int. J. Antimicrob. agents* 53 (4), 371–382. doi:10.1016/j.ijantimicag.2018.11.010

Ehrlich, P. (1913). Address in Pathology, on Chemotherapy: Delivered before the Seventeenth International Congress of Medicine. *Bmj* 2 (2746), 353–359. doi:10.1136/bmj.2.2746.353

Esposito, G., Corpetti, C., Pesce, M., Seguella, L., Annunziata, G., Del Re, A., et al. (2021). A Palmitoylethanolamide Producing *Lactobacillus Paracasei* Improves *Clostridium difficile* Toxin A-Induced Colitis. *Front. Pharmacol.* 12, 639728. doi:10.3389/fphar.2021.639728

Faruque, S. M., Albert, M. J., and Mekalanos, J. J. (1998). Epidemiology, Genetics, and Ecology of Toxigenic *Vibrio cholerae*. *Microbiol. Mol. Biol. Rev.* 62 (4), 1301–1314. doi:10.1128/MMBR.62.4.1301-1314.1998

Fasano, A. (2002). Toxins and the Gut: Role in Human Disease. *Gut* 50 (Suppl. 3), Iii9–iii14. doi:10.1136/gut.50.suppl_3.iii9

Fischetti, V. A. (2006). Using Phage Lytic Enzymes to Control Pathogenic Bacteria. *BMC Oral Health* 6 (Suppl. 1), S16. doi:10.1186/1472-6831-6-s1-s16

Fischetti, V. A. (2010). Bacteriophage Endolysins: a Novel Anti-infective to Control Gram-Positive Pathogens. *Int. J. Med. Microbiol.* 300 (6), 357–362. doi:10.1016/j.ijmm.2010.04.002

Flaugnatti, N., Isaac, S., Lemos Rocha, L. F., Stutzmann, S., Rendueles, O., Stoudmann, C., et al. (2021). Human Commensal Gut *Proteobacteria* Withstand Type VI Secretion Attacks through Immunity Protein-independent Mechanisms. *Nat. Commun.* 12 (1), 5751. doi:10.1038/s41467-021-26041-0

Fleming, A. (1929). On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to Their Use in the Isolation of *B. Influenzæ*. *Br. J. Exp. Pathol.* 10 (3), 226–236.

Focareta, A., Paton, J. C., Morona, R., Cook, J., and Paton, A. W. (2006). A Recombinant Probiotic for Treatment and Prevention of Cholera. *Gastroenterology* 130 (6), 1688–1695. doi:10.1053/j.gastro.2006.02.005

Fredens, J., Wang, K., de la Torre, D., Funke, L. F. H., Robertson, W. E., Christova, Y., et al. (2019). Total Synthesis of *Escherichia c*oli with a Recoded Genome. *Nature* 569 (7757), 514–518. doi:10.1038/s41586-019-1192-5

Freter, R. (1955). The Fatal Enteric Cholera Infection in the guinea Pig, Achieved by Inhibition of Normal Enteric Flora. *J. Infect. Dis.* 97 (1), 57–65. doi:10.1093/infdis/97.1.57

Gagliardi, A., Totino, V., Cacciotti, F., Iebba, V., Neroni, B., Bonfiglio, G., et al. (2018). Rebuilding the Gut Microbiota Ecosystem. *Ijerph* 15 (8), 1679. doi:10.3390/ijerph15081679

Gao, X., Ma, Y., Wang, Z., Bai, J., Jia, S., Feng, B., et al. (2019). Oral Immunization of Mice with a Probiotic *Lactobacillus Casei* Constitutively Expressing the Alpha-Toxoid Induces Protective Immunity against *Clostridium perfringens* Alpha-Toxin. *Virulence* 10 (1), 166–179. doi:10.1080/21505594.2019.1582975

Gardlik, R., Palffy, R., and Celec, P. (2012). Recombinant Probiotic Therapy in Experimental Colitis in Mice. *Folia Biol. (Praha)* 58 (6), 238–245.

Garrido, V., Piñero-Lambea, C., Rodriguez-Arce, I., Paetzold, B., Ferrar, T., Weber, M., et al. (2021). Engineering a Genome-reduced Bacterium to Eliminate *Staphylococcus aureus* Biofilms *In Vivo*. *Mol. Syst. Biol.* 17 (10), e10145. doi:10.15252/msb.202010145

Gehrke, E. J., Zhang, X., Pimentel-Elardo, S. M., Johnson, A. R., Rees, C. A., Jones, S. E., et al. (2019). Silencing Cryptic Specialized Metabolism in *Streptomyces* by the Nucleoid-Associated Protein Lsr2. *Elife* 8, e47691. doi:10.7554/eLife.47691

Gelpi, A., Gilbertson, A., and Tucker, J. D. (2015). Magic Bullet: Paul Ehrlich, Salvarsan and the Birth of Venereology. *Sex. Transm. Infect.* 91 (1), 2–69. doi:10.1136/sextrans-2014-051779

Gervasi, T., Horn, N., Wegmann, U., Dugo, G., Narbad, A., and Mayer, M. J. (2014a). Expression and Delivery of an Endolysin to Combat *Clostridium perfringens*. *Appl. Microbiol. Biotechnol.* 98 (6), 2495–2505. doi:10.1007/s00253-013-5128-y

Gervasi, T., Lo Curto, R., Minniti, E., Narbad, A., and Mayer, M. J. (2014b). Application of *Lactobacillus Johnsonii* Expressing Phage Endolysin for Control of *Clostridium perfringens*. *Lett. Appl. Microbiol.* 59 (4), 355–361. doi:10.1111/lam.12298

Ghatak, S., King, Z. A., Sastry, A., and Palsson, B. O. (2019). The Y-Ome Defines the 35% of *Escherichia coli* Genes that Lack Experimental Evidence of Function. *Nucleic Acids Res.* 47 (5), 2446–2454. doi:10.1093/nar/gkz030

Golneshin, A., Gor, M.-C., Williamson, N., Vezina, B., Van, T. T. H., May, B. K., et al. (2020). Discovery and Characterisation of Circular Bacteriocin Plantacyclin B21AG from *Lactiplantibacillus Plantarum* B21. *Heliyon* 6 (8), e04715. doi:10.1016/j.heliyon.2020.e04715

Gong, L., Wang, B., Zhou, Y., Tang, L., Zeng, Z., Zhang, H., et al. (2020). Protective Effects of *Lactobacillus Plantarum* 16 and *Paenibacillus Polymyxa* 10 against *Clostridium perfringens* Infection in Broilers. *Front. Immunol.* 11, 628374. doi:10.3389/fimmu.2020.628374

Gorbach, S., Chang, T.-W., and Goldin, B. (1987). Successful Treatment of Relapsing *Clostridium difficile* Colitis with *Lactobacillus GG*. *Lancet* 330 (8574), 1519. doi:10.1016/s0140-6736(87)92646-8

Graham, J. P., Boland, J. J., and Silbergeld, E. (2007). Growth Promoting Antibiotics in Food Animal Production: an Economic Analysis. *Public Health Rep.* 122 (1), 79–87. doi:10.1177/003335490712200111

Greenberg, R. N., Marbury, T. C., Foglia, G., and Warny, M. (2012). Phase I Dose Finding Studies of an Adjuvanted *Clostridium difficile* Toxoid Vaccine. *Vaccine* 30 (13), 2245–2249. doi:10.1016/j.vaccine.2012.01.065

Gunaratnam, S., Millette, M., McFarland, L. V., DuPont, H. L., and Lacroix, M. (2021). Potential Role of Probiotics in Reducing *Clostridioides Difficile* Virulence: Interference with Quorum Sensing Systems. *Microb. Pathog.* 153, 104798. doi:10.1016/j.micpath.2021.104798

Guo, S., Yan, W., McDonough, S. P., Lin, N., Wu, K. J., He, H., et al. (2015). The Recombinant *Lactococcus Lactis* Oral Vaccine Induces Protection against *C. difficile* Spore Challenge in a Mouse Model. *Vaccine* 33 (13), 1586–1595. doi:10.1016/j.vaccine.2015.02.006

Guo, S., Liu, D., Zhang, B., Li, Z., Li, Y., Ding, B., et al. (2017). Two Lactobacillus Species Inhibit the Growth and α-Toxin Production of *Clostridium perfringens* and Induced Proinflammatory Factors in Chicken Intestinal Epithelial Cells *In Vitro*. *Front. Microbiol.* 8, 2081. doi:10.3389/fmicb.2017.02081

Hassani, S., Pakbin, B., Brück, W. M., Mahmoudi, R., and Mousavi, S. (2022). Prevalence, Antibiotic Resistance, Toxin-Typing and Genotyping of *Clostridium perfringens* in Raw Beef Meats Obtained from Qazvin City, Iran. *Antibiotics* 11 (3), 340. doi:10.3390/antibiotics11030340

Heo, S., Kim, M. G., Kwon, M., Lee, H. S., and Kim, G. B. (2018). Inhibition of *Clostridium perfringens* Using Bacteriophages and Bacteriocin Producing Strains. *Korean J. Food Sci. Anim. Resour.* 38 (1), 88–98. doi:10.5851/kosfa.2018.38.1.88

Herrington, D. A., Hall, R. H., Losonsky, G., Mekalanos, J. J., Taylor, R. K., and Levine, M. M. (1988). Toxin, Toxin-Coregulated Pili, and the toxR Regulon Are Essential for *Vibrio cholerae* Pathogenesis in Humans. *J. Exp. Med.* 168 (4), 1487–1492. doi:10.1084/jem.168.4.1487

Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B.-C., and Herrmann, R. (1996). Complete Sequence Analysis of the Genome of the Bacterium *Mycoplasma Pneumoniae*. *Nucleic Acids Res.* 24 (22), 4420–4449. doi:10.1093/nar/24.22.4420

Hoang, T. H., Hong, H. A., Clark, G. C., Titball, R. W., and Cutting, S. M. (2008). Recombinant Bacillus Subtilis Expressing the *Clostridium perfringens* Alpha Toxoid Is a Candidate Orally Delivered Vaccine against Necrotic Enteritis. *Infect. Immun.* 76 (11), 5257–5265. doi:10.1128/IAI.00686-08

Holowko, M. B., Wang, H., Jayaraman, P., and Poh, C. L. (2016). Biosensing *Vibrio cholerae* with Genetically Engineered *Escherichia coli*. *ACS Synth. Biol.* 5 (11), 1275–1283. doi:10.1021/acssynbio.6b00079

Hooks, K. B., and O'Malley, M. A. (2017). Dysbiosis and its Discontents. *mBio* 8 (5), e01492–17. doi:10.1128/mBio.01492-17

Hopwood, D. A., Malpartida, F., Kieser, H. M., Ikeda, H., Duncan, J., Fujii, I., et al. (1985). Production of 'hybrid' Antibiotics by Genetic Engineering. *Nature* 314 (6012), 642–644. doi:10.1038/314642a0

Hsiao, A., and Zhu, J. (2020). Pathogenicity and Virulence Regulation of *Vibrio cholerae* at the Interface of Host-Gut Microbiome Interactions. *Virulence* 11 (1), 1582–1599. doi:10.1080/21505594.2020.1845039

Hutchings, M. I., Truman, A. W., and Wilkinson, B. (2019). Antibiotics: Past, Present and Future. *Curr. Opin. Microbiol.* 51, 72–80. doi:10.1016/j.mib.2019.10.008

Iqbal, Z., Ahmed, S., Tabassum, N., Bhattacharya, R., and Bose, D. (2021). Role of Probiotics in Prevention and Treatment of Enteric Infections: a Comprehensive Review. *3 Biotech.* 11 (5), 242. doi:10.1007/s13205-021-02796-7

Jayaraman, P., Holowko, M. B., Yeoh, J. W., Lim, S., and Poh, C. L. (2017). Repurposing a Two-Component System-Based Biosensor for the Killing of *Vibrio cholerae*. *ACS Synth. Biol.* 6 (7), 1403–1415. doi:10.1021/acssynbio.7b00058

Jiang, Y., Mo, H., Willingham, C., Wang, S., Park, J.-y., Kong, W., et al. (2015). Protection against Necrotic Enteritis in Broiler Chickens by Regulated Delayed LysisSalmonellaVaccines. *Avian Dis.* 59 (4), 475–485. doi:10.1637/11094-041715-Reg

Jones, D., Metzger, H. J., Schatz, A., and Waksman, S. A. (1944). Control of Gram-Negative Bacteria in Experimental Animals by Streptomycin. *Science* 100 (2588), 103–105. doi:10.1126/science.100.2588.103.b

Karyal, C., Hughes, J., Kelly, M. L., Luckett, J. C., Kaye, P. V., Cockayne, A., et al. (2021). Colonisation Factor CD0873, an Attractive Oral Vaccine Candidate against *Clostridioides Difficile*. *Microorganisms* 9 (2), 306. doi:10.3390/microorganisms9020306

Kaviani, A., Ince, D., and Axelrod, D. A. (2020). Management of Antimicrobial Agents in Abdominal Organ Transplant Patients in Intensive Care Unit. *Curr. Transpl. Rep.* 7 (1), 1–11. doi:10.1007/s40472-020-00268-0

Kelesidis, T., and Pothoulakis, C. (2012). Efficacy and Safety of the Probiotic *Saccharomyces Boulardii* for the Prevention and Therapy of Gastrointestinal Disorders. *Ther. Adv. Gastroenterol.* 5 (2), 111–125. doi:10.1177/1756283X11428502

Khalique, A., Zeng, D., Shoaib, M., Wang, H., Qing, X., Rajput, D. S., et al. (2020). Probiotics Mitigating Subclinical Necrotic Enteritis (SNE) as Potential Alternatives to Antibiotics in Poultry. *Amb. Expr.* 10 (1), 50. doi:10.1186/s13568-020-00989-6

Khan, I., Bai, Y., Zha, L., Ullah, N., Ullah, H., Shah, S. R. H., et al. (2021). Mechanism of the Gut Microbiota Colonization Resistance and Enteric Pathogen Infection. *Front. Cell. Infect. Microbiol.* 11, 716299. doi:10.3389/fcimb.2021.716299

Khoruts, A., Staley, C., and Sadowsky, M. J. (2021). Faecal Microbiota Transplantation for *Clostridioides Difficile*: Mechanisms and Pharmacology. *Nat. Rev. Gastroenterol. Hepatol.* 18 (1), 67–80. doi:10.1038/s41575-020-0350-4

Kirk, M. D., Pires, S. M., Black, R. E., Caipo, M., Crump, J. A., Devleesschauwer, B., et al. (2015). World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLoS Med.* 12 (12), e1001921. doi:10.1371/journal.pmed.1001921

Kiu, R., and Hall, L. J. (2018). An Update on the Human and Animal Enteric Pathogen *Clostridium perfringens*. *Emerg. Microbes Infect.* 7 (1), 1–15. doi:10.1038/s41426-018-0144-8

Kokai-Kun, J. F., Chanturiya, T., and Mond, J. J. (2007). Lysostaphin as a Treatment for Systemic *Staphylococcus aureus* Infection in a Mouse Model. *J. Antimicrob. Chemother.* 60 (5), 1051–1059. doi:10.1093/jac/dkm347

Krishnakumar, R., Assad-Garcia, N., Benders, G. A., Phan, Q., Montague, M. G., and Glass, J. I. (2010). Targeted Chromosomal Knockouts in *Mycoplasma Pneumoniae*. *Appl. Environ. Microbiol.* 76 (15), 5297–5299. doi:10.1128/aem.00024-10

Kuehne, S. A., Collery, M. M., Kelly, M. L., Cartman, S. T., Cockayne, A., and Minton, N. P. (2014). Importance of Toxin A, Toxin B, and CDT in Virulence of an Epidemic *Clostridium difficile* Strain. *J. Infect. Dis.* 209 (1), 83–86. doi:10.1093/infdis/jit426

Lamps, L. W. (2009). "Infectious Disorders of the GI Tract," in *Surgical Pathology of the GI Tract, Liver, Biliary Tract, and Pancreas (Second Edition)*. Editor R. D. Odze and J. R. Goldblum (Philadelphia: W. B. Saunders), Chap. 4, 51–79. doi:10.1016/B978-14160-4059-0.X50001-2

Larsen, J., Raisen, C. L., Ba, X., Sadgrove, N. J., Padilla-González, G. F., Simmonds, M. S. J., et al. (2022). Emergence of Methicillin Resistance Predates the Clinical Use of Antibiotics. *Nature* 602, 135–141. doi:10.1038/s41586-021-04265-w

Lee, J. S., Park, S. W., Lee, H. B., and Kang, S.-S. (2021). Bacteriocin-like Inhibitory Substance (BLIS) Activity of *Enterococcus Faecium* Db1 against Biofilm Formation by *Clostridium perfringens*. *Probiotics Antimicro. Prot.* 13 (5), 1452–1457. doi:10.1007/s12602-021-09813-4

Lessa, F. C., Mu, Y., Bamberg, W. M., Beldavs, Z. G., Dumyati, G. K., Dunn, J. R., et al. (2015a). Burden ofClostridium difficileInfection in the United States. *N. Engl. J. Med.* 372 (9), 825–834. doi:10.1056/NEJMoa1408913

Lessa, F. C., Winston, L. G., and McDonald, L. C.Emerging Infections Program C; difficile Surveillance Team (2015b). Burden of *Clostridium difficile* Infection in the United States. *N. Engl. J. Med.* 372 (24), 2368–2370. doi:10.1056/NEJMc1505190

Li, Y., Li, Z., Yamanaka, K., Xu, Y., Zhang, W., Vlamakis, H., et al. (2015). Directed Natural Product Biosynthesis Gene Cluster Capture and Expression in the Model Bacterium *Bacillus Subtilis*. *Sci. Rep.* 5, 9383. doi:10.1038/srep09383

Li, R., Wan, X., Takala, T. M., and Saris, P. E. J. (2021). Heterologous Expression of the *Leuconostoc* Bacteriocin Leucocin C in Probiotic Yeast *Saccharomyces Boulardii*. *Probiotics Antimicro. Prot.* 13 (1), 229–237. doi:10.1007/s12602-020-09676-1

Lindström, M., Heikinheimo, A., Lahti, P., and Korkeala, H. (2011). Novel Insights into the Epidemiology of *Clostridium perfringens* Type A Food Poisoning. *Food Microbiol.* 28 (2), 192–198. doi:10.1016/j.fm.2010.03.020

Logan, S. L., Thomas, J., Yan, J., Baker, R. P., Shields, D. S., Xavier, J. B., et al. (2018). The *Vibrio cholerae* Type VI Secretion System Can Modulate Host Intestinal Mechanics to Displace Gut Bacterial Symbionts. *Proc. Natl. Acad. Sci. U.S.A.* 115 (16), E3779–E3787. doi:10.1073/pnas.1720133115

Lubkowicz, D., Ho, C. L., Hwang, I. Y., Yew, W. S., Lee, Y. S., and Chang, M. W. (2018). Reprogramming Probiotic Lactobacillus Reuteri as a Biosensor for *Staphylococcus aureus* Derived AIP-I Detection. *ACS Synth. Biol.* 7 (5), 1229–1237. doi:10.1021/acssynbio.8b00063

Lv, J., Da, R., Cheng, Y., Tuo, X., Wei, J., Jiang, K., et al. (2020). Mechanism of Antibacterial Activity of *Bacillus Amyloliquefaciens C-1* Lipopeptide toward Anaerobic *Clostridium difficile*. *BioMed Res. Int.* 2020, 1–12. doi:10.1155/2020/3104613

Mao, D., Okada, B. K., Wu, Y., Xu, F., and Seyedsayamdost, M. R. (2018a). Recent Advances in Activating Silent Biosynthetic Gene Clusters in Bacteria. *Curr. Opin. Microbiol.* 45, 156–163. doi:10.1016/j.mib.2018.05.001

Mao, N., Cubillos-Ruiz, A., Cameron, D. E., and Collins, J. J. (2018b). Probiotic Strains Detect and Suppress Cholera in Mice. *Sci. Transl. Med.* 10 (445), eaao2586. doi:10.1126/scitranslmed.aao2586

Marra, A. R., Perencevich, E. N., Nelson, R. E., Samore, M., Khader, K., Chiang, H.-Y., et al. (2020). Incidence and Outcomes Associated WithClostridium difficileInfections. *JAMA Netw. Open* 3 (1), e1917597. doi:10.1001/jamanetworkopen.2019.17597

Martens, E. C., Neumann, M., and Desai, M. S. (2018). Interactions of Commensal and Pathogenic Microorganisms with the Intestinal Mucosal Barrier. *Nat. Rev. Microbiol.* 16 (8), 457–470. doi:10.1038/s41579-018-0036-x

Mathipa, M. G., and Thantsha, M. S. (2017). Probiotic Engineering: towards Development of Robust Probiotic Strains with Enhanced Functional Properties and for Targeted Control of Enteric Pathogens. *Gut Pathog.* 9, 28. doi:10.1186/s13099-017-0178-9

Mayer, M. J., Narbad, A., and Gasson, M. J. (2008). Molecular Characterization of a *Clostridium difficile* Bacteriophage and its Cloned Biologically Active Endolysin. *J. Bacteriol.* 190 (20), 6734–6740. doi:10.1128/jb.00686-08

Mazzantini, D., Calvigioni, M., Celandroni, F., Lupetti, A., and Ghelardi, E. (2021). Spotlight on the Compositional Quality of Probiotic Formulations Marketed Worldwide. *Front. Microbiol.* 12, 693973. doi:10.3389/fmicb.2021.693973

McDonald, L. C., Gerding, D. N., Johnson, S., Bakken, J. S., Carroll, K. C., Coffin, S. E., et al. (2018). Clinical Practice Guidelines for *Clostridium difficile* Infection in Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). *Clin. Infect. Dis.* 66 (7), e1–e48. doi:10.1093/cid/cix1085

McFarland, L. V., Surawicz, C. M., Greenberg, R. N., Fekety, R., Elmer, G. W., Moyer, K. A., et al. (1994). A Randomized Placebo-Controlled Trial of *Saccharomyces Boulardii* in Combination with Standard Antibiotics for *Clostridium difficile* Disease. *Jama* 271 (24), 1913–1918. doi:10.1001/jama.271.24.1913

McGovern, B. H., Ford, C. B., Henn, M. R., Pardi, D. S., Khanna, S., Hohmann, E. L., et al. (2021). SER-109, an Investigational Microbiome Drug to Reduce Recurrence after *Clostridioides Difficile* Infection: Lessons Learned from a Phase 2 Trial. *Clin. Infect. Dis.* 72 (12), 2132–2140. doi:10.1093/cid/ciaa387

Meader, E., Mayer, M. J., Steverding, D., Carding, S. R., and Narbad, A. (2013). Evaluation of Bacteriophage Therapy to Control *Clostridium difficile* and Toxin

Production in an *In Vitro* Human Colon Model System. *Anaerobe* 22, 25–30. doi:10.1016/j.anaerobe.2013.05.001

Medema, M. H., and Fischbach, M. A. (2015). Computational Approaches to Natural Product Discovery. *Nat. Chem. Biol.* 11 (9), 639–648. doi:10.1038/nchembio.1884

Mingyar, E., Mühling, L., Kulik, A., Winkler, A., Wibberg, D., Kalinowski, J., et al. (2021). A Regulator Based "Semi-targeted" Approach to Activate Silent Biosynthetic Gene Clusters. *Ijms* 22 (14), 7567. doi:10.3390/ijms22147567

Mishra, N., and Smyth, J. A. (2017). Oral Vaccination of Broiler Chickens against Necrotic Enteritis Using a Non-virulent NetB Positive Strain of *Clostridium perfringens* Type A. *Vaccine* 35, 6858–6865. doi:10.1016/j.vaccine.2017.10.030

Mondal, S. I., Draper, L. A., Ross, R. P., and Hill, C. (2020). Bacteriophage Endolysins as a Potential Weapon to Combat *Clostridioides Difficile* Infection. *Gut Microbes* 12 (1), 1813533. doi:10.1080/19490976.2020.1813533

Moono, P., Foster, N. F., Hampson, D. J., Knight, D. R., Bloomfield, L. E., and Riley, T. V. (2016). *Clostridium difficile* Infection in Production Animals and Avian Species: A Review. *Foodborne Pathog. Dis.* 13 (12), 647–655. doi:10.1089/fpd.2016.2181

Moore, P. R., Evenson, A., Luckey, T. D., McCoy, E., Elvehjem, C. A., and Hart, E. B. (1946). Use of Sulfasuxidine, Streptothricin, and Streptomycin in Nutritional Studies with the Chick. *J. Biol. Chem.* 165 (2), 437–441. doi:10.1016/s0021-9258(17)41154-9

Mora, Z. V.-d. l., Macías-Rodríguez, M. E., Arratia-Quijada, J., Gonzalez-Torres, Y. S., Nuño, K., and Villarruel-López, A. (2020). *Clostridium perfringens* as Foodborne Pathogen in Broiler Production: Pathophysiology and Potential Strategies for Controlling Necrotic Enteritis. *Animals* 10 (9), 1718. doi:10.3390/ani10091718

Mousavi Khaneghah, A., Abhari, K., Eş, I., Soares, M. B., Oliveira, R. B. A., Hosseini, H., et al. (2020). Interactions between Probiotics and Pathogenic Microorganisms in Hosts and Foods: A Review. *Trends Food Sci. Technol.* 95, 205–218. doi:10.1016/j.tifs.2019.11.022

Murray, C. J. L., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., et al. (2022). Global Burden of Bacterial Antimicrobial Resistance in 2019: a Systematic Analysis. *Lancet* 399 (10325), 629–655. doi:10.1016/S0140-6736(21)02724-0

Nagarajan, V., Peng, M., Tabashsum, Z., Salaheen, S., Padilla, J., and Biswas, D. (2019). Antimicrobial Effect and Probiotic Potential of Phage Resistant *Lactobacillus Plantarum* and its Interactions with Zoonotic Bacterial Pathogens. *Foods* 8 (6), 194. doi:10.3390/foods8060194

Nale, J. Y., Spencer, J., Hargreaves, K. R., Buckley, A. M., Trzepiński, P., Douce, G. R., et al. (2016). Bacteriophage Combinations Significantly Reduce *Clostridium difficile* Growth *In Vitro* and Proliferation *In Vivo*. *Antimicrob. Agents Chemother.* 60 (2), 968–981. doi:10.1128/aac.01774-15

Nataraj, B. H., and Mallappa, R. H. (2021). Antibiotic Resistance Crisis: An Update on Antagonistic Interactions between Probiotics and Methicillin-Resistant *Staphylococcus aureus* (MRSA). *Curr. Microbiol.* 78 (6), 2194–2211. doi:10.1007/s00284-021-02442-8

Navarro, M., McClane, B., and Uzal, F. (2018). Mechanisms of Action and Cell Death Associated with *Clostridium perfringens* Toxins. *Toxins* 10 (5), 212. doi:10.3390/toxins10050212

Newton, G. G. F., and Abraham, E. P. (1955). Cephalosporin C, a New Antibiotic Containing Sulphur and D-α-Aminoadipic Acid. *Nature* 175 (4456), 548. doi:10.1038/175548a0

Nicolaou, K. C., and Rigol, S. (2018). A Brief History of Antibiotics and Select Advances in Their Synthesis. *J. Antibiot.* 71 (2), 153–184. doi:10.1038/ja.2017.62

O'Neil, J. (2016). "Tackling Drug-Resistant Infections Globally: Final Report and Recommendations," in *Review on Antimicrobial Resistance* London, UK.

Olsufyeva, E. N., and Yankovskaya, V. S. (2020). Main Trends in the Design of Semi-synthetic Antibiotics of a New Generation. *Russ. Chem. Rev.* 89 (3), 339–378. doi:10.1070/RCR4892

Onderdonk, A. B., and Garrett, W. S. C. P. M. (2020). "Diseases Caused by Clostridium," in *Principles and Practice of Infectious Diseases*. Editors J. E. Bennett, R. Dolin, and M. J. B. Ninth (Philadelphia, PA: Elesvier), 2960–2968.e2962.

Osawa, S., Muto, A., Jukes, T. H., and Ohama, T. (1990). Evolutionary Changes in the Genetic Code. *Proc. Biol. Sci.* 241 (1300), 19–28. doi:10.1098/rspb.1990.0060

O'Shea, E. F., Gardiner, G. E., O'Connor, P. M., Mills, S., Ross, R. P., and Hill, C. (2009). Characterization of Enterocin- and Salivaricin-Producing Lactic Acid Bacteria from the Mammalian Gastrointestinal Tract. *FEMS Microbiol. Lett.* 291 (1), 24–34. doi:10.1111/j.1574-6968.2008.01427.x

Otto, M. (2013). *Staphylococcal* Infections: Mechanisms of Biofilm Maturation and Detachment as Critical Determinants of Pathogenicity. *Annu. Rev. Med.* 64, 175–188. doi:10.1146/annurev-med-042711-140023

Paik, W., Alonzo, F., 3rd, and Knight, K. L. (2019). Probiotic Exopolysaccharide Protects against Systemic *Staphylococcus aureus* Infection, Inducing Dual-Functioning Macrophages that Restrict Bacterial Growth and Limit Inflammation. *Infect. Immun.* 87 (1), e00791–18. doi:10.1128/IAI.00791-18

Paiva, D., Walk, C., and McElroy, A. (2014). Dietary Calcium, Phosphorus, and Phytase Effects on Bird Performance, Intestinal Morphology, Mineral Digestibility, and Bone Ash during a Natural Necrotic Enteritis Episode. *Poult. Sci.* 93 (11), 2752–2762. doi:10.3382/ps.2014-04148

Panpetch, W., Phuengmaung, P., Cheibchalard, T., Somboonna, N., Leelahavanichkul, A., and Tumwasorn, S. (2021). *Lacticaseibacillus Casei* Strain T21 Attenuates *Clostridioides Difficile* Infection in a Murine Model through Reduction of Inflammation and Gut Dysbiosis with Decreased Toxin Lethality and Enhanced Mucin Production. *Front. Microbiol.* 12, 745299. doi:10.3389/fmicb.2021.745299

Park, E. D., Lightner, D. V., and Park, D. L. (1994). Antimicrobials in Shrimp Aquaculture in the United States: Regulatory Status and Safety Concerns. *Rev. Environ. Contam. Toxicol.* 138, 1–20. doi:10.1007/978-1-4612-2672-7_1

Patel, S. J., Wellington, M., Shah, R. M., and Ferreira, M. J. (2020). Antibiotic Stewardship in Food-Producing Animals: Challenges, Progress, and Opportunities. *Clin. Ther.* 42 (9), 1649–1658. doi:10.1016/j.clinthera.2020.07.004

Pedrolli, D. B., Ribeiro, N. V., Squizato, P. N., de Jesus, V. N., Cozetto, D. A., Tuma, R. B., et al. (2019). Engineering Microbial Living Therapeutics: The Synthetic Biology Toolbox. *Trends Biotechnol.* 37 (1), 100–115. doi:10.1016/j.tibtech.2018.09.005

Peng, Z., Wang, S., Gide, M., Zhu, D., Lamabadu Warnakulasuriya Patabendige, H. M., Li, C., et al. (2018). A Novel Bacteriophage Lysin-Human Defensin Fusion Protein Is Effective in Treatment of *Clostridioides Difficile* Infection in Mice. *Front. Microbiol.* 9, 3234. doi:10.3389/fmicb.2018.03234

Pepin, J., Alary, M.-E., Valiquette, L., Raiche, E., Ruel, J., Fulop, K., et al. (2005). Increasing Risk of Relapse after Treatment of *Clostridium difficile* Colitis in Quebec, Canada. *Clin. Infect. Dis.* 40 (11), 1591–1597. doi:10.1086/430315

Pereira, F. C., Wasmund, K., Cobankovic, I., Jehmlich, N., Herbold, C. W., Lee, K. S., et al. (2020). Rational Design of a Microbial Consortium of Mucosal Sugar Utilizers Reduces *Clostridiodes Difficile* Colonization. *Nat. Commun.* 11 (1), 5104. doi:10.1038/s41467-020-18928-1

Permpoonpattana, P., Hong, H. A., Phetcharaburanin, J., Huang, J.-M., Cook, J., Fairweather, N. F., et al. (2011). Immunization with *Bacillus* Spores Expressing Toxin A Peptide Repeats Protects against Infection with *Clostridium difficile* Strains Producing Toxins A and B. *Infect. Immun.* 79 (6), 2295–2302. doi:10.1128/iai.00130-11

Pipiya, S. O., Mokrushina, Y. A., Gabibov, A. G., Smirnov, I. V., and Terekhov, S. S. (2020). Selective Eradication of *Staphylococcus aureus* by the Designer Genetically Programmed Yeast Biocontrol Agent. *Antibiotics* 9 (9), 527. doi:10.3390/antibiotics9090527

Popoff, M. R., and Bouvet, P. (2009). *Clostridial* Toxins. *Future Microbiol.* 4 (8), 1021–1064. doi:10.2217/fmb.09.72

Pruteanu, M., and Shanahan, F. (2013). Digestion of Epithelial Tight Junction Proteins by the commensalClostridium Perfringens. *Am. J. Physiol.-Gastrointest. Liver Physiol.* 305 (10), G740–G748. doi:10.1152/ajpgi.00316.2012

Pulingam, T., Parumasivam, T., Gazzali, A. M., Sulaiman, A. M., Chee, J. Y., Lakshmanan, M., et al. (2022). Antimicrobial Resistance: Prevalence, Economic Burden, Mechanisms of Resistance and Strategies to Overcome. *Eur. J. Pharm. Sci.* 170, 106103. doi:10.1016/j.ejps.2021.106103

Raplee, I., Walker, L., Xu, L., Surathu, A., Chockalingam, A., Stewart, S., et al. (2021). Emergence of Nosocomial Associated Opportunistic Pathogens in the Gut Microbiome after Antibiotic Treatment. *Antimicrob. Resist Infect. Control* 10 (1), 36. doi:10.1186/s13756-021-00903-0

Reading, N. C., and Sperandio, V. (2006). Quorum sensing: the Many Languages of Bacteria. *FEMS Microbiol. Lett.* 254 (1), 1–11. doi:10.1111/j.1574-6968.2005.00001.x

Riley, J. C. (2005). Estimates of Regional and Global Life Expectancy, 1800-2001. *Popul. Dev. Rev.* 31 (3), 537–543. doi:10.1111/j.1728-4457.2005.00083.x

Ritter, S. C., and Hackel, B. J. (2019). Validation and Stabilization of a Prophage Lysin of *Clostridium perfringens* by Using Yeast Surface Display and Coevolutionary Models. *Appl. Environ. Microbiol.* 85 (10), e00054–19. doi:10.1128/aem.00054-19

Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., et al. (2006). Production of the Antimalarial Drug Precursor Artemisinic Acid in Engineered Yeast. *Nature* 440 (7086), 940–943. doi:10.1038/nature04640

Robertson, W. E., Funke, L. F. H., de la Torre, D., Fredens, J., Elliott, T. S., Spinck, M., et al. (2021). Sense Codon Reassignment Enables Viral Resistance and Encoded Polymer Synthesis. *Science* 372 (6546), 1057–1062. doi:10.1126/science.abg3029

Ryan, E. T., Butterton, J. R., Smith, R. N., Carroll, P. A., Crean, T. I., and Calderwood, S. B. (1997). Protective Immunity against *Clostridium difficile* Toxin A Induced by Oral Immunization with a Live, Attenuated *Vibrio cholerae* Vector Strain. *Infect. Immun.* 65 (7), 2941–2949. doi:10.1128/iai.65.7.2941-2949.1997

Saidi, N., Owlia, P., Marashi, S. M. A., and Saderi, H. (2019). Inhibitory Effect of Probiotic Yeast *Saccharomyces cerevisiae* on Biofilm Formation and Expression of α-hemolysin and Enterotoxin A Genes of *Staphylococcus aureus*. *Iran. J. Microbiol.* 11 (3), 246–254.

Saleh, O., Bonitz, T., Flinspach, K., Kulik, A., Burkard, N., Mühlenweg, A., et al. (2012). Activation of a Silent Phenazine Biosynthetic Gene Cluster Reveals a Novel Natural Product and a New Resistance Mechanism against Phenazines. *Med. Chem. Commun.* 3 (8), 1009–1019. doi:10.1039/C2MD20045G

Sazawal, S., Hiremath, G., Dhingra, U., Malik, P., Deb, S., and Black, R. E. (2006). Efficacy of Probiotics in Prevention of Acute Diarrhoea: a Meta-Analysis of Masked, Randomised, Placebo-Controlled Trials. *Lancet Infect. Dis.* 6 (6), 374–382. doi:10.1016/S1473-3099(06)70495-9

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., et al. (2011). Foodborne Illness Acquired in the United States-Major Pathogens. *Emerg. Infect. Dis.* 17 (1), 7–15. doi:10.3201/eid1701.p11101

Schmelcher, M., Donovan, D. M., and Loessner, M. J. (2012). Bacteriophage Endolysins as Novel Antimicrobials. *Future Microbiol.* 7 (10), 1147–1171. doi:10.2217/fmb.12.97

Schoster, A., Kokotovic, B., Permin, A., Pedersen, P. D., Bello, F. D., and Guardabassi, L. (2013). *In Vitro* inhibition of *Clostridium difficile* and *Clostridium perfringens* by Commercial Probiotic Strains. *Anaerobe* 20, 36–41. doi:10.1016/j.anaerobe.2013.02.006

Sekurova, O. N., Schneider, O., and Zotchev, S. B. (2019). Novel Bioactive Natural Products from Bacteria via Bioprospecting, Genome Mining and Metabolic Engineering. *Microb. Biotechnol.* 12 (5), 828–844. doi:10.1111/1751-7915.13398

Sekurova, O. N., Sun, Y.-Q., Zehl, M., Rückert, C., Stich, A., Busche, T., et al. (2021). Coupling of the Engineered DNA "Mutator" to a Biosensor as a New Paradigm for Activation of Silent Biosynthetic Gene Clusters in *Streptomyces*. *Nucleic Acids Res.* 49 (14), 8396–8405. doi:10.1093/nar/gkab583

Selle, K., Fletcher, J. R., Tuson, H., Schmitt, D. S., McMillan, L., Vridhambal, G. S., et al. (2020). *In Vivo* Targeting of *Clostridioides Difficile* Using Phage-Delivered CRISPR-Cas3 Antimicrobials. *mBio* 11 (2), e00019–20. doi:10.1128/mBio.00019-20

Sholeh, M., Krutova, M., Forouzesh, M., Mironov, S., Sadeghifard, N., Molaeipour, L., et al. (2020). Antimicrobial Resistance in *Clostridioides (Clostridium) Difficile* Derived from Humans: a Systematic Review and Meta-Analysis. *Antimicrob. Resist Infect. Control* 9 (1), 158. doi:10.1186/s13756-020-00815-5

Sikorska, H., and Smoragiewicz, W. (2013). Role of Probiotics in the Prevention and Treatment of Meticillin-Resistant *Staphylococcus aureus* Infections. *Int. J. Antimicrob. Agents* 42 (6), 475–481. doi:10.1016/j.ijantimicag.2013.08.003

Slavić, D., Boerlin, P., Fabri, M., Klotins, K. C., Zoethout, J. K., Weir, P. E., et al. (2011). Antimicrobial Susceptibility of *Clostridium perfringens* Isolates of Bovine, Chicken, Porcine, and turkey Origin from Ontario. *Can. J. Vet. Res.* 75 (2), 89–97.

Sleator, R. D. (2010). The Human Superorganism - of Microbes and Men. *Med. Hypotheses* 74 (2), 214–215. doi:10.1016/j.mehy.2009.08.047

Sluijter, M., Kaptein, E., Spuesens, E. B. M., Hoogenboezem, T., Hartwig, N. G., Van Rossum, A. M. C., et al. (2010). The M*ycoplasma Genitalium* MG352-Encoded Protein Is a Holliday Junction Resolvase that Has a Non-functional Orthologue in *Mycoplasma Pneumoniae*. *Mol. Microbiol.* 77 (5), 1261–1277. doi:10.1111/j.1365-2958.2010.07288.x

Smith, A. L., Powers, C., and Beal, R. K. (2014). "Chapter 13 - the Avian Enteric Immune System in Health and Disease," in *Avian Immunology*. Editors K. A. Schat, B. Kaspers, and P. Kaiser. Second Edition (Boston: Academic Press), 227–250. doi:10.1016/b978-0-12-396965-1.00013-3

Sola-Oladokun, B., Culligan, E. P., and Sleator, R. D. (2017). Engineered Probiotics: Applications and Biological Containment. *Annu. Rev. Food Sci. Technol.* 8, 353–370. doi:10.1146/annurev-food-030216-030256

Song, B., Cui, H., Ju, L., Song, L., Tang, L., and Li, Y. (2018). Expression of the Alpha Toxin of *Clostridium perfringens* in *Lactobacillus Casei* Genome and Evaluation of its Immune Effects in Mice. *Microb. Pathog.* 118, 1–8. doi:10.1016/j.micpath.2018.03.010

Spinler, J. K., Auchtung, J., Brown, A., Boonma, P., Oezguen, N., Ross, C. L., et al. (2017). Next-Generation Probiotics Targeting *Clostridium difficile* through Precursor-Directed Antimicrobial Biosynthesis. *Infect. Immun.* 85 (10), e00303–17. doi:10.1128/IAI.00303-17

Suetens, C., Latour, K., Kärki, T., Ricchizzi, E., Kinross, P., Moro, M. L., et al. (2018). Prevalence of Healthcare-Associated Infections, Estimated Incidence and Composite Antimicrobial Resistance Index in Acute Care Hospitals and Long-Term Care Facilities: Results from Two European Point Prevalence Surveys, 2016 to 2017. *Euro Surveill.* 23 (46), 1800516. doi:10.2807/1560-7917.Es.2018.23.46.1800516

Sukhithasri, V., Nisha, N., Biswas, L., Anil Kumar, V., and Biswas, R. (2013). Innate Immune Recognition of Microbial Cell Wall Components and Microbial Strategies to Evade Such Recognitions. *Microbiol. Res.* 168 (7), 396–406. doi:10.1016/j.micres.2013.02.005

Swift, S., Seal, B., Garrish, J., Oakley, B., Hiett, K., Yeh, H.-Y., et al. (2015). A Thermophilic Phage Endolysin Fusion to a *Clostridium Perfringens*-Specific Cell Wall Binding Domain Creates an Anti-*Clostridium* Antimicrobial with Improved Thermostability. *Viruses* 7 (6), 3019–3034. doi:10.3390/v7062758

Tejada-Simon, M. V., Lee, J. H., Ustunol, Z., and Pestka, J. J. (1999). Ingestion of Yogurt Containing *Lactobacillus Acidophilus* and *Bifidobacterium* to Potentiate Immunoglobulin A Responses to Cholera Toxin in Mice. *J. Dairy Sci.* 82 (4), 649–660. doi:10.3168/jds.S0022-0302(99)75281-1

Thanissery, R., Winston, J. A., and Theriot, C. M. (2017). Inhibition of Spore Germination, Growth, and Toxin Activity of Clinically Relevant *C. difficile* Strains by Gut Microbiota Derived Secondary Bile Acids. *Anaerobe* 45, 86–100. doi:10.1016/j.anaerobe.2017.03.004

Tomm, H. A., Ucciferri, L., and Ross, A. C. (2019). Advances in Microbial Culturing Conditions to Activate Silent Biosynthetic Gene Clusters for Novel Metabolite Production. *J. Indust. Microbiol. Biotechnol.* 46 (9), 1381–1400. doi:10.1007/s10295-019-02198-y

Tullus, K., Aronsson, B., Marcus, S., and Möllby, R. (1989). Intestinal Colonization withClostridium Difficile in Infants up to 18 Months of Age. *Eur. J. Clin. Microbiol. Infect. Dis.* 8 (5), 390–393. doi:10.1007/bf01964052

Van Immerseel, F., Rood, J. I., Moore, R. J., and Titball, R. W. (2009). Rethinking Our Understanding of the Pathogenesis of Necrotic Enteritis in Chickens. *Trends Microbiol.* 17 (1), 32–36. doi:10.1016/j.tim.2008.09.005

Vardakas, K. Z., Trigkidis, K. K., Boukouvala, E., and Falagas, M. E. (2016). *Clostridium difficile* Infection Following Systemic Antibiotic Administration in Randomised Controlled Trials: a Systematic Review and Meta-Analysis. *Int. J. Antimicrob. Agents* 48 (1), 1–10. doi:10.1016/j.ijantimicag.2016.03.008

Vasilescu, I.-M., Chifiriuc, M.-C., Pircalabioru, G. G., Filip, R., Bolocan, A., Lazăr, V., et al. (2021). Gut Dysbiosis and *Clostridioides Difficile* Infection in Neonates and Adults. *Front. Microbiol.* 12, 651081. doi:10.3389/fmicb.2021.651081

Vedantam, G., Kochanowsky, J., Lindsey, J., Mallozzi, M., Roxas, J. L., Adamson, C., et al. (2018). An Engineered Synthetic Biologic Protects against *Clostridium difficile* Infection. *Front. Microbiol.* 9, 2080. doi:10.3389/fmicb.2018.02080

Ventola, C. L. (2015). The Antibiotic Resistance Crisis: Part 1: Causes and Threats. *P. T* 40 (4), 277–283.

Vezina, B., Allnutt, T., Keyburn, A. L., Wade, B., Van, T. T. H., Johanesen, P., et al. (2021). Stable Recombinant-Gene Expression from a Ligilactobacillus Live Bacterial Vector via Chromosomal Integration. *Appl. Environ. Microbiol.* 87 (11), e00392–21. doi:10.1128/aem.00392-21

Waksman, S. A., Schatz, A., and Reynolds, D. M. (2010). Production of Antibiotic Substances by Actinomycetes*†. *Ann. N. Y. Acad. Sci.* 1213, 112–124. doi:10.1111/j.1749-6632.2010.05861.x

Waksman, S. A., and Lechevalier, H. A. (1949). Neomycin, a New Antibiotic Active against Streptomycin-Resistant Bacteria, Including Tuberculosis Organisms. *Science* 109 (2830), 305–307. doi:10.1126/science.109.2830.305

Wan, M. L. Y., Forsythe, S. J., and El-Nezami, H. (2019). Probiotics Interaction with Foodborne Pathogens: a Potential Alternative to Antibiotics and Future Challenges. *Crit. Rev. Food Sci. Nutr.* 59 (20), 3320–3333. doi:10.1080/10408398.2018.1490885

Wang, T., Sun, H., Zhang, J., Liu, Q., Wang, L., Chen, P., et al. (2014). The Establishment of *Saccharomyces Boulardii* Surface Display System Using a Single Expression Vector. *Fung. Genet. Biol.* 64, 1–10. doi:10.1016/j.fgb.2013.11.006

Wang, C., Chang, T., Yang, H., and Cui, M. (2015). Antibacterial Mechanism of Lactic Acid on Physiological and Morphological Properties of *Salmonella Enteritidis, Escherichia coli* and *Listeria Monocytogenes*. *Food control.* 47, 231–236. doi:10.1016/j.foodcont.2014.06.034

Wang, Y., Wang, S., Bouillaut, L., Li, C., Duan, Z., Zhang, K., et al. (2018). Oral Immunization with Nontoxigenic *Clostridium difficile* Strains Expressing Chimeric Fragments of TcdA and TcdB Elicits Protective Immunity against *C. difficile* Infection in Both Mice and Hamsters. *Infect. Immun.* 86 (11), e00489–18. doi:10.1128/IAI.00489-18

Wang, B., Hussain, A., Zhou, Y., Zeng, Z., Wang, Q., Zou, P., et al. (2020). *Saccharomyces Boulardii* Attenuates Inflammatory Response Induced by *Clostridium perfringens* via TLR4/TLR15-MyD8 Pathway in HD11 Avian Macrophages. *Poult. Sci.* 99 (11), 5356–5365. doi:10.1016/j.psj.2020.07.045

Ward, S. J., Douce, G., Figueiredo, D., Dougan, G., and Wren, B. W. (1999). Immunogenicity of a *Salmonella typhimurium aroA aroD* Vaccine Expressing a Nontoxic Domain of *Clostridium difficile* Toxin A. *Infect. Immun.* 67 (5), 2145–2152. doi:10.1128/iai.67.5.2145-2152.1999

Whittle, M. J., Bilverstone, T. W., van Esveld, R. J., Lücke, A. C., Lister, M. M., Kuehne, S. A., et al. (2022). A Novel Bacteriophage with Broad Host Range against *Clostridioides Difficile* Ribotype 078 Supports SlpA as the Likely Phage Receptor. *Microbiol. Spectr.* 10 (1), e0229521. doi:10.1128/spectrum.02295-21

Wilcox, M. H., Gerding, D. N., Poxton, I. R., Kelly, C., Nathan, R., Birch, T., et al. (2017). Bezlotoxumab for Prevention of RecurrentClostridium difficileInfection. *N. Engl. J. Med.* 376 (4), 305–317. doi:10.1056/NEJMoa1602615

Wilde, S., Jiang, Y., Tafoya, A. M., Horsman, J., Yousif, M., Vazquez, L. A., et al. (2019). *Salmonella*-vectored Vaccine Delivering Three *Clostridium perfringens* Antigens Protects Poultry against Necrotic Enteritis. *PLoS One* 14 (2), e0197721. doi:10.1371/journal.pone.0197721

World Health Organization (2018). Cholera Vaccine: WHO Position Paper, August 2017 - Recommendations. *Vaccine* 36 (24), 3418–3420. doi:10.1016/j.vaccine.2017.09.034

Wu, C., Shang, Z., Lemetre, C., Ternei, M. A., and Brady, S. F. (2019). Cadasides, Calcium-dependent Acidic Lipopeptides from the Soil Metagenome that Are Active against Multidrug-Resistant Bacteria. *J. Am. Chem. Soc.* 141 (9), 3910–3919. doi:10.1021/jacs.8b12087

Yamanaka, K., Reynolds, K. A., Kersten, R. D., Ryan, K. S., Gonzalez, D. J., Nizet, V., et al. (2014). Direct Cloning and Refactoring of a Silent Lipopeptide Biosynthetic Gene Cluster Yields the Antibiotic Taromycin A. *Proc. Natl. Acad. Sci. U.S.A.* 111 (5), 1957–1962. doi:10.1073/pnas.1319584111

Yang, J., and Yang, H. (2021). Evaluation of the Therapeutic Effect and Dose-Effect of *Bifidobacterium Breve* on the Primary *Clostridioides Difficile* Infected Mice. *Appl. Microbiol. Biotechnol.* 105 (24), 9243–9260. doi:10.1007/s00253-021-11668-2

Yang, X.-q., Zhao, Y.-g., Chen, X.-q., Jiang, B., and Sun, D.-y. (2013). The Protective Effect of Recombinant *Lactococcus Lactis* Oral Vaccine on a *Clostridium Difficile*-Infected Animal Model. *BMC Gastroenterol.* 13, 117. doi:10.1186/1471-230x-13-117

Yang, J., Meng, L., and Yang, H. (2022). Therapeutic Effects of *Bifidobacterium Breve YH68* in Combination with Vancomycin and Metronidazole in a Primary *Clostridioides Difficile*-Infected Mouse Model. *Microbiol. Spectr.* 10, e0067222. doi:10.1128/spectrum.00672-22

Zhang, M., Sun, K., Wu, Y., Yang, Y., Tso, P., and Wu, Z. (2017). Interactions between Intestinal Microbiota and Host Immune Response in Inflammatory Bowel Disease. *Front. Immunol.* 8, 942. doi:10.3389/fimmu.2017.00942

Zhang, X., Hindraand Elliot, M. A. (2019). Unlocking the Trove of Metabolic Treasures: Activating Silent Biosynthetic Gene Clusters in Bacteria and Fungi. *Curr. Opin. Microbiol.* 51, 9–15. doi:10.1016/j.mib.2019.03.003

Zhang, S., Liu, P., Wang, Y., Shen, Z., and Wang, S. (2021). Multiresistance Gene Cfr(C) in *Clostridium perfringens* of Cattle Origin from China. *J. Antimicrob. Chemother.* 76 (12), 3310–3312. doi:10.1093/jac/dkab339

Zhao, L., Guo, Z., Liu, J., Wang, Z., Wang, R., Li, Y., et al. (2017). Recombinant Lactobacillus Casei Expressing *Clostridium perfringens* Toxoids α, β2, ε and β1 Gives Protection against *Clostridium perfringens* in Rabbits. *Vaccine* 35 (32), 4010–4021. doi:10.1016/j.vaccine.2017.05.076

Zhu, J., and Mekalanos, J. J. (2003). Quorum sensing-dependent Biofilms Enhance Colonization in *Vibrio cholerae*. *Dev. Cell.* 5 (4), 647–656. doi:10.1016/s1534-5807(03)00295-8

Zuckerman, J. N., Rombo, L., and Fisch, A. (2007). The True Burden and Risk of Cholera: Implications for Prevention and Control. *Lancet Infect. Dis.* 7 (8), 521–530. doi:10.1016/S1473-3099(07)70138-X

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 03 June 2022
doi: 10.3389/fbioe.2022.837944

Construction of an Antibiotic-Free Vector and its Application in the Metabolic Engineering of *Escherichia Coli* for Polyhydroxybutyrate Production

Ying-Cheng Liao[1], *Boonyawee Saengsawang*[1], *Jun-Wei Chen*[1], *Xiao-Zhen Zhuo*[1] and *Si-Yu Li*[1,2]*

[1]Department of Chemical Engineering, National Chung Hsing University, Taichung, Taiwan, [2]Innovation and Development Center of Sustainable Agriculture, National Chung Hsing University, Taichung, Taiwan

An antibiotic- and inducer-free culture condition was proposed for polyhydroxybutyrate (PHB) production in recombinant *Escherichia coli*. First, antibiotic-free vectors were constructed by installing the plasmid maintenance system, *alp7, hok/sok*, and the *hok/sok* and *alp7* combination into the pUC19 vector. The plasmid stability test showed that pVEC02, the pUC19 vector containing the *hok/sok* system, was the most effective in achieving antibiotic-free cultivation in the *E. coli* B strain but not in the K strain. Second, the putative *phaCAB* operon derived from *Caldimonas manganoxidans* was inserted into pVEC02 to yield pPHB01 for PHB production in *E. coli* BL21 (DE3). The putative *phaCAB* operon was first shown function properly for PHB production and thus, inducer-free conditions were achieved. However, the maintenance of pPHB01 in *E. coli* requires antibiotics supplementation. Finally, an efficient *E. coli* ρ factor-independent terminator, thrLABC (ECK120033737), was inserted between the *phaCAB* operon and the *hok/sok* system to avoid possible transcriptional carry-over. The newly constructed plasmid pPHB01-1 facilitates an antibiotic- and inducer-free culture condition and induces the production of PHB with a concentration of 3.0 on0.2 g/L, yield of 0.26 /L0.07 g/g-glucose, and content of 44 /g3%. The PHB production using *E. coli* BL21 (DE3)/pPHB01-1 has been shown to last 84 and 96 h in the liquid and solid cultures.

Keywords: antibiotic-free vector, hok/sok, Alp7A, inducer-free fermentation, polyhydroxybutyrate

OPEN ACCESS

Edited by:
Shota Atsumi,
University of California, Davis, United States

Reviewed by:
Jian Wang,
University of Georgia, United States
Min-Kyu Oh,
Korea University, South Korea

***Correspondence:**
Si-Yu Li
syli@dragon.nchu.edu.tw

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *17 December 2021*
Accepted: *11 March 2022*
Published: *03 June 2022*

Citation:
Liao Y-C, Saengsawang B, Chen J-W, Zhuo X-Z and Li S-Y (2022) Construction of an Antibiotic-Free Vector and its Application in the Metabolic Engineering of Escherichia Coli for Polyhydroxybutyrate Production. Front. Bioeng. Biotechnol. 10:837944. doi: 10.3389/fbioe.2022.837944

INTRODUCTION

Plasmid-based genetic manipulation is a common approach in metabolic engineering or whole-cell bio-catalysis. One advantage of employing plasmid-based genetic manipulation is the permit of *in vitro* genetic manuscription. The maintenance of recombinant plasmids in bacterial cells usually depends on the use of selective conditions, such as antibiotics supplementations. However, the scaling up of antibiotics use is expensive.

Strategies for the maintenance of plasmids during cell growth under antibiotic-free conditions have been extensively studied (Zielenkiewicz and Cegłowski, 2001; Kroll et al., 2010; Akiyama et al., 2011; Laguna et al., 2015; Singha et al., 2017). Among such strategies, post-segregational killing (PSK)

and active partition systems have been proven to increase plasmid stability (Nordström and Austin, 1989; Sayeed et al., 2000). The *hok/sok* (formerly *parB*) locus, following the PSK mechanism, is a toxin-antidote system that simultaneously produces a toxin (Hok) and a short-lived antidote (Sok), such that in the event of plasmid loss, the cell will be killed by the long-lived toxin (Wu and Wood, 1994; Pecota et al., 2003; Friehs, 2004; Danino et al., 2015). In other words, only the cell that harbors the *hok/sok* locus can propagate because of Sok production.

The *alp7* system, found in *Bacillus subtilis*, is an active partition system consisting of two acting proteins (Alp7A and Alp7R; actin-like proteins) and a centromere-like cis-acting site (alp7C). Alp7A protein is an ATPase which interact with the Alp7R protein. The Alp7R protein recognizes and binds to the *alp7C* site, and the formation of the Alp7CR complex with the help of Alp7A pushes plasmids to the cellular poles during plasmid replication, ensuring equal segregation during cell division (Friehs, 2004; Derman et al., 2009; Derman et al., 2012; Million-Weaver et al., 2014).

Wu and Wood showed that *Escherichia coli* BK6 harboring a *hok/sok*-containing system maintained the plasmid for over 300 h without antibiotics (reaching a 90% proportion of antibiotic-resistant cells) (Wu and Wood, 1994). A recombinant plasmid has been shown to be maintained in *E. coli* Nissle 1917 over 72 h with the help of a *hok*/*sok* and *alp7* combination so that the recombinant *E. coli* Nissle 1917 can serve as a programmable probiotic for cancer detection (Danino et al., 2015). A recombinant plasmid harboring the *alp7* system can be maintained in *B. subtilis* over 14 generations without the supplementation of antibiotics (Derman et al., 2009; Derman et al., 2012). A *trpA*-containing plasmid can be maintained in *E. coli* M72 (K-strain) over 100 generations by integrating the *hok/sok* locus in the plasmid backbone (Kim and Kang, 1996).

Poly-3-hydroxybutyrate (PHB) has attracted much attention owing its biodegradability and biocompatibility, particularly because of the increase in global environmental concerns in recent years. The common method for PHB biosynthesis from acetyl-CoA in microorganisms comprises three steps. First, two molecules of acetyl-CoA are condensed by β-ketothiolase, encoded by *phaA* gene, to form acetoacetyl-CoA. Second, acetoacetyl-CoA reductase (encoded by *phaB*) converts acetoacetyl-CoA to 3-hydroxybutyryl-CoA using NADPH. Finally, the enzyme PHA synthase (encoded by *phaC*) polymerizes 3-hydroxybutyryl-CoA monomers to PHB, liberating CoA (Bernd, 2003; Stubbe et al., 2005; Peña et al., 2014; Boontip et al., 2021). In our previous study, we expressed the PHB biosynthesis genes, *phaCAB*, from the thermophilic *Caldimonas manganoxidans* in *E. coli* and achieved PHB production with a concentration of 16.8 ± 0.6 g/L, content of 74%, and yield of 0.28 g/g glucose (Lin et al., 2017). The expression of the three genes was achieved by inserting *phaA* and *phaB* into the commercial vector, pCDF-Duet1, and inserting *phaC* into a vector, T-BAD, which is ampicillin resistant.

In this study, we constructed three pUC19-based vectors, pVEC01, pVEC02, and pVEC03, which contain *alp7, hok/sok*, and the *alp7-hok/sok* combination, respectively. The stability of pVEC01-03 without the ampicillin supplementation was tested in various *E. coli* strains. The most stable vector among the three was chosen for application in PHB production. The putative *phaCAB* operon of *C. manganoxidans* was first amplified from the *C. manganoxidans* genome, where the amplicon (3,919 bp in total) contained the *phaCAB* gene cluster and a 194-bp fragment upstream of *phaC*. In this manner, the expression of *phaCAB* genes will be under the control of the promoter and ribosome binding site (RBS) from *C. manganoxidans*. This is practically important because any possible inducer, such as IPTG, can be avoided for the economics purpose, where IPTG was calculated to cost 10% of total cost of recombinant β-glucosidase production (Ferreira et al., 2018). While the *phaCAB* operon has been shown to be functionalized in *E. coli* in this study, the function of the recombinant plasmid containing the putative *phaCAB* operon in *E. coli* can only be shown with ampicillin supplementation in the first trial. The latter part of the study proposes a solution to take advantage of the *hok/sok* for antibiotic-free PHB production.

MATERIALS AND METHODS

Bacteria Strains and Plasmids

All strains and plasmids used in this study are listed in **Table 1**.

DNA Manipulation and Transformation

Plasmids were constructed by sequence- and ligation-independent cloning (Li and Elledge, 2007; Jeong et al., 2012; Islam et al., 2017) and schematic diagrams of plasmid construction are shown in **Figure 1** The primers used in this study are listed in **Table 2**. To construct the recombinant plasmid pVEC01, *alp7* was amplified from pTKW106alp7 (Addgene, Watertown, United States) by polymerase chain reaction (PCR), and pUC19 was digested with *Xba*I and *Bam*HI. To construct the recombinant plasmids pVEC02 and pVEC03, the *hok/sok* locus was amplified from pTKW106alp7 by PCR, and the vectors pUC19 and pVEC01 were digested with *Sal*I. To construct the recombinant plasmid pPHB01, the *phaCAB* operon was amplified from the genomic DNA of *C. manganoxidans* and pVEC02 was digested with *Sph*I. After the preparation of inserts and vectors, the DNA fragments were purified using the Gene-Spin™ 1-4-3 DNA Extraction Kit (PROTECH, Taipei, Taiwan). The insert and corresponding vector were mixed at an appropriate ratio and incubated at 37°C for 1 min with T4 DNA polymerase to generate 5′-overhangs. Thereafter, the reaction mixture was placed on ice for 20 min for single-strand annealing and was introduced into competent *E. coli* DH5α cells for transformation.

To construct the recombinant plasmid pPHB01-1, the Q5® site-directed mutagenesis kit (NEB, Ipswich, United States) was used to insert a 57-bp *E. coli* ρ factor-independent terminator, thrLABC (ECK120033737) between the *phaCAB* operon and the *hok/sok* locus in pPHB01. The primers used for constructing pPHB01-1 were In/del-F- pPHB01-1 and In/del-R-pPHB01-1 (**Table 2**). After PCR, the amplified DNA was treated with a Kinase-Ligase-DpnI enzyme mix at room temperature for 5 min for circularization and template removal. Mutant selection was

TABLE 1 | List of bacterial strains and plasmids used in this study.

Name	Descriptions	References
Bacterial strains		
E. coli DH5α	F⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17 (rK^-mK^+), λ^-	Lab stock
E. coli BL21 (DE3)	F⁻, *dcm*, *ompT*, *gal*, *lon*, $hsdS_B$(rB^-, mB^-), λ(DE3 [*lac*I, *lac*UV5-T7 gene 1, *ind*1, *sam7*,*nin5*])	Lab stock
E. coli MG1655	F⁻ lambda⁻ *ilvG- rfb*-50 *rph*-1	Lab stock
E. coli MZLF	*E. coli* BL21 (DE3) Δ*zwf*, Δ*ldhA*, Δ*frd*	Yang et al. (2016)
C. manganoxidans	JCM 10698 (BCRC 17858)	Lab stock
Plasmids		
pUC19	Commonly used cloning vector that conveys Tet and Amp resistance.	Lab stock
pTKW106alp7A	Recombinant plasmid carries *alp7CAR* gene and *hok/sok* gene	Danino et al. (2015)
pVEC01	Recombinant plasmid carries *alp7CAR* gene (derived from *B. subtilis* pLS20) to maintain the plasmid stability.	This study
pVEC02	Recombinant plasmid carries *hok/sok* gene (derived from *E. coli* R1 plasmid) to maintain the plasmid stability.	This study
pVEC03	Recombinant plasmid carries *alp7CAR* gene and *hok/sok* gene to maintain the plasmid stability.	This study
pPHB01	pVEC02-based plasmid carries *hok/sok* gene and *phaCAB* operon (derived from *C. manganoxidans*) under the control of the native promoter of *phaCAB* operon from *C. manganoxidans*.	This study
pPHB01-1	pPHB01 derived recombinant plasmid where a terminator was inserted between *hok/sok* gene and *phaCAB* operon	This study

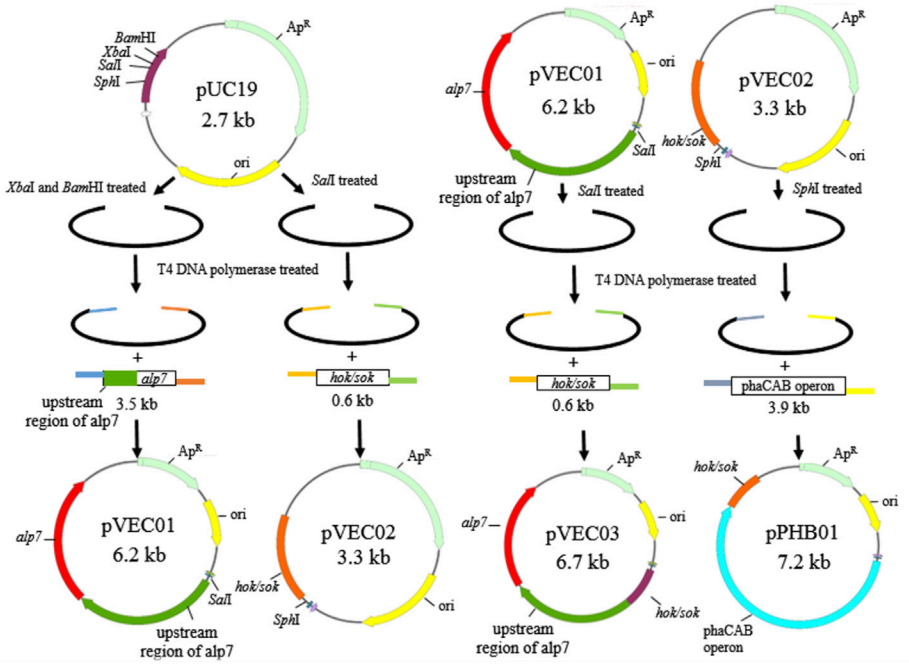


FIGURE 1 | Schematic of the construction of pVEC01, pVEC02, pVEC03, and pPHB01.

enriched by simultaneous ligation and *Dpn*I treatment. Next, 5 μL of the KLD mix was directly introduced into competent *E. coli* DH5α cells. Colony PCR was used for pPHB01-1 screening with the primers CK-F-pPHB01-1 (in-del) and CK-R-pPHB01-1 (in-del) (**Table 2**). Finally, pPHB01-1 was confirmed by DNA sequencing. All recombinant plasmids constructed in this study were verified by colony PCR, restriction enzyme digestion, and DNA sequencing.

Plasmid Stability Test

Escherichia coli strains DH5α, MG1655, BL21 (DE3), and MZLF (Yang et al., 2016) were used as hosts to test plasmid stability. Cells for the plasmid stability test were grown from -80°C stocks by streaking on an LB plate with ampicillin. After 16–24 h, a single colony was picked and used to inoculate LB containing antibiotics for 12 h. Then, 10 μL of the cell suspensions was inoculated into 3 ml LB without antibiotics to start the stability experiment. The inoculum was diluted and plated on LB plates supplemented with or without antibiotics.

After 12 h, 10 μL of the cell suspensions was transferred to a new test tube. Every 12 h, 10 μL of the sample was transferred to a new flask, and the cells were diluted and plated on LB plates supplemented with or without antibiotics. Plasmid stability (segregation) was estimated by the equation:

TABLE 2 | Primer sequence in this study.

Primer	Sequence[a]
SLIC-F-alp7-i	TGCCTGCAGGTCGACTCTAGaccacctaggtcattagcct
SLIC-R-alp7-i	GAGCTCGGTACCCGGGGATCtcagggcgtctgtgttgcaa
SLIC-F-parB-01	CTTGCATGCCTGCAGGTCGAaacaaactccgggaggcagc
SLIC-R-parB-01	CCTAGGTGGTCTAGAGTCGAacaacatcagcaaggagaaa
SLIC-F-phaCAB1-01	CCATGATTACGCCAAGCTTGggtacatggagcagatgcgc
SLIC-R-phaCAB1-01	AGTTTGTTTCGACCTGCAGGcgagttgatcgccaacgaag
gko-F-CAB200-01	ATGAAGCGAACCACCTCC
gko-R-CAB200-01	CAAGCTTGGCGTAATCATG
In/del-F- pPHB01-1	CCCGCACCTGACAGTGCGGGCTTTTTTTTTCGACCAAAGGagtgccgctcttcgttggcgatcaactcg
In/del-R- pPHB01-1	CTTTTTTctgtgtttccATTGTTAGACGAGAGTGTGCTCAGTTGTCAAGccgtccgccgggacggccgcatgtcctggcttac
CK-F-pPHB01-1 (in-del)	GAAGAAATCGCTTCGATCG
CK-R-pPHB01-1 (in-del)	AAACCACCTTCACGTCATG

[a]*Uppercase letters indicate regions of homology with vector; lowercase letters indicate the primer for insert DNA.*

$$\text{Ampicillin} - \text{resistant cells } (\%) = \frac{\text{colonies on } LB + \text{ampicillin}}{\text{colonies on LB}} \times 100\%.$$

Characterization of the Putative Promoter of the *phCAB* Operon From *C. Manganoxidans*

The putative promoter of the *phaCAB* operon was predicted by BPROM (http://www.softberry.com/berry.phtml) and putative ribosome binding sites (RBSs) were reported in a previous study (Lin et al., 2017).

Culture Conditions for PHB Production

The strains used for shake flask experiments were grown aerobically at 200 rpm and 37°C in fresh 25-ml LB medium supplemented with 20 g/L glucose, 0.6 g/L $MgSO_4$ $7H_2O$, 0.07 g/L $CaCl_2$ $2H_2O$, 0.04 g/L KH_2PO_4, 0.04 g/L K_2HPO_4, 0.4 g/L $NaHCO_3$ and 100 μg/ml ampicillin. In the shake flask experiments, the initial OD_{600} was adjusted to 0.05. Samples were taken and stored every 12 h until 48 h for the subsequent analysis.

Analytical Methods

Cell density was monitored by measuring the optical density at 600 nm (GENESYS 10S, Thermo Scientific, United States). For the measurement of biomass concentration, 3 ml of the cultured cells were centrifuged for 1 min at 6,791 rcf. The cells were frozen at −20°C and lyophilized. After freeze-drying, the weight of the sample was recorded to calculate the biomass concentration. The PHB produced by the recombinant *E. coli* cells was quantified by gas chromatography (GC), and the samples were treated as described before (Hsiao et al., 2016). Briefly, an appropriate amount of biomass was transferred to a clean spiral test tube and 1 ml of chloroform, 0.85 ml of methanol, and 0.15 ml of sulfuric acid were added. The tube was incubated in a water bath for 140 min at 80°C and cooled down to room temperature. Then 1 ml of DI H_2O was mixed well by vortexing. After standing and layering, the organic phase was removed and filtered through a 0.2-μm PVDF filter and then analyzed by GC. The temperatures of the injector and detector were 230 and 275°C, respectively. The temperature of the column was set at 100°C and increased to 200°C at a rate of 10°C/min and maintained at 200°C for 2 min. The nitrogen was used as the carrier gas at 3 ml/min. The split mode with the split ratio of 1:10 was used (Chen S.-K. et al., 2013).

The metabolites were determined using a Thermo Scientific™ Dionex™ Ulitmate 3000 LC system equipped with an ORH-801 column. Sulfuric acid (5 mM) was used as the mobile phase, and the flow rate was maintained at 0.6 ml/min. Samples for quantification were collected from the culture media after the removal of the cells by centrifugation for 1 min at 6,791 rcf. The supernatant was passed through a 0.2-μm PVDF filter, and the sample injection (10 μL) was performed using an auto-sampler.

Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis

The bacterial culture solutions were centrifuged at 4°C and 9,072 × g for 10 min and resuspended in lysis buffer (50 mM Tris, 0.1 M NaCl, 1 mM EDTA, pH 7.8) to obtain a final OD of 20. The concentrated bacterial solutions were subjected to ultrasonication (10 s break and 5 s intermittent, total 60 cycle) (QSonica Q700, United States). The cell lysates were then diluted with 4x Laemmli sample buffer (Bio-Rad Laboratories, Inc., United States) and heated in a dry batch for 10 min. Samples (10 μL) were loaded into the wells for SDS-PAGE analysis (12% acrylamide).

Nile Red Stain Assay

To confirm *in vivo* PHB production, the recombinant *E. coli* BL21 (DE3)/pPHB01 was grown on LB plates containing 10 g/L glucose and 0.8 mg/ml nile red. After 24 h of cultivation at 37°C, the plates were exposed to monochromatic light at 470 nm for observation.

RESULTS

Stability of pVEC01, pVEC02, and pVEC03 in *E. coli*

The recombinant plasmids pVEC01, pVEC02, and pVEC03 were tested for plasmid stability in different *E. coli* hosts.

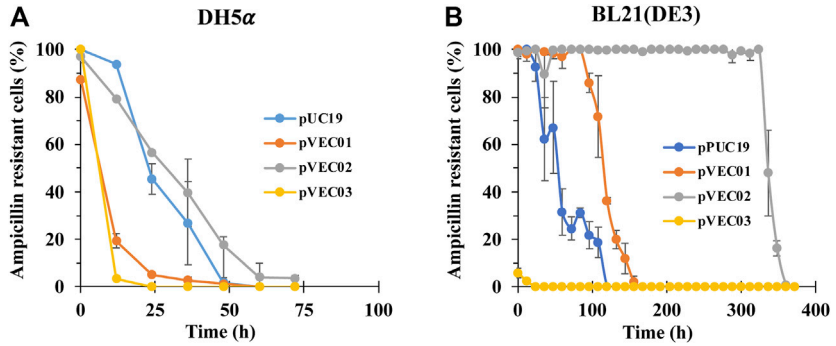


FIGURE 2 | The plasmid stability test of pUC19, pVEC01, pVEC02, and pVEC03 in *E. coli* strains **(A)** DH5α and **(B)** BL21 (DE3). Errors represent standard deviation with $n = 3$.

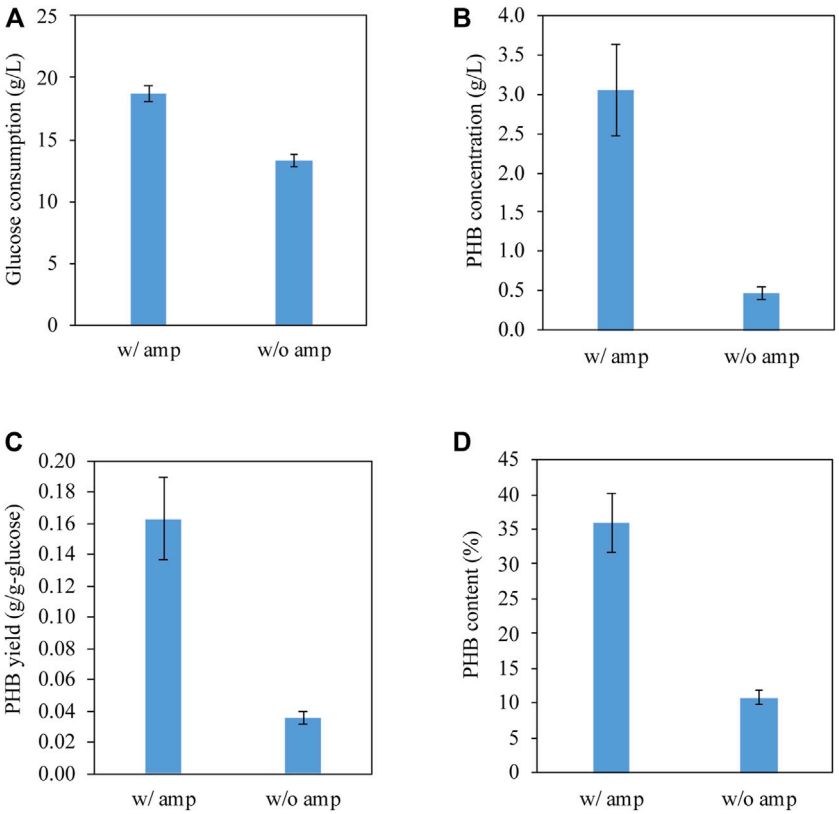


FIGURE 3 | (A) Glucose consumption, **(B)** PHB concentration, **(C)** PHB yield, and **(D)** PHB content of *E. coli* BL21 (DE3)/pPHB01 with and without ampicillin supplementation. The initial glucose concentration for conditions of with and without amp were 19.6 BL0.4 and 18.7 .40.5 g/L, respectively. Errors represent standard deviation with $n = 3$.

Compared to the parental plasmid pUC19, pVEC02 showed good stability in *E. coli* DH5α *E. coli* BL21 (DE3) and with time required to reach 50% ampicillin-resistant cells of 26 and 336 h, respectively (while pUC19 had times of 25 and 54 h, respectively, **Figures 2A,B**). The *hok/sok* system increased the segregational stability in BL21 (DE3) by 6 fold, but not in *E. coli* K strain. In fact, pVEC01, pVEC02, and pVEC03 were not stable in *E. coli* K strains DH5α. pVEC01 was slightly more stable than pUC19 in BL21 (DE3) cells and pVEC03 showed instability. Because of the stability of pVEC02 in this test, it was chosen as the new vector for the subsequent induction of PHB production.

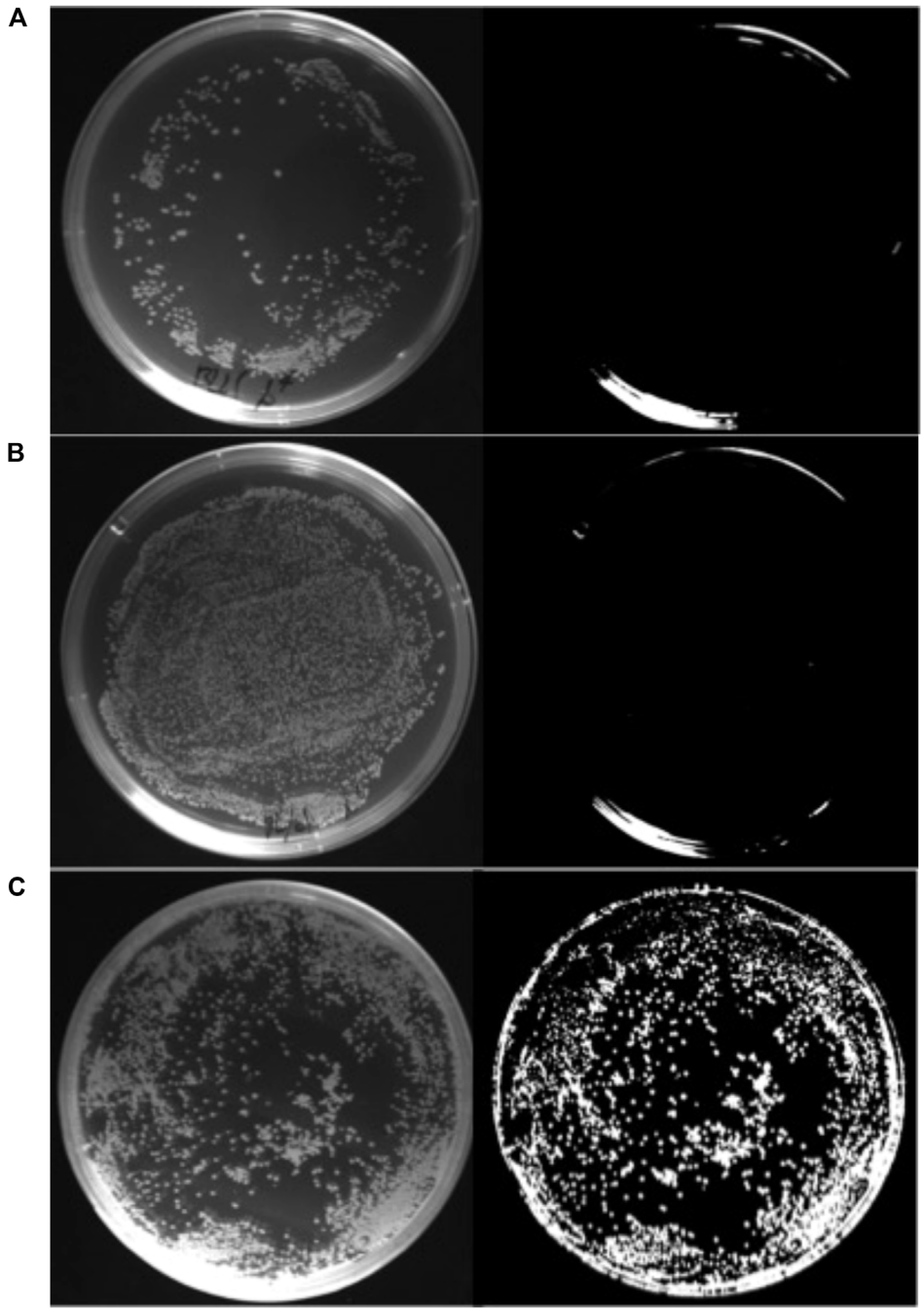


FIGURE 4 | Nile red stain assay of **(A)** *E. coli* BL21 (DE3), **(B)** *E. coli* BL21 (DE3)/pPHB01 **(C)** *E. coli* BL21 (DE3)/pPHB01-1. The LB agar plate contained 10 g/L glucose and 0.8 mg/ml nile red without ampicillin. The plate on the left and right in each panel was taken under white light and monochromatic light at 470 nm, respectively.

PHB Production by *E. Coli* BL21(DE3)/ pPHB01 in Shake Flasks

We tested PHB production by conducting shake flask experiments at 37°C under the control of the native promoter of *phaCAB* operon from *C. manganoxidans*. **Figure 3A** shows that *E. coli* BL21 (DE3)/pPHB01 consumed 18 DE1 and 13 a1 g/L glucose with and without ampicillin, respectively. In the presence of ampicillin, *E. coli* BL21 (DE3) harboring pPHB01 was able to effectively produce PHB and achieved a PHB concentration of 3.1 DE0.6 g/L (**Figure 3B**), yield of 0.16 , 0.03 g/g-glucose, and content of 36 /g4% (**Figure 3D**). In contrast, in the absence of ampicillin, *E. coli* BL21 (DE3) harboring pPHB01 produced PHB

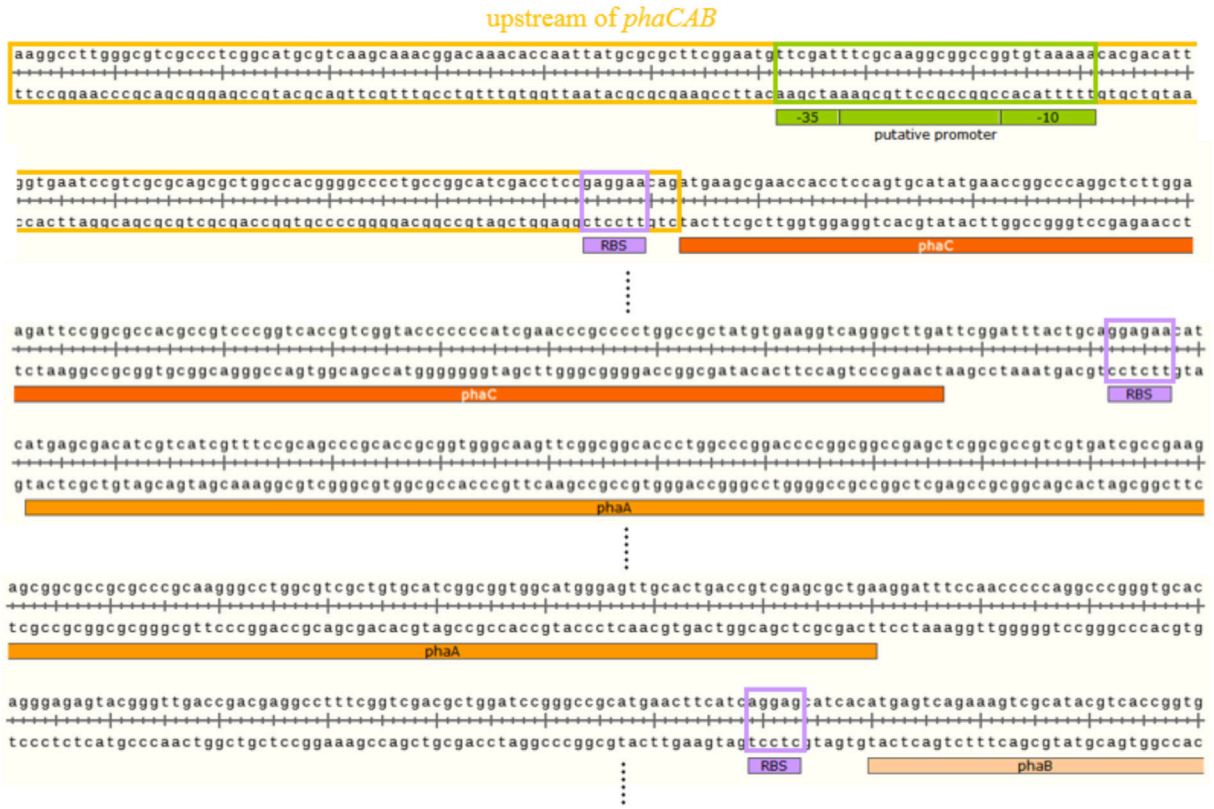


FIGURE 5 | The putative promoter of *phaCAB* operon and the putative ribosome binding sites (RBSs) of *phaCAB* genes.

with a decreased concentration of 0.5 w0.1 g/L, as shown in **Figures 3B–D**. The effective PHB production in *E. coli* BL21 (DE3)/pPHB01 justifies the high glucose consumption in the presence of ampicillin.

The capability of *E. coli* BL21 (DE3)/pPHB01-1 to produce PHB without the inducer and antibiotics was confirmed by the Nile red assay, as shown in **Figure 4**.

Characterization of the *phaCAB* Operon From *C. manganoxidans*

The putative promoter of the *phaCAB* operon was predicted by BPROM, and the sequences of the putative promoter, -35 box and -10 box sequences were TTCGATTTCGCAAGGCGGCCGGTG TAAAAA, TTCGAT, and GTGTAAAAA, respectively (**Figure 5**). Furthermore, RBSs of the *phaCAB* operon were labeled according to a previous study (**Figure 5**) (Lin et al., 2017).

Isolation of the *phaCAB* Operon and *Hok/Sok* Genetic Units by Inserting a Terminator

While pVEC02 (*hok/sok*) was shown to be a stable vector in *E. coli* without ampicillin supplementation, the low PHB production by *E. coli* BL21 (DE3)/pPHB01 in the absence of ampicillin indicates the instability of pPHB01 in *E. coli* BL21 (DE3). It is speculated that the insertion of the *phaCAB* operon in pVEC02 disrupted the balance of hok/sok transcripts. In other words, the promoter of the *phaCAB* operon may carry an additional transcript of *hok* and make pPHB01 toxic to *E. coli* (**Figure 6**). To verify this speculation, an efficient *E. coli* ρ factor-independent terminator, thrLABC (ECK120033737) (Chen Y.-J. et al., 2013), was inserted between the *phaCAB* operon and the *hok/sok* locus to isolate the two genetic units (**Figure 6**). The newly constructed plasmid was named pPHB01-1.

E. coli BL21 (DE3) harboring the newly constructed plasmid pPHB01-1 in was tested for the PHB production with and without ampicillin supplementation (**Figure 7**). In the presence of ampicillin, *E. coli* BL21 (DE3)/pPHB01-1 was able to produce PHB effectively. All the fed glucose of 13 DE3 g/L was completely consumed in 48 h (**Figure 7A**), and the PHB concentration, yield, and content were 3.1 , 0.3 g/L, 0.26 /L0.08 g/g-glucose, and 45 /g4%, respectively (**Figures 7B–D**). In the absence of ampicillin supplementation, *E. coli* BL21 (DE3)/pPHB01-1 was still able to provide a comparable PHB performance, where the PHB concentration, yield, and content was 3.0 PH0.2 g/L, 0.26 /L0.07 g/g-glucose, and 44 /g3%, respectively (**Figures 7B–D**).

To characterize the long-term stability of PHB production in *E. coli* BL21 (DE3)/pPHB01-1, *E. coli* BL21 (DE3)/pPHB01-1 was first cultivated in the liquid culture without the ampicillin (the medium composition was described in 2.5) and transferred to a fresh medium every 12 h. During the transfer, the grown-culture solution

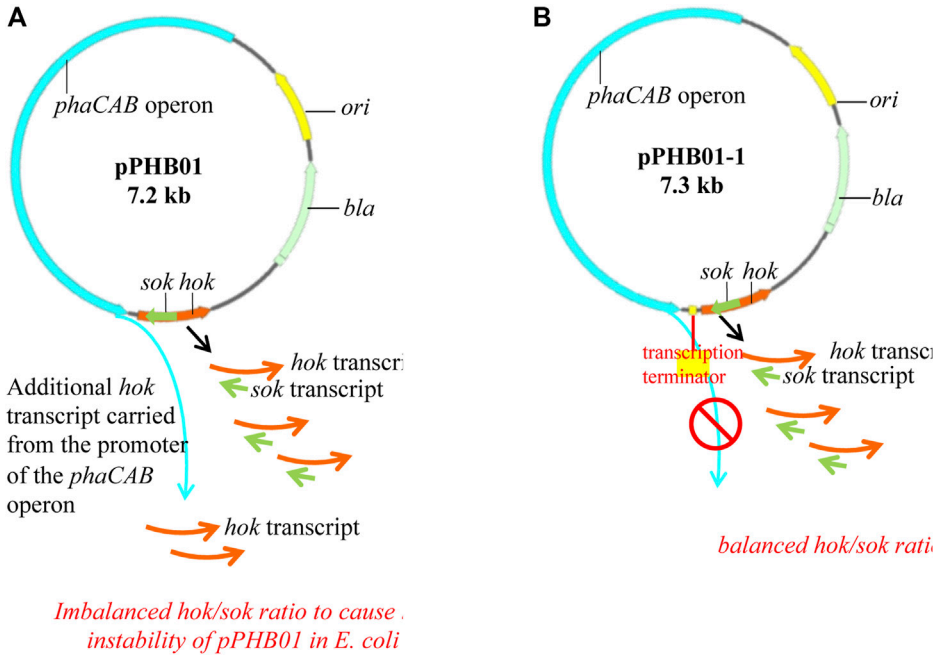


FIGURE 6 | Maps of **(A)** pPHB01 and **(B)** pPHB01-1, where pPHB01-1 was constructed by inserting a 57-bp *E. coli* ρ factor-independent terminator (thrLABC, ECK120033737) into the pPHB01 to isolate the *phaCAB* operon and the *hok/sok* system.

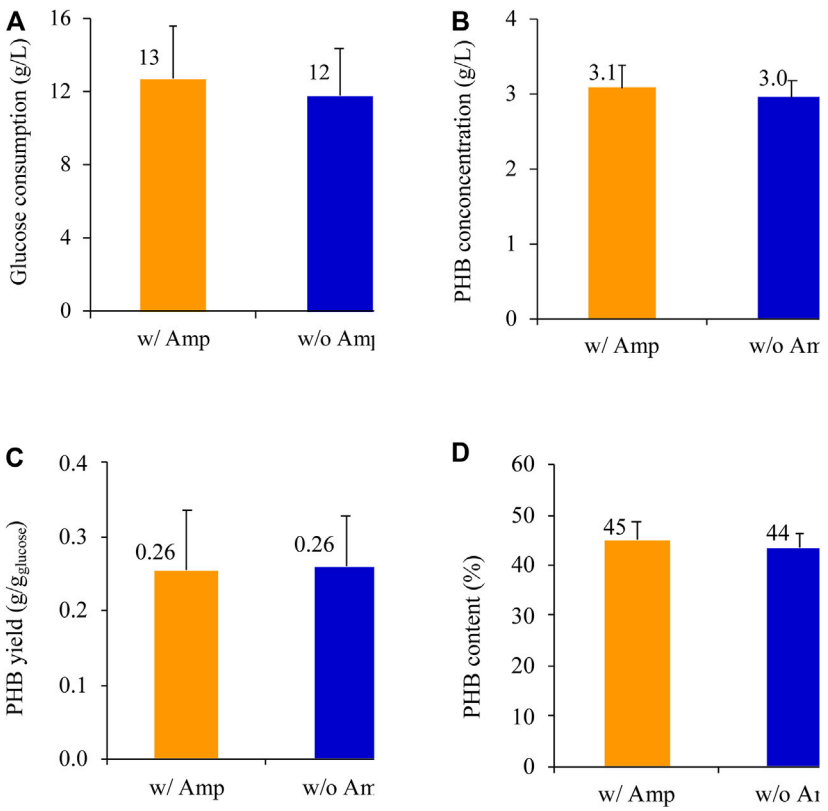


FIGURE 7 | (A) Glucose consumption, **(B)** PHB conconcentration, **(C)** PHB yield, and **(D)** PHB content of *E. coli* BL21 (DE3)/pPHB01-1 with out and with 100 μg/ml ampicillin. The initial glucose concentration for conditions of with and without amp were 13 am3 and 12 a3 g/L, respectively. Errors represent standard deviation with $n = 3$.

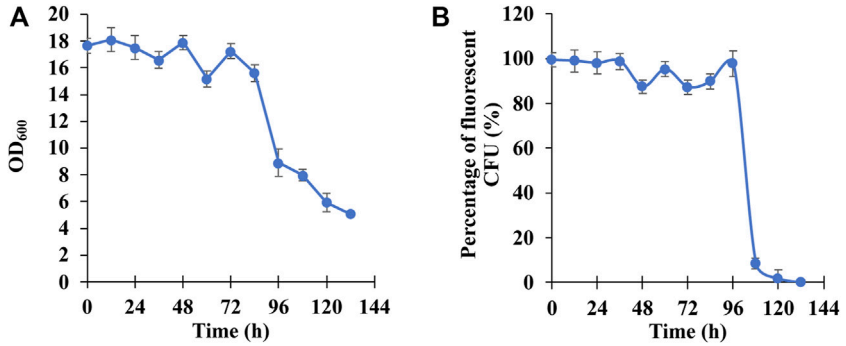


FIGURE 8 | (A) OD_{600} and **(B)** percentage of fluorescent colony forming unit (CFU) of *E. coli* BL21 (DE3)/pPHB01-1 for the long-term PHB production. Errors represent standard deviation with $n = 3$.

was spread on the agar plate without the ampicillin but 0.8 mg/ml nile red. After the formation of CFU, the total CFU as well as the fluorescent CFU were counted to quantify the percentage of CFU that can produce PHB. It can be seen in **Figure 8A** that the OD_{600}, which is positively correlated to the PHB production, can be maintained in the range of 15 and 18 in the first 84 h and gradually dropped to 5.1 a0.1 at 132 h. **Figure 8B** shows that the *E. coli* BL21 (DE3)/pPHB01-1 maintained its ability to produce PHB in the first 96 h and less than 20% of cell population can produce PHB afterwards.

SDS-PAGE of *E. Coli* BL21(DE3)/pPHB01 and *E. Coli* BL21(DE3)/pPHB01-1

SDS-PAGE was used to confirm the expression of the *phaCAB* operon derived from *C. manganoxidans* in *E. coli*. It can be clearly seen in **Figure 9** (lanes 7–12) that three overexpressed signals were found in both *E. coli* BL21 (DE3)/pPHB01 and *E. coli* BL21 (DE3)/pPHB01-1 samples, which correspond to PhaC (61 kDa), PhaA (43 kDa), and PhaB (27 kDa). PhaA and PhaB in *E. coli* BL21 (DE3)/pPHB01 and *E. coli* BL21 (DE3)/pPHB01-1 were mainly found in the soluble fractions (lanes 8 and 11 of **Figure 9**), while PhaC was mainly found in the insoluble fraction (lanes 9 and 12 of **Figure 9**).

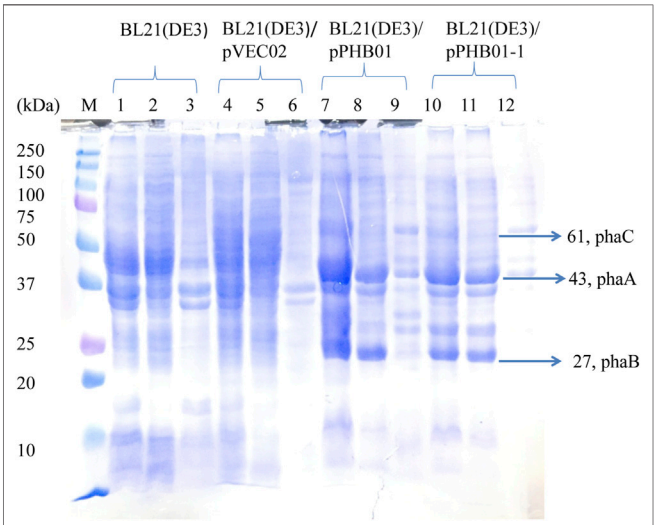


FIGURE 9 | SDS-PAGE of *E. coli* BL21 (DE3)/pPHB01 and BL21 (DE3)/pPHB01-1. M: marker; lane 1- lane 3: BL21 (DE3); lane 4 - lane 6: BL21 (DE3)/pVEC02; lane 7- lane 9: BL21 (DE3)/pPHB01; lane 10 - lane 12: BL21 (DE3)/pPHB0-1. Lane 1, 4, 7, 10: whole cells; lane 2, 5, 8, 11: soluble protein; lane 3, 6, 9, 12: insoluble fractions.

DISCUSSION

Synthetic biology has gradually become the keystone of our current technology-based society (Tseng et al., 2018; Voigt, 2020; Pelletier et al., 2021). The management of synthetic biology relies on the characterization of genetic units and their interactions. Plasmid system is a general platform to carry genetic information and therefore plasmid instability becomes a concern to maintain the newly installed traits. In previous studies, *hok/sok* was found to be a potential system to increase the plasmid stability in *E. coli* strains (Gerdes et al., 1986; Schweder et al., 1992; Lee et al., 1994). Multiple copies of the *hok/sok* locus in the target plasmid can further improve the stability of the plasmid in *E. coli* (Gerdes et al., 1986; Lee et al., 1994; Kim and Kang, 1996; Pecota et al., 1997). However, it is common to see the expression of recombinant proteins cause the imbalance of the *hok/sok* system, where transcription and translation are simultaneously involved in the plasmid maintenance (Lee et al., 1994; Wu and Wood, 1994). Therefore, the combination of plasmid maintenance systems as a hybrid one was proposed to increase the plasmid stability and the time to reach 50% plasmid-bearing cells was extended to ~ 72 h (Pecota et al., 2003).

In this study, a high-copy number backbone, pUC19, was chosen since it may benefit recombinant protein production and metabolic engineering. PSK and active partition systems are usually found in a low-copy number plasmid. It will be interesting to see how well PSK and active partition systems maintain the high-copy pUC19 in *E. coli*. This study clearly demonstrated that *alp7* system was incompatibility with the high-copy backbone. This may because the active partition of the

plasmid consumes ATP during DNA replication (Baxter and Funnell, 2015). Energy consumption is critical for the function of *alp7*, and the installation of *alp7* in a high-copy number backbone may consume too much ATP, making *alp7*/pUC19 incompatible with *E. coli*. In contrast, the genetic unit *hok/sok* in pVEC02 was found to be an efficient antibiotic-free selection marker in *E. coli* B strains, including *E. coli* BL21 (DE3) and MZLF. The time to reach 50% ampicillin-resistant cells can go up to 360 h, which is competitive among literatures. Perplexedly, this study presented a result that the *hok/sok* system was not compatible with *E. coli* K strains, including MG1655 and DH5α. *E. coli* K and B strains bear 5 and 6 *hok/sok* loci in the chromosome, respectively (Schneider et al., 2002), and all of them are considered inactive yet may be induced by unknown signal (Pedersen and Gerdes, 1999). More study can be conducted in detail in the future to discriminate the compatibility of the *hok/sok* system among *E. coli* strains.

The application of the antibiotic-free vector pVEC02 was further investigated in metabolic engineering for PHB production by constructing the recombinant pPHB01. *E. coli* BL21 (DE3)/pPHB01 can effectively produce PHB; however, only in the presence of antibiotics. The ineffective function of the *hok/sok* system in pPHB01 demonstrates that the *hok/sok* system in the engineering perspective needs more detailed investigation. We focused on the interaction among genetic units to determine the antibiotic demand of pPHB01. It is speculated that the transcription activity of the *phaCAB* operon passes downstream of the *hok/sok* gene unit. The transcriptional carry-over increases *hok* transcription; therefore, the *hok/sok* balance is interrupted. This carry-over of the transcription activity can be especially severe when the downstream *hok* is a short transcript. To prevent transcriptional carry-over, an efficient *E. coli* ρ factor-independent terminator, thrLABC (ECK120033737) (Chen Y.-J. et al., 2013), was inserted between the *phaCAB* operon and the *hok/sok* gene unit (**Figure 7**). The isolation of the two genetic units was shown in this study to be effective in finding an application for pVEC02 in metabolic engineering. In addition, a putative promoter region of the *phaCAB* operon from *C. manganoxidans* was first shown properly while effectively function in *E. coli*. To our knowledge, this is the first study to demonstrate that the *phaCAB* operon from *C. manganoxidans* can be used for effective PHB production in *E. coli*. A previous study showed that the expression level of PhaC was important for the phaCAB-mediated pathway (Hiroe et al., 2012). As shown in **Figure 9** that PhaC expression level was not strong, the PhaC expression level of the *phaCAB* operon in *E. coli* can be further optimized in the future. In summary, the inducer-free PHB production can be achieved by adopting heterologous promoter from *C. manganoxidans*. The antibiotic-free PHB production involves the interplay among bacteria chassis, antibiotic-free genetic markers, and the genetic operon for PHB production. This study presents a right combination to achieve the inducer- and antibiotic-free system for PHB production.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Y-CL and BS conducted experiments and wrote the draft of the manuscript. J-WC and X-ZZ conducted experiments. Y-CL, BS, J-WC, X-ZZ, and S-YL analyzed the data. S-YL conceived and supervised the study along with editing the manuscript.

FUNDING

This work was funded by the Ministry of Science and Technology Taiwan MOST-108-2628-E-005-002-MY3, MOST-110-2622-E-005-011-CC2, and the Ministry of Education, Taiwan, R.O.C. under the Higher Education Sprout Project.

REFERENCES

Akiyama, H., Okuhata, H., Onizuka, T., Kanai, S., Hirano, M., Tanaka, S., et al. (2011). Antibiotics-free Stable Polyhydroxyalkanoate (PHA) Production from Carbon Dioxide by Recombinant Cyanobacteria. *Bioresour. Technol.* 102, 11039–11042. doi:10.1016/j.biortech.2011.09.058

Baxter, J. C., and Funnell, B. E. (2015). "Plasmid Partition Mechanisms," in *Plasmids: Biology and Impact in Biotechnology and Discovery* American Society for Microbiology ASM, 135–155. doi:10.1128/9781555818982.ch8

Bernd, H. (2003). Polyester Synthases: Natural Catalysts for Plastics. *Biochem. J.* 376, 15–33.

Boontip, T., Waditee-Sirisattha, R., Honda, K., and Napathorn, S. C. (2021). Strategies for Poly(3-Hydroxybutyrate) Production Using a Cold-Shock Promoter in *Escherichia coli*. *Front. Bioeng. Biotechnol.* 9, 666036. doi:10.3389/fbioe.2021.666036

Chen, S.-K., Chin, W.-C., Tsuge, K., Huang, C.-C., and Li, S.-Y. (2013a). Fermentation Approach for Enhancing 1-butanol Production Using Engineered Butanologenic *Escherichia coli*. *Bioresour. Technol.* 145, 204–209. doi:10.1016/j.biortech.2013.01.115

Chen, Y.-J., Liu, P., Nielsen, A. A. K., Brophy, J. A. N., Clancy, K., Peterson, T., et al. (2013b). Characterization of 582 Natural and Synthetic Terminators and Quantification of Their Design Constraints. *Nat. Methods* 10, 659–664. doi:10.1038/nmeth.2515

Danino, T., Prindle, A., Kwong, G. A., Skalak, M., Li, H., Allen, K., et al. (2015). Programmable Probiotics for Detection of Cancer in Urine. *Sci. Transl Med.* 7, 289ra84. doi:10.1126/scitranslmed.aaa3519

Derman, A. I., Becker, E. C., Truong, B. D., Fujioka, A., Tucey, T. M., Erb, M. L., et al. (2009). Phylogenetic Analysis Identifies many Uncharacterized Actin-like Proteins (Alps) in Bacteria: Regulated Polymerization, Dynamic Instability and Treadmilling in Alp7A. *Mol. Microbiol.* 73, 534–552. doi:10.1111/j.1365-2958.2009.06771.x

Derman, A. I., Nonejuie, P., Michel, B. C., Truong, B. D., Fujioka, A., Erb, M. L., et al. (2012). Alp7R Regulates Expression of the Actin-like Protein Alp7A in Bacillus Subtilis. *J. Bacteriol.* 194, 2715–2724. doi:10.1128/jb.06550-11

Ferreira, R. d. G., Azzoni, A. R., and Freitas, S. (2018). Techno-economic Analysis of the Industrial Production of a Low-Cost Enzyme Using *E. coli*: the Case of Recombinant β-glucosidase. *Biotechnol. Biofuels* 11, 81. doi:10.1186/s13068-018-1077-0

Friehs, K. (2004). "Plasmid Copy Number and Plasmid Stability," in *New Trends and Developments in Biochemical Engineering* (Springer), 47–82. doi:10.1007/b12440

Gerdes, K., Rasmussen, P. B., and Molin, S. (1986). Unique Type of Plasmid Maintenance Function: Postsegregational Killing of Plasmid-free Cells. *Proc. Natl. Acad. Sci. U.S.A.* 83, 3116–3120. doi:10.1073/pnas.83.10.3116

Hiroe, A., Tsuge, K., Nomura, C. T., Itaya, M., and Tsuge, T. (2012). Rearrangement of Gene Order in the phaCAB Operon Leads to Effective Production of Ultrahigh-Molecular-Weight Poly[(R)-3-Hydroxybutyrate] in Genetically Engineered *Escherichia coli*. *Appl. Environ. Microbiol.* 78, 3177–3184. doi:10.1128/aem.07715-11

Hsiao, L.-J., Lin, J.-H., Sankatumvong, P., Wu, T.-M., and Li, S.-Y. (2016). The Feasibility of Thermophilic Caldimonas Manganoxidans as a Platform for Efficient PHB Production. *Appl. Biochem. Biotechnol.* 180, 852–871. doi:10.1007/s12010-016-2138-0

Islam, M. N., Lee, K. W., Yim, H.-S., Lee, S. H., Jung, H. C., Lee, J.-H., et al. (2017). Optimizing T4 DNA Polymerase Conditions Enhances the Efficiency of One-step Sequence- and Ligation-independent Cloning. *BioTechniques* 63, 125–130. doi:10.2144/000114588

Jeong, J.-Y., Yim, H.-S., Ryu, J.-Y., Lee, H. S., Lee, J.-H., Seen, D.-S., et al. (2012). One-Step Sequence- and Ligation-independent Cloning as a Rapid and Versatile Cloning Method for Functional Genomics Studies. *Appl. Environ. Microbiol.* 78, 5440–5443. doi:10.1128/aem.00844-12

Kim, J.-Y., and Kang, H. A. (1996). Comparison of Effects of the Par Locus and the parB Locus on Plasmid Stability and Host Cell Growth. *J. Ferment. Bioeng.* 82, 495–497. doi:10.1016/s0922-338x(97)86990-5

Kroll, J., Klinter, S., Schneider, C., Voss, I., and Steinbüchel, A. (2010). Plasmid Addiction Systems: Perspectives and Applications in Biotechnology. *Microb. Biotechnol.* 3, 634–657. doi:10.1111/j.1751-7915.2010.00170.x

Laguna, R., Young, S. J., Chen, C.-C., Ruiz, N., Yang, S.-T., and Tabita, F. R. (2015). Development of a Plasmid Addicted System that Is Independent of Co-inducers, Antibiotics and Specific Carbon Source Additions for Bioproduct (1-butanol) Synthesis in *Escherichia coli*. *Metab. Eng. Commun.* 2, 6–12. doi:10.1016/j.meteno.2014.12.001

Lee, S. Y., Yim, K. S., Chang, H. N., and Chang, Y. K. (1994). Construction of Plasmids, Estimation of Plasmid Stability, and Use of Stable Plasmids for the Production of Poly(3-Hydroxybutyric Acid) by Recombinant *Escherichia coli*. *J. Biotechnol.* 32, 203–211. doi:10.1016/0168-1656(94)90183-x

Li, M. Z., and Elledge, S. J. (2007). Harnessing Homologous Recombination *In Vitro* to Generate Recombinant DNA via SLIC. *Nat. Methods* 4, 251–256. doi:10.1038/nmeth1010

Lin, J.-H., Lee, M.-C., Sue, Y.-S., Liu, Y.-C., and Li, S.-Y. (2017). Cloning of phaCAB Genes from Thermophilic Caldimonas Manganoxidans in *Escherichia coli* for Poly(3-Hydroxybutyrate) (PHB) Production. *Appl. Microbiol. Biotechnol.* 101, 6419–6430. doi:10.1007/s00253-017-8386-2

Million-Weaver, S., Camps, M., and Camps, M. (2014). Mechanisms of Plasmid Segregation: Have Multicopy Plasmids Been Overlooked? *Plasmid* 75, 27–36. doi:10.1016/j.plasmid.2014.07.002

Nordström, K., and Austin, S. J. (1989). Mechanisms that Contribute to the Stable Segregation of Plasmids. *Annu. Rev. Genet.* 23, 37–69.

Pecota, D. C., Kim, C. S., Wu, K., Gerdes, K., and Wood, T. K. (1997). Combining the Hok/sok, parDE, and Pnd Postsegregational Killer Loci to Enhance Plasmid Stability. *Appl. Environ. Microbiol.* 63, 1917–1924. doi:10.1128/aem.63.5.1917-1924.1997

Pecota, D., Osapay, G., Selsted, M., and Wood, T. (2003). Antimicrobial Properties of the R1 Plasmid Host Killing Peptide. *J. Biotechnol.* 100, 1–12. doi:10.1016/s0168-1656(02)00240-7

Pedersen, K., and Gerdes, K. (1999). Multiple Hok Genes on the Chromosome of *Escherichia coli*. *Mol. Microbiol.* 32, 1090–1102. doi:10.1046/j.1365-2958.1999.01431.x

Pelletier, J. F., Sun, L., Wise, K. S., Assad-Garcia, N., Karas, B. J., Deerinck, T. J., et al. (2021). Genetic Requirements for Cell Division in a Genomically Minimal Cell. *Cell* 184, 2430–2440. e2416. doi:10.1016/j.cell.2021.03.008

Peña, C., Castillo, T., García, A., Millán, M., and Segura, D. (2014). Biotechnological Strategies to Improve Production of Microbial Poly-(3-hydroxybutyrate): a Review of Recent Research Work. *Microb. Biotechnol.* 7, 278–293.

Sayeed, S., Reaves, L., Radnedge, L., and Austin, S. (2000). The Stability Region of the Large Virulence Plasmid of Shigella Flexneri Encodes an Efficient Postsegregational Killing System. *J. Bacteriol.* 182, 2416–2421. doi:10.1128/jb.182.9.2416-2421.2000

Schneider, D., Duperchy, E., Depeyrot, J., Coursange, E., Lenski, R. E., and Blot, M. (2002). Genomic Comparisons Among *Escherichia coli* Strains B, K-12, and O157:H7 Using IS Elements as Molecular Markers. *BMC Microbiol.* 2, 18. doi:10.1186/1471-2180-2-18

Schweder, T., Schmidt, I., Herrmann, H., Neubauer, P., Hecker, M., and Hofmann, K. (1992). An Expression Vector System Providing Plasmid Stability and Conditional Suicide of Plasmid-Containing Cells. *Appl. Microbiol. Biotechnol.* 38, 91–93. doi:10.1007/BF00169425

Singha, T. K., Gulati, P., Mohanty, A., Khasa, Y. P., Kapoor, R. K., and Kumar, S. (2017). Efficient Genetic Approaches for Improvement of Plasmid Based Expression of Recombinant Protein in *Escherichia coli* : A Review. *Process Biochem.* 55, 17–31. doi:10.1016/j.procbio.2017.01.026

Stubbe, J., Tian, J., He, A., Sinskey, A. J., Lawrence, A. G., and Liu, P. (2005). Nontemplate-dependent Polymerization Processes: Polyhydroxyalkanoate Synthases as a Paradigm. *Annu. Rev. Biochem.* 74, 433–480. doi:10.1146/annurev.biochem.74.082803.133013

Tseng, I.-T., Chen, Y.-L., Chen, C.-H., Shen, Z.-X., Yang, C.-H., and Li, S.-Y. (2018). Exceeding the Theoretical Fermentation Yield in Mixotrophic Rubisco-Based Engineered *Escherichia coli*. *Metab. Eng.* 47, 445–452. doi:10.1016/j.ymben.2018.04.018

Voigt, C. A. (2020). Synthetic Biology 2020-2030: Six Commercially-Available Products that Are Changing Our World. *Nat. Commun.* 11, 6379. doi:10.1038/s41467-020-20122-2

Wu, K., and Wood, T. K. (1994). Evaluation of Thehok/sok Killer Locus for Enhanced Plasmid Stability. *Biotechnol. Bioeng.* 44, 912–921. doi:10.1002/bit.260440807

Yang, C. H., Liu, E. J., Chen, Y. L., Ou-Yang, F. Y., and Li, S. Y. (2016). The Comprehensive Profile of Fermentation Products during *In Situ* CO2 Recycling by Rubisco-Based Engineered *Escherichia coli*. *Microb. Cel Fact* 15, 133–210. doi:10.1186/s12934-016-0530-7

Zielenkiewicz, U., and Cegłowski, P. (2001). Mechanisms of Plasmid Stable Maintenance with Special Focus on Plasmid Addiction Systems. *Acta Biochim. Pol.* 48, 1003–1023. doi:10.18388/abp.2001_3863

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Bioengineering and Biotechnology

REVIEW
published: 08 June 2022
doi: 10.3389/fbioe.2022.907861

From a Hetero- to a Methylotrophic Lifestyle: Flash Back on the Engineering Strategies to Create Synthetic Methanol-User Strains

*Camille Peiro, Cláudia M. Vicente, Denis Jallet and Stephanie Heux**

TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

Engineering microorganisms to grow on alternative feedstocks is crucial not just because of the indisputable biotechnological applications but also to deepen our understanding of microbial metabolism. One-carbon (C1) substrate metabolism has been the focus of extensive research for the prominent role of C1 compounds in establishing a circular bioeconomy. Methanol in particular holds great promise as it can be produced directly from greenhouse gases methane and carbon dioxide using renewable resources. Synthetic methylotrophy, i.e. introducing a non-native methanol utilization pathway into a model host, has therefore been the focus of long-time efforts and is perhaps the pinnacle of metabolic engineering. It entails completely changing a microorganism's lifestyle, from breaking up multi-carbon nutrients for growth to building C-C bonds from a single-carbon molecule to obtain all metabolites necessary to biomass formation as well as energy. The frontiers of synthetic methylotrophy have been pushed further than ever before and in this review, we outline the advances that paved the way for the more recent accomplishments. These include optimizing the host's metabolism, "copy and pasting" naturally existing methylotrophic pathways, "mixing and matching" enzymes to build new pathways, and even creating novel enzymatic functions to obtain strains that are able to grow solely on methanol. Finally, new approaches are contemplated to further advance the field and succeed in obtaining a strain that efficiently grows on methanol and allows C1-based production of added-value compounds.

Keywords: synthetic methylotrophs, methanol, metabolic engineering, modelling, biotechnology

OPEN ACCESS

Edited by:
Jean Marie François, Institut Biotechnologique de Toulouse (INSA), France

Reviewed by:
Hiroya Yurimoto, Kyoto University, Japan
Ethan I. Lan, National Yang Ming Chiao Tung University, Taiwan

****Correspondence:***
Stephanie Heux
stephanie.heux@insa-toulouse.fr

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *30 March 2022*
Accepted: *16 May 2022*
Published: *08 June 2022*

Citation:
Peiro C, Vicente CM, Jallet D and Heux S (2022) From a Hetero- to a Methylotrophic Lifestyle: Flash Back on the Engineering Strategies to Create Synthetic Methanol-User Strains. Front. Bioeng. Biotechnol. 10:907861. doi: 10.3389/fbioe.2022.907861

INTRODUCTION

Synthetic methylotrophy refers to the design and engineering of methanol assimilation pathways into established non-methylotrophic production hosts making use of their vast biotechnological potential (Becker et al., 2015) and providing access to methanol as feedstock.

To date, attempts to introduce methylotrophy into biotechnologically relevant microbes have been described for *Escherichia coli* (Muller et al., 2015; Whitaker et al., 2017), *Corynebacterium glutamicum* (Lessmeier et al., 2015; Witthoff et al., 2015), *Pseudomonas putida* (Koopman et al., 2009), *Saccharomyces cerevisiae* (Dai et al., 2017) and *Yarrowia lipolytica* (Vartiainen et al., 2019). Efforts to engineer this non-native nutrient catabolism have relied on four different levels of engineering (Erb et al., 2017). In the first level, existing pathways are used to allow methanol assimilation in the non-methylotrophic host (Espinosa et al., 2020). In the next level, synthetic

methylotrophs were obtained by "copying and pasting" naturally occurring methanol assimilation pathways (Muller, et al., 2015; Dai, et al., 2017). In the third level, synthetic methanol assimilation pathways were developed by "mixing and matching" enzymes from these methanol assimilation pathways (De Simone et al., 2020). In the last level, novel methanol assimilation pathways were created from known or new enzyme mechanisms (Siegel et al., 2015). However, none of these synthetic strains was able to grow on methanol alone.

If the implementation of methylotrophy in non-native methylotrophs can be seen as a quite straightforward approach, researchers quickly realised that the complexity of methylotrophy could not be reduced to a simple metabolic transplant. In contrast with heterotrophic metabolism where multi-carbon substrates are essentially broken down to obtain metabolites and biomass building blocks, in methylotrophy all carbon-carbon bonds essential to life must be built from a single carbon (C1) molecule. Furthermore, in all these implemented pathways carbon is assimilated in the form of formaldehyde, a central but highly toxic intermediate that can lead to cell death in case of imbalance between dissimilation and assimilation. In the engineering strategies based on naturally occurring methanol assimilation pathways, formaldehyde assimilation is achieved by a cyclic process and requires a C1-acceptor that enables the formation of a C-C bond (Wendisch et al., 2021). The efficiency of methanol assimilation is then determined by the capability of the cells to produce and regenerate the C1-acceptor, especially when C1 is assimilated as formaldehyde. In contrast, the novel pathways, created from novel reactions, are linear and thus independent from any C1-acceptor. However, as formaldehyde is the only intermediate, a high concentration of formaldehyde is required to support a sustainable rate of formaldehyde condensation that enables the formation of a C-C bond. The efficiency of methanol assimilation is then determined by the capability of the cells to tolerate a higher level of formaldehyde. To overcome these obstacles, different rational and evolutionary engineering strategies have been applied during the last decade. The ultimate growth phenotype was recently achieved by Chen et al., 2020 and Kim et al., 2020 who observed growth on methanol as the sole carbon source by a synthetic methylotrophic *E. coli* (Chen et al., 2020; Kim et al., 2020). This review proposes to present the various strategies implemented to reprogram the food diet of non-methylotrophic production hosts.

EXISTING PATHWAYS

In this approach, existing pathways within the natural host are engineered (i.e. through gene deletion or overexpression) to improve their overall capacity towards a given function. A few studies have followed this approach to implement synthetic methylotrophy.

In *S. cerevisiae*, when grown with methanol (plus yeast extract), the final optical density increased by 39% relative to medium without methanol. In addition, ^{13}C from labeled methanol was incorporated into intracellular metabolites (60% of the acetyl-coA pool was fully labeled). Based on these observations, Espinosa et al. decided to further boost the assimilation of methanol in *S. cerevisiae* by adaptive laboratory evolution (ALE). Genomic analysis of the resulting evolved strains helped in the identification of beneficial mutations allowing improvement of methanol utilization. These mutations were implemented in a wild type strain and the reconstructed strain exhibited a higher biomass yield on methanol (a 21% increase in OD600) than the parental strain. In addition, a ^{13}C-methanol tracer analysis revealed that the reconstructed strain had a higher percentage of fully ^{13}C-labelled intracellular metabolites (33% of fructose-1,6-bisphosphate and 60% of acetyl-CoA fully labelled) compared with the parental strain (Espinosa, et al., 2020).

It is interesting to notice that other non-methylotrophic microorganisms possess endogenous enzymes that may confer a native capacity for methylotrophy. Endogenous methanol oxidizing activities, which may stem from promiscuous alcohol dehydrogenases, have been observed in Y. *lipolitica* (Koopman, et al., 2009) and *P. putida* (Vartiainen, et al., 2019). *Bacillus subtilis* possesses a Ribulose Monophosphate (RuMP) pathway which may act as a detoxification system for formaldehyde (Yasueda et al., 1999). However, no studies report the use of ALE to optimize the native methylotrophic metabolism of those microorganisms.

COPY AND PASTE ENZYMES

In this more advanced approach, naturally occurring metabolic pathways for methanol assimilation are implemented in another host and optimized by shaping the metabolic network of the host to fit the acquired property. Aerobic natural methylotrophy is supported by four different metabolic pathways found among methylotrophic bacteria and yeast: the RuMP pathway, the Xylulose Monophosphate (XuMP) pathway, the Calvin-Benson-Bassham (CBB) cycle and the Serine cycle (Wendisch, et al., 2021). In thoses studies aiming at "copying and pasting" existing methylotrophic pathways into non-methylotrophic hosts, the RuMP pathway has been used far more frequently than the others.

The RuMP Pathway and Its Optimization

In this pathway only three heterologous enzymes, i.e. a methanol dehydrogenase (Mdh), a 3-hexulose-6-phophate synthase (Hps) and a 6-phospho-3-hexuloisomerase (Phi) are needed to integrate the methylotrophic module (Yurimoto et al., 2009). The challenging parts often consist in connecting methanol assimilation to the host central metabolism and in ensuring proper C1-acceptor regeneration.

Rational Engineering

Screening and Engineering of Mdh

The oxidation of methanol to formaldehyde, the first step in methanol assimilation, is ensured by an Mdh. Depending on the electron acceptor, Mdh enzymes can be classified into three groups: the periplasmic pyrrolo-quinoline-quinone (PQQ)-

dependent Mdh found in Gram-negative methylotrophs, the cytoplasmic NAD-dependent Mdh common in Gram-positive methylotrophs and the cytoplasmic NDMA (N,N-dimethyl-4-nitrosoaniline)-dependent Mdh found in *Mycobacterium* and that uses mycofactocin as an *in vivo* electron acceptor (Heux et al., 2018; Dubey et al., 2019). The PQQ- and NDMA-dependent Mdhs are often multimeric proteins (Keltjens et al., 2014; Dubey, et al., 2019) and require the synthesis of PQQ and mycofactocin through complex biosynthetic pathways absent in a lot of hosts. The PQQ-dependent Mdh requires oxygen and specific cellular locations for its proper function while the NAD-dependent Mdh is located in the cytoplasm and functions under both aerobic and anaerobic conditions. For all these reasons, the choice of a NAD-dependent Mdh appeared to be the simplest way to implement synthetic methylotrophy. The *mdh* gene from *Bacillus methanolicus* was the first to be used (Lessmeier, et al., 2015; Muller, et al., 2015; Witthoff, et al., 2015). However, its efficiency towards methanol conversion remains relatively low and the same states for several reported variants. Thereby it required a high amount of methanol to yield high amounts of intracellular labelling (25% ^{13}C-enrichment in phosphoenolpyruvate (PEP) with 1M of ^{13}C-methanol) (Muller, et al., 2015). Scientists therefore aimed to find new NAD-dependent Mdhs with better kinetic parameters by screening Mdhs coming from different donor organisms. Among them, an alcohol dehydrogenase from the non-methylotrophic *Bacillus stearothermophilus* was found to use methanol as a substrate with a 10 times lower K_m value (Sheehan et al., 1988). Whitaker and colleagues used this Mdh combined with Hps and Phi from *B. methanolicus* in *E. coli* and showed that the combination resulted in a 43% increase in biomass yield and a 10 times higher ^{13}C-enrichment in intracellular metabolites compared to the association using the Mdh from *B. methanolicus* in a media containing 16 times less methanol (i.e. 60 mM vs 1M) but yeast extract (Whitaker, et al., 2017). Another alcohol dehydrogenase from *Cupriavidus necator* was found to exhibit similar activity towards methanol at 30°C to the Mdh from *B. methanolicus* at 45°C (its optimal temperature for activity). To further improve the kinetic parameters of Mdh, engineering strategies were applied. By using directed molecular evolution, a variant of the Mdh from *C. necator* with a catalytic efficiency for methanol 6-fold higher than the wild-type was obtained (Wu et al., 2016). In another study, by using phage assisted evolution on the Mdh2 of *B. methanolicus*, a two times higher methanol incorporation was observed in the resulting strain compared to the strain expressing the native Mdh2 (Roth et al., 2019).

Engineering of the Cell Redox State

Cellular redox state directly affects the Mdh activity since high NADH/NAD$^+$ ratios are unfavourable to methanol oxidation. Results obtained after ALE experiments aiming at improving methanol assimilation in *E. coli* showed mutations in the *nadR* gene encoding for a transcriptional regulator of genes involved in NAD$^+$ transport and *de novo* synthesis (Meyer et al., 2018). The activity of this repressor was found to be reduced, highlighting the importance of the NADH/NAD$^+$ ratio during growth on methanol. To balance the redox state of the cell, authors also showed that when the NAD-dependent malate dehydrogenase gene *maldh* was knocked-out, growth was improved on methanol and gluconate. This was also observed in *E. coli* during growth on methanol and yeast extract even if methanol assimilation was not improved (Rohlhill et al., 2020). In a similar way, the methanol oxidation rate was improved when Mdh was coupled with a "NADH sink" by using lactate dehydrogenase to recycle NADH into NAD$^+$ (Price et al., 2016).

Screening of Hps and Phi Candidates

The implementation of a methylotrophic pathway in *E. coli* was done by testing a series of Hps and Phi candidates from different donor organisms *in vitro* and *in vivo* (Muller, et al., 2015; Whitaker, et al., 2017; Fan et al., 2018). In all these studies, Hps and Phi from *B. methanolicus* were the best performing enzymes both *in vitro* and *in vivo*. In *C. glutamicum*, methanol utilization has been achieved by expressing Mdh from *B. methanolicus* together with *hxlA* (3-hexulose-phosphate synthase) and *hxlB* (6-phospho-3-hexuloisomerase) from *B. subtilis* (Lessmeier, et al., 2015; Witthoff, et al., 2015). In the resulting strains, the incorporation of ^{13}C-label from ^{13}C-methanol into central metabolites was detected, demonstrating the *in vivo* operation of the synthetic methanol utilization pathway (Lessmeier, et al., 2015; Witthoff, et al., 2015). However, expression of the same genes in *S. cerevisiae* failed to allow methanol consumption and cell growth in a minimal medium containing methanol as the sole carbon source (Dai, et al., 2017). In the bacterium *P. putida*, introduction of the *hps* and *phi* genes from *Bacillus brevis* allowed the strain to utilize methanol and formaldehyde as auxiliary substrates (Koopman, et al., 2009). The Hps and Phi expressing strain showed a two times higher biomass yield compared to the control strain when grown on a medium containing formaldehyde plus glucose. Furthermore, the strain was also able to grow when replacing formaldehyde by methanol while the control strain did not reach steady state under these conditions. However, authors did not show any evidence that a functional RuMP was operating *in vivo*. Overall, these studies demonstrate that the efficiency of Hps and Phi is not only linked with their origin but also with the host in which they are expressed. Improvement of Hps and Phi catalytic efficiency has been obtained by fusing both enzymes together and this will be discussed in the section Spatial engineering.

Engineering of the Dissimilatory and Recycling Pathways

In order to drive the metabolism towards methanol assimilation, a common strategy is to delete one or more genes encoding for the formaldehyde detoxification pathway to avoid a carbon loss as CO_2 (Whitaker, et al., 2017; Bennett et al., 2018; Vartiainen, et al., 2019; De Simone, et al., 2020; Rohlhill, et al., 2020). The importance of this mutation has been confirmed in ALE experiments where all the evolved strains with improved methylotrophic capacity had a deletion in one or more genes of the formaldehyde detoxification operon (Chen et al., 2018; Meyer, et al., 2018; Chen, et al., 2020).

Another strategy to stimulate methanol assimilation was to favour the recycling of the C1-acceptor (i.e. Ribulose-5-

phosphate, Ru5P). In synthetic methylotrophs, Ru5P regeneration is ensured by the non-oxidative part of the pentose phosphate pathway (PPP). In the natural methylotroph *B. methanolicus*, it was shown that *pfk*, *rpe*, *tkt*, *glpX*, and *fba* (i.e. part of the sedoheptulose-1,7-biphosphatase (SBPase) variant of the RuMP pathway) are key genes involved in Ru5P regeneration. Their deletion resulted in the loss of the capacity of the bacteria to grow on methanol (Brautaset et al., 2004). Therefore, enhancing the host's capacity to regenerate Ru5P by overexpressing heterologous enzymes appears critical to achieve efficient synthetic methylotrophy. To boost Ru5P regeneration, Bennett et al. overexpressed PPP enzymes by integrating *pfk*, *rpe*, *tkt*, *glpX* and *fba* from *B. methanolicus* under a strong inducible promoter into the *E. coli* chromosome (Bennett, et al., 2018). Even if culture media had to be supplemented with yeast extract, the resulting strain achieved a 20% improvement in biomass production during growth on methanol compared to the parental strain and 59% of ^{13}C-enrichment was reached in PEP. PEP is an interesting metabolite to follow methanol assimilation and Ru5P recycling. Indeed, the more carbons in PEP are labelled, the more methanol has been assimilated and so the more Ru5P has been recycled. Overexpressing the PPP enzymes does not suffice, as a fine balance must be found between the amounts of methanol carbon directed towards recycling vs biomass formation. Woolston et al. used iodoacetate to block glyceraldehyde-3-phosphate dehydrogenase (Ga3PDH). Authors hypothesised that by inhibiting this reaction, which connects the RuMP pathway to lower glycolysis, larger intermediate pools of the upper glycolytic and RuMP pathways could be maintained in starved cells. This strategy led to a higher methanol incorporation into central metabolites in *E. coli*. The ^{13}C-enrichment in fructose-6-phosphate (F6P) reached 27.5% when iodoacetate was present. Moreover, the intracellular concentrations of F6P, Ru5P and sedoheptulose-7-phosphate (S7P) were higher (Woolston et al., 2018). Rohlhill et al. went one step further by modulating the expression of *rpe* and *tkt* using the native formaldehyde-inducible promoter P_{frm} of *E. coli*. When combined with the disruption of the malate dehydrogenase, methanol carbon incorporation into intracellular metabolites was improved, but yeast extract was still needed during the experiment (Rohlhill, et al., 2020).

To boost both assimilation and recycling, Chen et al. first integrated two operons in *E. coli*'s genome. The first one included *mdh*, *hps* and *phi* genes while the second one was composed by the same *mdh* and *phi* genes plus the *hps, tkt* and *tal* genes from various organisms. Adding the *tkt* and *tal* genes in the operon helped enhancing Ru5P recycling (Chen, et al., 2020). Moreover, after the first ALE experiment, using Ensemble Modelling for Robustness Analysis (EMRA), authors identified that the high activity of phosphofructokinase and Ga3PDH was channelling the flux away from the RuMP cycle, which tended to unbalance the metabolic system. To reduce the activity of these two enzymes, *pfkA* was knockout and the *gapA* gene was replaced by another *gapC* from *E. coli* BL21 encoding for a less efficient Ga3PDH than the native one.

Engineering of the Regulation of Gene Expression

Natural methylotrophs have developed complex regulatory networks allowing them to express the genes required for the methanol metabolism via sensing methanol and/or formaldehyde (Jakobsen et al., 2006). In contrast, synthetic methylotrophs do not possess such genetic machinery to recognize methanol. Therefore, attempts to make synthetic methylotrophs sensing methanol have been carried out. A chimeric two-component system was created in *E. coli*, MxaYZ, by fusing the periplasmic methanol sensing domain MxaY of the Mdh from *Paracoccus denitrificans* with the cytoplasmic catalytic transmitter domain of EnvZ from *E. coli* (Ganesh et al., 2017). MxaYZ sensed extracellular methanol and could activate the expression of a fluorescent reporter. Formaldehyde on the other hand is naturally encountered by *E. coli* via the formaldehyde-inducible promoter (P_{frm}) which activates the dissimilatory pathway to avoid any intracellular formaldehyde accumulation. Several examples of engineered P_{frm} in synthetic methylotrophs have been described (Denby et al., 2016; Rohlhill et al., 2017). In *E. coli,* P_{frm} is situated upstream to the *frmRAB* operon and is repressed by FrmR, a transcriptional repressor. When formaldehyde is present in the cells, it interacts with FrmR and triggers conformational changes, leading to the dissociation of FrmR from its recognized DNA motif within P_{frm} and to the derepression of the transcription of the operon (Denby, et al., 2016). When placed upstream the *mdh*, *hps* and *phi* operon, both the native and engineered P_{frm} enabled improved biomass production in the engineered strain (Rohlhill, et al., 2017). In another study, FrmR and P_{frm} were used to build a formaldehyde biosensor in *E. coli* and to modulate the expression of *mdh*, *hps* and *phi* (Woolston, et al., 2018).

Evolutionary Engineering

Despite these efforts, rational engineering strategies were not sufficient to ensure full synthetic C1-assimilation in microorganisms. Rational engineering had to be associated with evolutionary engineering to improve the microorganism's performance. However, as no growth on pure methanol was initially observed in the synthetic methylotrophs, the strategy adopted by several groups was to first build an "auxotrophic" strain in which methanol is necessary to ensure growth when another carbon source (i.e. glucose, xylose, ribose, gluconate or pyruvate) is present and then to subject the strain to ALE (Chen, et al., 2018; Meyer, et al., 2018; Bennett et al., 2020b; Chen, et al., 2020; Keller et al., 2020).

Chassis Optimization to Engineer Methanol-dependent Strain for Growth

To ensure methanol is co-utilized with the other carbon source, cells are engineered in order to gain a benefit in utilizing methanol. To do so, the common strategy was to design a strain dependent on methanol utilization for growth while the other carbon source is used only to ensure Ru5P production. This is achieved by blocking either the conversion of glucose to F6P or the Ru5P catabolic pathways depending on the carbon source used (**Figure 1**). That way, cells have no other choice than producing F6P from the co-assimilation of methanol and the other carbon source while ensuring a high pool of the key Ru5P. This strategy was applied by Meyer et al. to engineer an *E. coli* methanol dependent-strain (Δ*rpiAB*Δ*edd*Δ*maldh*) co-utilizing

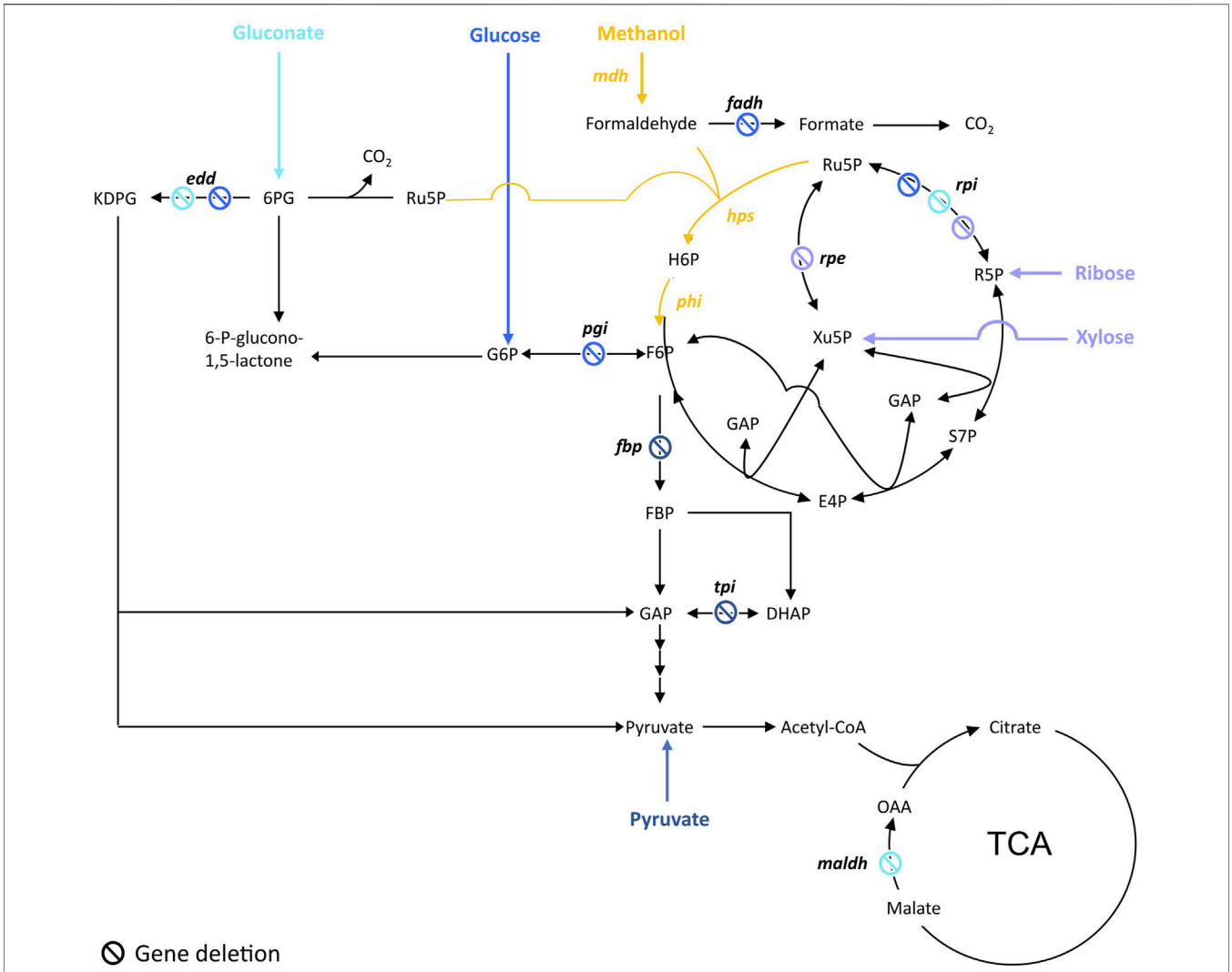


FIGURE 1 | Targeted genes to construct methanol-dependent strains in *E. coli*. Adapted from (Wang et al., 2020b). Engineered synthetic methanol-dependent strain co-assimilating gluconate (in light blue) by (Meyer, et al., 2018); co-assimilating glucose (in blue) by (Bennett, et al., 2020b); co-assimilating ribose by (Chen, et al., 2018) or xylose by (Chen, et al., 2018; Chen, et al., 2020) (in purple); co-assimilating pyruvate (in dark blue) by (Keller, et al., 2020). Enzymes are written in bold. edd, 6-phosphogluconate isomerase; fadh, formaldehyde dehydrogenase; hps, 3-hexulose-6-phosphate synthase; maldh, malate dehydrogenase; mdh, methanol dehydrogenase; pgi, glucose-6-phosphate isomerase; phi, 6-phospho-3-hexuloisomerase; rpe, ribulose-3-phosphate epimerase; rpi, ribose-5-phosphate isomerase. Ac-CoA, acetyl-CoA; E4P, erythrose-4-phosphate; F6P, fructose-6-phosphate; FBP, fructose-1,6-biphosphate; GAP, glyceraldehyde-3-phosphate; DHAP, dihydroxyacetone-phosphate, H6P, hexulose-6-phosphate; KDPG, 2-keto-3-deoxy-6-phosphogluconate; 6 PG, 6-phosphogluconate; R5P, ribose-5-phosphate; Ru5P, ribulose-5-phosphate; S7P, sedoheptulose-7-phosphate; SBP, sedoheptulose-1,7-biphosphate; Xu5P, xylulose-5-phosphate.

gluconate (Meyer, et al., 2018). Similarly, Chen et al. engineered an *E. coli* strain (Δ*rpiAB*) co-utilizing xylose and another one (Δ*rpe*) co-utilizing ribose (Chen, et al., 2018). Later, to build their fully methylotrophic strain, Chen et al. adopted the same strategy than in 2018 by deleting *rpiAB* but switched from *E. coli* BL21 to *E. coli* K12 BW25113 to facilitate the genetic engineering efforts (Chen, et al., 2020). To enable the co-utilization of methanol and glucose, Bennet et al. first deleted the phosphoglucose isomerase (*pgi*) that converts glucose-6-phosphate (G6P) to F6P, so that G6P is only used by the oxidative part of the PPP to directly fill the pool of Ru5P. The *frmA* gene, encoding for formaldehyde dehydrogenase, was also deleted in order to push formaldehyde towards assimilation (Bennett, et al., 2018). Later, authors went ahead by knocking-out *edd* and *rpiAB* in their Δ*frmA*Δ*pgi E. coli* strain to completely abolish glucose assimilation but amino acids supplementation in the cultivation medium was then required (Bennett, et al., 2020b). Keller et al. used a different strategy by applying first a flux balance analysis and identifying candidate deletions *in silico* that could lead to methanol-auxotrophy with a complete RuMP cycle and a high degree of methanol incorporation. Authors investigated two out of 1,200 candidate strains, one with a deletion of fructose-1,6-bisphosphatase (*fbp*) and another with triosephosphate isomerase (*tpiA*) deleted. Those strains were methanol dependent and showed a 99%

TABLE 1 | Overview of ^{13}C-methanol incorporation level at the intracellular level or in final products in strains after ALE experiment.

Organism	Cultivations Conditions	^{13}C-Enrichment	References
E. coli BW25113	500 mM ^{13}C-methanol + 5 mM gluconate	21% in PEP	Meyer, et al. (2018)
E. coli BL21 DE3	250 mM ^{13}C-methanol + 50 mM xylose	22% in butanol	Chen, et al. (2018)
E. coli BW25113	500 mM ^{13}C-methanol + 200 mM glucose	22% in acetone	Bennett, et al. (2020b)
C. glutamicum ATCC 13032	125 mM ^{13}C-methanol + 27 mM xylose	22% in PEP	Tuyishime, et al. (2018)
C. glutamicum ATCC 13032	469 mM ^{13}C-methanol + 27 mM xylose	20–30% in amino acids	Wang, et al. (2020a)
C. glutamicum ATCC 13032	500 mM ^{13}C-methanol +20 mM gluconate+ 0.5 g/L yeast extract	30% in cadaverine	Hennig, et al. (2020)

fractional incorporation of methanol into RuMP cycle metabolites (Keller, et al., 2020).

If most of the attempts were done in *E. coli*, methanol dependent strains were also obtained in *C. glutamicum* by generating Δ*aldh*Δ*fadH*Δ*rpi* and Δ*aldh*Δ*fadH*Δ*rpe* mutants, respectively using xylose or gluconate and ribose or gluconate as co-substrates (Hennig et al., 2020; Tuyishime et al., 2020). The *aldH* and *fadH* genes are respectively coding for an acetaldehyde dehydrogenase and a formaldehyde dehydrogenase, so their deletion prevented formaldehyde oxidation to CO_2. More recently, synthetic methanol auxotrophy has also been engineered in *Bacillus subtilis* by deleting *rpiAB* and *rpe* (Gao et al., 2022).

Adaptive Laboratory Evolution

By engineering methanol-dependent strains, methanol assimilation is coupled with cell growth thus opening the way to ALE to improve methanol assimilation. ALE is used on microorganisms cultivated under defined conditions for weeks to months in order to select improved phenotypes on relevant carbon sources. Microorganisms can either be cultured in shake flasks with sequential serial passages or cells can be grown in chemostats with a controlled environment in which one component of the medium is limiting. Shake flasks have the great advantage to enable to run many culture conditions in parallel but pH, oxygenation and cell density can vary during the experiment. By using a chemostat, these parameters as well as growth rate can be closely monitored and kept constant, and higher cell densities can be reached. However, the operation costs and the know-how of this device are higher than for shake flasks (Dragosits et al., 2013).

Mutants evolving towards a faster growth can be selected, and methanol utilization can be tested by measuring the incorporation of label in intracellular metabolites from ^{13}C-methanol (**Table 1**). Meyer et al. decided to knock out *maldh* in a Δ*rpiAB*Δ*edd* background after *in silico* analysis, and to apply ALE to this *E. coli* strain. A mutant co-utilizing methanol and gluconate at a growth rate of $\mu = 0.08\ h^{-1}$ and with a methanol uptake rate of 13 mmol. $gCDW^{-1}$. h^{-1} - which is close to the one reported for natural methylotrophs—was selected. In this condition, 21% of the PEP carbon atoms came from methanol (Meyer, et al., 2018). When Chen et al. used ALE on their Δ*rpiAB* strain, they succeeded to select a strain with a growth rate of $\mu = 0.17\ h^{-1}$. This strain was able to produce butanol from the co-consumption of methanol and xylose, and 22% of butanol carbons came from methanol (Chen, et al., 2018). After the ALE experiment, Bennet et al. isolated an *E. coli* strain that was growing on glucose and methanol with a specific growth rate of $\mu = 0.15\ h^{-1}$ and without the need for amino acid supplementation. The strain was able to produce acetone, for which 22% of carbon atoms derived from methanol (Bennett, et al., 2020b). An ALE experiment was successful for Chen et al. as they obtained an evolved methylotrophic *E. coli* strain that exhibited a growth rate of $\mu = 0.08\ h^{-1}$ on pure methanol (Chen, et al., 2020). However, no proof of *in vivo* functionality of the introduced methanol pathway was shown.

Tuyshime et al. performed ALE experiments on their *C. glutamicum* mutant and selected a strain growing on xylose and methanol at a growth rate of $0.03\ h^{-1}$ with a methanol uptake rate of 0.86 mmol. $gCDW^{-1}$. h^{-1} (Tuyishime et al., 2018). Based on transcriptome analysis, further work was done on this strain to explore the metabolic regulation that operated during methylotrophic conditions (Fan et al., 2021). Results demonstrated that the evolved *C. glutamicum* used the SBPase variant to recycle Ru5P as *fba* and *glpX* were upregulated while *tal* was downregulated. Nitrate was also shown to serve as a complementary electron acceptor during aerobic methanol metabolism as the operon *narKGHJI* was upregulated. This operon encodes for a respiratory nitrate reductase which is usually active during anaerobic growth. Amino acids biosynthesis was also shown to limit growth on methanol as genes involved in the biosynthesis of several L-aspartate derived amino acids were downregulated. This strain was subjected to a second ALE experiment which enabled to select a new evolved strain exhibiting a growth rate of $0.052\ h^{-1}$ on xylose and methanol as well as a higher tolerance towards methanol (Wang et al., 2020a). In another study, Henning et al. selected, in *C. glutamicum*, an evolved Δ*aldh*Δ*fadH*Δ*rpe* strain co-utilizing methanol and gluconate. This strain was able to produce cadaverine, for which 43% of carbon atoms came from methanol. After a new round of ALE experiments started from the previous strain, another one was selected on ribose and methanol. This strain grew without yeast extract and exhibited a specific growth rate of $\mu = 0.10\ h^{-1}$ (Hennig, et al., 2020).

Resulting Metabolic Adaptation After ALE

As in nature, during ALE genomic changes occur to generate an improved phenotype enabling cells to cope with their environment. It is of great interest to get insight on which genes are essential (or non-essential) for methanol assimilation by investigating the mutations and their consequences. Mutations found in evolved strains can be classified in three categories,

mutations affecting: 1) the dissimilatory pathway and the recycling pathways; 2) the redox and energy balance; 3) the substrate uptake and other enzymes connected to metabolism (Wang et al., 2020b).

Mutations in the *frm* operon genes leading to the inactivation of the formaldehyde dissimilatory pathway were found in all the evolved *E. coli* strains (Chen, et al., 2018; Meyer, et al., 2018; Chen, et al., 2020). These mutations are consistent with the fact that efficient methanol-dependent growth required formaldehyde to be entirely redirected towards C1-assimilation. Mutations leading to *pykF* and *zwf* inactivation possibly enabled to increase the F6P pool size for Ru5P regeneration via the PPP (Chen et al., 2018). Similarly, the *deoD* mutation was thought to increase the Ru5P pool size for formaldehyde assimilation (Chen, et al., 2018).

Mutations were found in *nadR* (Meyer, et al., 2018), *cyaA* (Chen, et al., 2018) in *E. coli* and *mtrA* in *C. glutamicum* (Tuyishime, et al., 2018). The enzymes encoded by these genes are involved in the redox state or in the energy supply of the cell. The *nadR* and *mtrA* genes encode for enzymes involved in the balance of the redox state. Both mutations are thought to lead to the increase of NAD^+ availability in the cell, thereby most probably promoting methanol oxidation. *cyaA* encodes for an enzyme producing cyclic AMP (cAMP) from ATP. The transcription of most of the TCA cycle enzymes is activated by cAMP. Most probably a *cyaA* mutation would lower the TCA cycle activity and consequently NAD(P)H production, at a beneficial cost for methanol oxidation. Moreover, inactivation of *frmA* and *fdoG* (encoding a formate dehydrogenase) for the synthetic methylotrophic strain, help limiting further the production of NADH (Chen, et al., 2020).

Mutations in *gnd* encoding for 6-phosphogluconate dehydrogenase were found in a strain assimilating methanol better than the parental one but still requiring xylose for growth (Chen, et al., 2020). GntR and AltR are transcriptional regulators and mutations in genes encoding these proteins were found respectively in the *E. coli* strain co-utilizing methanol and gluconate and *C. glutamicum* strain co-utilizing methanol and xylose (Meyer, et al., 2018; Tuyishime, et al., 2018). GntR represses two enzymes involved in gluconate uptake, therefore *gntR* mutations may have improved gluconate uptake. Similarly, AltR regulates among other genes the expression of *xylB* encoding for a xylulose kinase and *adhA* encoding for an alcohol dehydrogenase functioning as a methanol dehydrogenase in *C. glutamicum*. Therefore, an *altR* mutation may have modified *xylB* and *adhA* expression leading to improved co-utilization of methanol and xylose by increasing substrate uptake. In the strain exhibiting higher tolerance towards methanol, two mutations in *cgl0653* and *cgl0833,* respectively encoding for an *O*-acetyl-L-homoserine sulfhydrylase and for a methanol-induced membrane-bound transporter were found (Wang, et al., 2020a).

Spatial Engineering

Once produced, formaldehyde has to be quickly condensed with Ru5P to avoid its toxic effect. To optimize formaldehyde assimilation, another strategy has been proposed that consists in enhancing substrate channelling by co-localizing the enzymes in a delimited space. Orita et al. were the first to address this question by fusing *hps* and *phi* from *Mycobacterium gastri* (Orita et al., 2007). *In vivo* assays showed that when expressed in *E. coli*, the fusion protein of Hps and Phi leads to an increased growth rate on formaldehyde than when the proteins were separated. Later, Price et al. advanced the field by engineering a supramolecular enzyme complex with Mdh from *B. methanolicus* and Hps and Phi from *Mycobacterium gastri* (Price, et al., 2016). They took advantage of the decameric structure of Mdh to design a supramolecular complex able to self-assemble by using SH3-ligand in order to "plug" on it the fusion protein of Hps and Phi previously described by Orita et al. This engineered complex enabled a faster conversion of methanol into F6P compared to unassembled proteins *in vitro*. In the same way, Fan et al. used an alternative strategy to fuse Mdh from *B. stearothermophilus* with Hps and Phi from *B. methanolicus* by using flexible linkers $(GGGGS)_n$. They also demonstrated an improvement in methanol conversion to F6P *in vitro* (Fan, et al., 2018).

The XuMP Pathway

Dai et al. introduced the XuMP pathway in the non-methylotrophic yeast *S. cerevisae* by integrating an alcohol oxidase, a catalase, a dihydroxyacetone synthase (Das) and a dihydroxyacetone kinase (DhaK) from *Pichia pastoris* in its genome (Dai, et al., 2017). By using the native peroxisome targeting peptide of *P. pastoris*, enzymes were addressed to the *S. cerevisae* peroxisome. When yeast extract was added to the medium, methanol consumption was further improved from 1.04 g/L to 2.35 g/L, suggesting that Xu5P recycling was enhanced by yeast extract addition. However, no proof of *in vivo* functionality of the introduced methanol pathway was shown.

In the yeast *Y. lipolitica*, the expression of a dihydroxyacetone synthase from *Candida boidini* restored the formaldehyde tolerance of a formaldehyde sensitive strain (i.e. deleted for the FaDH) but was not enough to convert methanol into biomass (Vartiainen, et al., 2019).

MIX AND MATCH ENZYMES

Different from the above strategies that mainly build on existing pathways, the "mix and match" approach allowed combining existing enzymes from different microbial sources into a synthetic pathway and implementing it in a host. Considering the methylotrophic metabolic diversity, we can estimate that there are more than 500 unique methanol assimilation pathways from methanol to biomass (Heux, et al., 2018). Therefore, such an approach is perfectly adapted to explore this large space of metabolic solutions while identifying the most powerful combinations. At the moment, only four hybrid pathways have been studied (**Figure 2**).

The Serine and Homoserine Cycles

Yu and Liao implemented a modified serine cycle in *E. coli* (**Figure 2B**). Methanol is assimilated via the H_4F-dependent

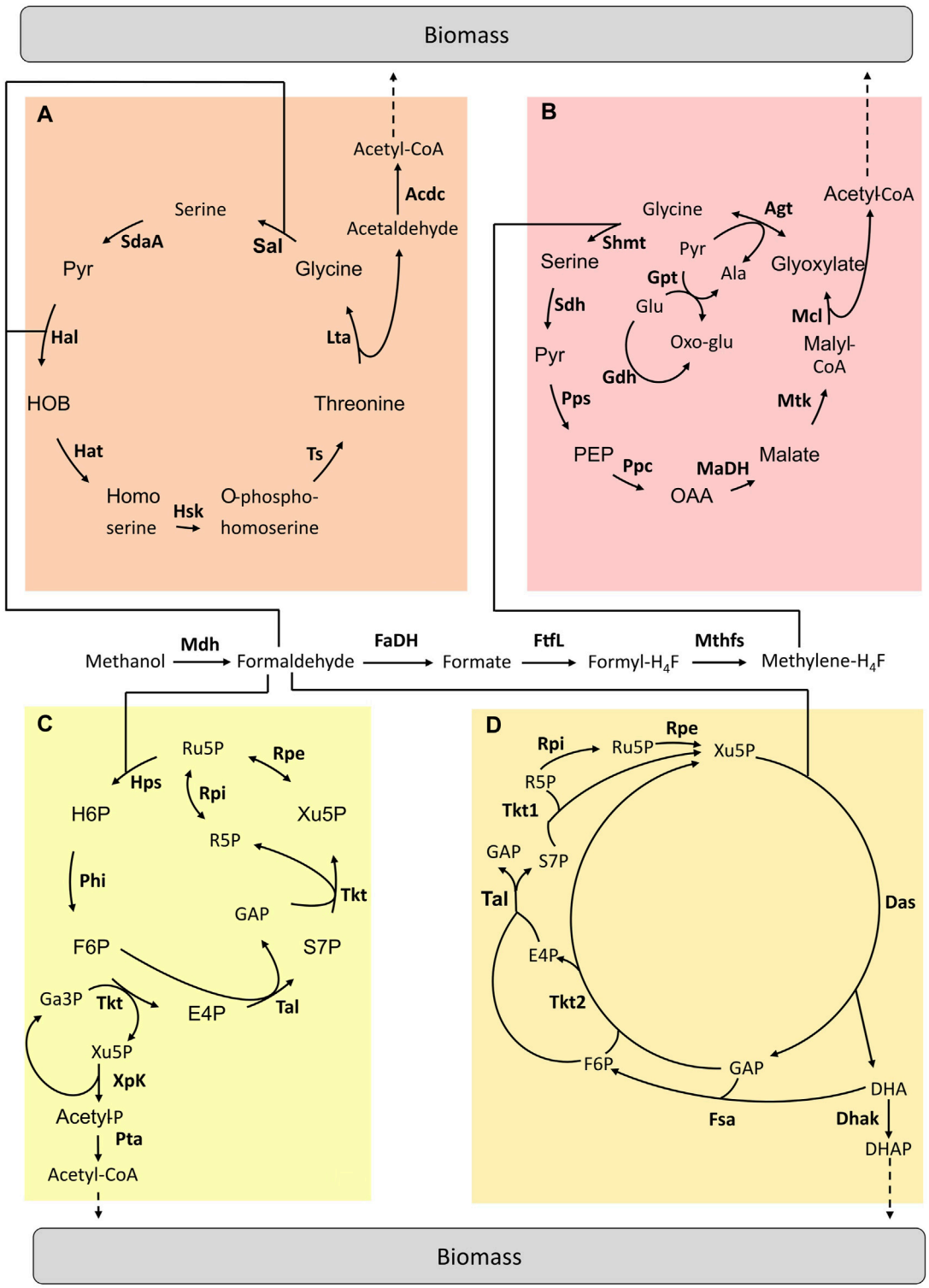


FIGURE 2 | Hybrid pathways implemented for C1-assimilation. **(A)** Homoserine cycle pathway **(B)** Modified serine cycle pathway **(C)** Methanol condensation cycle pathway **(D)** Hybrid Mdh-Das pathway. Adapted from (Tuyishime, et al., 2020). Enzymes are written in bold. Acdc, acetaldehyde dehydrogenase; Agt, Alanine-glyoxylate aminotransferase; Dhak, DHA kinase; FaDH, formaldehyde dehydrogenase; FtfL, formate-H_4F ligase; Fsa, Fructose-6-aldolase; Gdh, glutamate dehydrogenase; Gpt; Glutamate-pyruvate transaminase; Hal, 4-hydroxy-2-oxobutanoate aldolase; Hat, 4-hydroxy-2-oxobutanoate aminotransferase; Hsk, homoserine kinase; Lta, threonine aldolase; MaDH, malate dehydrogenase; Mcl, malyl-CoA lyase; Mdh, methanol dehydrogenase; Mthfs, methylene-H_4F dehydrogenase; Mtk, malate thiokinase; Ppc, phosphoenolpyruvate carboxylase; Pps, phosphoenolpyruvate synthetase; Rpe, ribulose-3-phosphate epimerase; Rpi, ribose-5-phosphate

(*Continued*)

FIGURE 2 | isomerase; Sal, serine aldolase; SdaA, serine deaminase; Sdh, serine, dehydratase; Shmt, serine-H_4F hydroxymethyltransferase; Tal, transaldolase; Tkt1 & Tkt2, transketolase type 1 & 2; Ts, threonine synthase; Ala, alanine; E4P, erythrose-4-phosphate; F6P, fructose-6-phosphate; GAP, glyceraldehyde-3-phosphate; Glu, glutamate; H6P, hexulose-6-phosphate; OAA, oxaloacetate; Oxo-Glu, 2-oxoglutarate; PEP, phosphoenolpyruvate; Pyr, pyruvate; R5P, ribose-5-phosphate; Ru5P, ribulose-5-phosphate; S7P, sedoheptulose-7-phosphate; SBP, sedoheptulose-1,7-biphosphate; Xu5P, xylulose-5-phosphate.

pathway by being converted into formate. Rather than "copy-pasting" the serine cycle, the authors decided to select a set of enzymes to replace key reactions in the natural pathway by other known reactions from other microorganisms to fit the activity of this pathway with *E. coli* metabolism. The *E. coli* metabolic network was modified to become dependent on methanol for growth. Once the operation of the two modules was confirmed (i.e. formate conversion to CH_2H_4F and the modified serine cycle), the complete pathway was tested *in vivo* by using ^{13}C-methanol and xylose. Label derived from methanol was found in intracellular metabolites (Yu et al., 2018). Later, He et al. designed a complete homoserine cycle based on *E. coli* native enzymes and the promiscuous activity of the formaldehyde aldolase (**Figure 2A**). Unlike the serine cycle, here the C1 intermediate for assimilation is formaldehyde. A codon-optimized Mdh from *C. glutamicum* was used to oxidise methanol to formaldehyde. Then, the *E. coli* native metabolic network was modified in order to design a strain in which methanol assimilation is required for serine biosynthesis. The functionality of the homoserine cycle was tested *in vivo* with labelling experiments by using *E. coli* auxotrophic strains in which the complete formaldehyde dissimilation pathway was blocked. However, optimization is still required to achieve *E. coli* growth using the homoserine cycle (He et al., 2020).

The Methanol Condensation Cycle

In 2014, Bogorad et al. designed a non-natural pathway enabling methanol assimilation named methanol condensation cycle (MCC) (**Figure 2C**) (Bogorad et al., 2014). In the MCC pathway, existing enzymes from various microorganisms were combined to convert methanol into acetyl-CoA. These enzymes do not work together in nature. The MCC pathway was first constructed *in silico* and then tested in a cell-free environment. This pathway ensures the conservation of phosphates in the catalytic cycle as well as the balance of the redox state. The MCC was found to be functional and acetyl-CoA production was achieved *in vitro*. Acetyl-CoA is a precursor for the production of biofuels (i.e. ethanol and butanol). However, there is no report of the *in vivo* functionality of the MCC pathway yet.

The Hybrid Mdh-Das Pathway

The hybrid Mdh-Das (HMD) pathway (**Figure 2D**) was built in *E. coli* using an iterative combination of dry and wet lab approaches to design, implement and optimize this metabolic trait (De Simone, et al., 2020). Through *in silico* modelling, a new route that "mixed and matched" two methylotrophic enzymes, a bacterial methanol dehydrogenase (Mdh) and a dihydroxyacetone synthase (Das) from yeast was designed. To identify the best combination of enzymes to introduce into *E. coli*, a library of 266 pathway variants containing different combinations of Mdh and Das homologues was built and then screened using high-throughput ^{13}C-labeling experiments. The highest level of incorporation of methanol into central metabolism intermediates (i.e. 22% into PEP) was obtained using a variant composed of a Mdh from *Acinetobacter gerneri* and a codon-optimized version of *Pichia angusta* Das. Finally, the activity of the HMD pathway was further improved by engineering strategic metabolic targets identified using omics and modelling approaches. The final synthetic strain had 1.5 to 5.9 times higher methanol assimilation in intracellular metabolites and proteinogenic amino acids than the starting strain but was still unable to grow on methanol as the sole carbon and energy source (De Simone, et al., 2020).

A chimeric methanol assimilation pathway was engineered in *Y. lipolytica* by introducing both the RuMP and the XumP pathways (Wang et al., 2021). In this strain, methanol was oxidised by the Mdh from *B. stearothermophilus*. For the RuMP pathway, Hps and Phi from *B. methanolicus* were selected while Das and DhaK from *P. pastoris* were chosen. The formaldehyde dissimilation pathway was blocked by deleting *FLD1* the gene, the homologue of *frmA* in *E. coli*. Recycling of the C1-acceptors, Ru5P and Xu5P, was enhanced by the overexpression of several key genes of the glycolysis and the PPP (i.e. *TKL1*, *PFK*, *FBA*, and *RPE1*) as well as *glpX* from *B. methanolicus* that encodes for a SBPase and a FBPase. After ALE experiments, two evolved strains with improved methanol assimilation were obtained. Interestingly, in one strain, only the expression of the genes of the RuMP pathway was upregulated while in the other one the expression of both RuMP and XuMP pathways was increased. This highlights the capacity of the microbial strain to invoke several strategies in response to externally introduced genetic perturbations.

CREATE NOVEL ENZYMES

Creating new reactions that do not exist in nature is the ultimate strategy to fully explore the metabolic solution space. With the advances in enzyme engineering and *de novo*-enzyme design, it becomes possible to create these reactions and integrate them into a synthetic pathway to perform a novel function. So far, only four studies have implemented synthetic methylotrophic pathways that involve novel catalytic transformations (**Figure 3**).

The Formolase Pathway

Wang et al. were the first in succeeding to implement a linear synthetic pathway. This pathway relies on two enzymes, a NAD-dependent Mdh and a formolase (Fls) (**Figure 3A**). Fls is a synthetic, computationally designed enzyme (Siegel, et al., 2015) that condensates three molecules of formaldehyde to produce one molecule of DHA. The cooperation of the Mdh from *B. methanolicus* MGA3 or PB1 were first tested with Fls

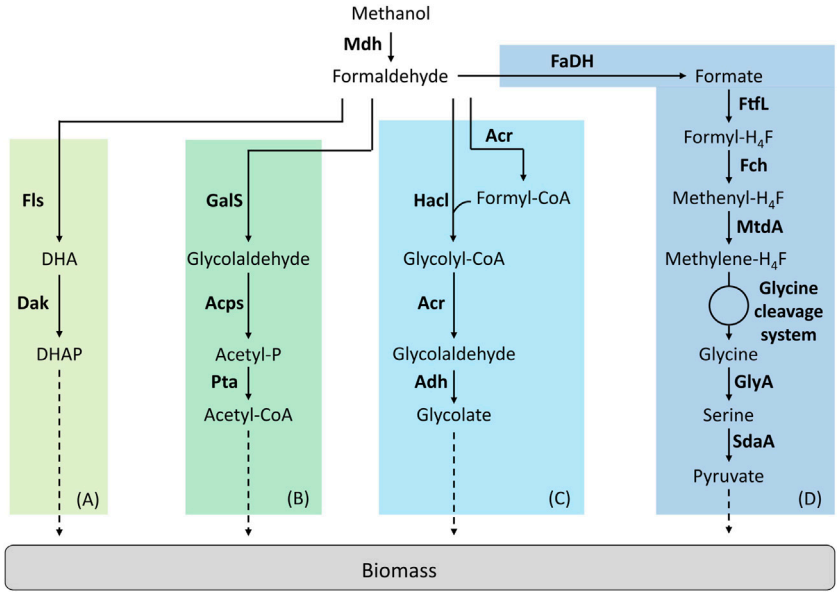


FIGURE 3 | Synthetic linear pathways implemented for C1-assimilation. **(A)** Formolase pathway, **(B)** Synthetic acetyl-CoA pathway, **(C)** 2-hydroxyacyl-CoA lyase pathway, **(D)** Reductive glycine pathway. Adapted from (Tuyishime, et al., 2020). Enzymes are written in bold. Acps, acetylphosphate synthase; Acr, acyl-CoA reductase; Adh, aldehyde dehydrogenase; Dak, dihydroxyacetone kinase; FaDH, formaldehyde dehydrogenase; Fch, methenyl H_4F-cyclohydrolase; Fls, formolase; FtfL, formate-H_4F ligase; GalS, glycolaldehyde synthase; GlyA, serine hydroxymethyltransferase; Hacl, 2-hydroxyacyl-CoA lyase; Mdh, methanol dehydrogenase; MtdA, methylene-H_4F dehydrogenase/methylene-H_4MPT dehydrogenase; Pta, phosphate acetyltransferase; SdaA, serine deaminase. DHA, dihydroxyacetone; DHAP, dihydroxyacetone phosphate.

in vitro. From this experiment, authors decided to introduce the Mdh from *B. methanolicus* PB1 combined with Fls in an *E. coli* strain deleted for *frmA*. To improve the performance of this strain, ALE was used. By using ^{13}C-methanol, the resulting evolved strain was shown to have a higher level of labelling in the proteinogenic amino acids than the parental strain but failed to growth on pure methanol (Wang et al., 2017).

The Synthetic Acetyl-CoA and Glycolaldehyde Assimilation Pathway

The synthetic acetyl-CoA (SACA) is a linear pathway that was designed and constructed by Lu et al. (**Figure 3B**). This pathway was first designed to produce acetyl-CoA from formaldehyde and relies on three enzymes. A glycolaldehyde synthase (GalS) condensates two molecules of formaldehyde to produce one molecule of glycolaldehyde. Then, glycolaldehyde is converted into acetyl-phosphate by the acetyl-phosphate synthase. Acetyl-phosphate is subsequently used to produce acetyl-CoA by the phosphate acetyltransferase. The authors first designed GalS and checked for glycolaldehyde production. Then, they optimized GalS by using directed evolution to improve the kinetic properties of the enzyme. The SACA pathway was functional *in vitro* and the authors then decided to implement the pathway in *E. coli*. The functionality of the pathway was tested *in vivo* in rich media. By combining the SACA pathway with the Mdh from *B. stearothermophilus*, a slight improvement in the final OD was observed. When tested on minimal media by using ^{13}C-methanol, label was also found in some proteinogenic amino acids after 26 h but no growth was observed (Lu et al., 2019).

Based on a computational analysis of metabolic reactions from MetaCyc and Atlas databases, Yang et al. designed a new formaldehyde assimilation pathway named the glycolaldehyde assimilation (GAA) pathway (Yang et al., 2019). This pathway was tested *in vitro* using engineered versions of GalS and transaldolase B and reached a carbon yield of 88%. However, no proof of *in vivo* functionality of this pathway was shown. More recently a new GalS was engineered based on a newly discovered glyoxylate carboligase enzyme found in *E. coli* (Jo et al., 2022). One of the variants showed a 10 times higher affinity and a 2 times higher catalytic efficiency for formaldehyde compared to the previously described GalS (Lu, et al., 2019; Yang, et al., 2019). When this optimized GalS was associated *in vitro* to a lactaldehyde reductase (FucO), 66% of the formaldehyde was converted into ethylene glycol (Jo, et al., 2022).

The 2-Hydroxyacyl-CoA Lyase Pathway

The 2-hydroxyacyl-CoA lyase (Hacl) is a synthetic pathway designed by Chou et al. where Hacl condenses formaldehyde with formyl-CoA to produce glycolyl-CoA (**Figure 3C**). Glycolyl-CoA is subsequently converted to glycolaldehyde by an acyl-CoA reductase (Acr). Finally, a aldehyde dehydrogenase produces glycolate from glycolaldehyde. Hacl is a mammalian enzyme involved in α-oxidation. However, after a BLAST research limited to prokaryotes, the authors identified and tested one Hacl from *Rhodospiralleles bacterium* (RuHacl) exhibiting the condensation activity of formaldehyde with formyl-CoA. They

tested the pathway *in vivo* in an *E. coli* strain deleted for Δ*frmA*, Δ*fdhF*Δ*fdnG*Δ*fdoG* and Δ*glcD* to avoid any competitive reactions using formaldehyde, formate or gluconate. The strain produced 0.5 g/L of glycolate corresponding to a yield of 67%. After enzyme engineering using directed evolution, a variant of RuHacl with improved enzyme kinetics was selected. By expressing this variant, *E. coli* produced 1.2 g/L of glycolate. However, the pathway has not been tested yet *in vitro* or *in vivo* on methanol with the addition of a Mdh (Chou et al., 2019). Later, the Hacl pathway was used to establish the formyl-CoA elongation (FORCE) pathway (Chou et al., 2021). The FORCE pathway was designed as an orthogonal platform for C1 utilization. Based on thermodynamics and stoichiometric analyses, different FORCE pathways were evaluated. Several were tested both *in vitro* and *in vivo* in *E. coli*. The conversion of formate, formaldehyde and methanol into glycolate, ethylene glycol and glycerate was demonstrated among other products.

It was shown that Hacls were poorly produced in *E. coli*, which could limit its efficiency (Chou, et al., 2019; Burgener et al., 2020). Together with Hacl, Oxalyl-CoA decarboxylase (Oxc) are members of a family enzyme catalyzing the condensation of formyl-CoA with formaldehyde to produce glycolyl-CoA. Oxc was thus repurposed to improve its glycolyl-CoA synthase activity under physiological conditions (Nattermann et al., 2021). Oxc from *Methylorubrum extorquens* (MeOxc) was subjected to several rounds of iterative site mutagenesis. A quadruple variant MeOxc4 was selected showing affinity towards formaldehyde similar to natural formaldehyde converting-enzymes and with an improvement for the carboligation activity of 200-fold compared to wild-type MeOxc. To test this enzyme *in vivo*, RuHacl was replaced with MeOxc4 in the *E. coli* strain previously designed for whole-cell bioconversion (Chou, et al., 2019; Nattermann, et al., 2021). Glycolate production was two-fold higher using MeOxc4 than RuHacl. With higher production rates in *E. coli*, this enzyme offers great possibilities for the engineering of new linear C1 pathways.

The Reductive Glycine Pathway

Kim et al. designed a fourth linear pathway: the reductive glycine pathway which, combined with Mdh, enabled methanol assimilation in *E. coli* (**Figure 3D**). In this pathway, formate is the key C1-intermediate that enables carbon assimilation (Kim, et al., 2020). The authors introduced the H_4F-pathway from *Methylobacter extorquens* associated with a glycine cleavage system in an *E. coli* strain auxotrophic for serine, glycine and C1 moieties (Δ*serA*Δ*kbl*Δ*ltaE*Δ*aceA*). It has been previously described that even if growth could not be supported on formate, the latter compound could be still assimilated in *E. coli* via the H_4F-dependent pathway (Tashiro et al., 2018). Kim et al. optimized the operation of this pathway first on formate. Then, the authors optimized the strain using ALE before implementing Mdh. The selected strain was able to grow on formate and CO_2 at a growth rate of $\mu = 0.086\ h^{-1}$. In this strain, methanol to be assimilated is first oxidized to formaldehyde by Mdh and then formaldehyde is oxidized to formate *via* the native GSH-dependent pathway of *E. coli* encoded by the operon *frmRAB*. Several Mdhs were tested and only the Mdh from *B. stearothermophilus* supported growth on methanol and CO_2 at a growth rate of $\mu = 0.013\ h^{-1}$. To confirm that the methanol pathway was functional, they confirmed that a *frmA* deletion blocked growth on methanol. By using ^{13}C-methanol, authors showed that methanol carbon atoms were recovered in proteinogenic amino acids. The labelling pattern found was the same than when the strain grew on ^{13}C-formate and $^{12}CO_2$.

SUMMARY AND PERSPECTIVES

Among all the attempts to engineer synthetic methylotrophy in the platform host microorganisms cited above, it is very interesting to notice that synthetic methylotrophic growth was only achieved in *E. coli* and using both a natural cyclic pathway (Chen, et al., 2020) and a synthetic linear methylotrophic pathway (Kim, et al., 2020). In both examples, rational and evolutionary approaches have been combined. Even if different NAD-dependent Mdhs were used, both are exhibiting improved kinetic parameters for methanol oxidation (i.e. engineered version of Mdh from *C. necator* (Chen, et al., 2020) and Mdh from *B. stearothermophilus* (Kim, et al., 2020)). Moreover, in both cases, auxotrophic strains were first designed prior to applying ALE, which has been determinant to reach the final phenotype. Indeed, when comparing the flux distribution within the metabolic network of the native methylotroph *B. methanolicus* in both methylotrophic and non-methylotrophic conditions (Delepine et al., 2020), we quickly realized that the number of metabolic adaptations required to enable growth on methanol is too high to be solely achieved by a targeted engineering approach. ALE thus appears as the most reasonable strategy to achieve the complete reorganisation of the central metabolism required to fit a methylotrophic lifestyle. In all these studies (Chen, et al., 2018; Meyer, et al., 2018; Bennett, et al., 2020b; Chen, et al., 2020; Keller, et al., 2020), ALE helped in rerouting carbon fluxes towards methanol assimilation by limiting competing pathways and improving substrate uptakes in order to enhance efficient biomass production while adjusting the energy and redox state of the cell. Nevertheless, sometimes rational engineering was necessary in between two rounds of ALE to obtain a fully methylotrophic lifestyle (Chen, et al., 2020). The length of ALE varied according to the study but at least 6 months were necessary to select a strain growing solely on methanol. When looking at the genetic adaptations level, ALE revealed that some are very specific of the co-assimilated carbon source used while others are more general, e.g. the mutations in the genes involved in redox homeostasis or formaldehyde detoxification. It would be very interesting to reconstruct a strain with those mutations in order to identify the minimal set of mutations required for building a true synthetic methylotroph.

However, despite these encouraging accomplishments, synthetic methylotrophy remains a major challenge as the observed growth rates are largely sub-optimal and can barely be considered as active growth. Doubling times of 50 h on a mixture of methanol-CO_2 and 9 h on methanol (corresponding to growth rates of 0.01 and 0.08 h^{-1}, respectively) are far from the

performance of natural methylotrophs that is 0.20 h^{-1} for *Bacillus methanollicus* at 37°C (Müller et al., 2014); 0.10–0.15 h^{-1} for *Pichia Pastoris* at 25°C (Tomàs-Gamisans et al., 2018); and 0.17 h^{-1} for *M. extorquens* at 28°C (Kiefer et al., 2009). Both strains are not yet fully adapted to these non-native carbon sources and further work is necessary. So far, efforts to engineer this non-native substrate catabolism have all taken the straightforward approach of metabolic pathway overexpression, ignoring coordination with the overall cellular processes that include central metabolism, stress-response and cell doubling. However, natural systems use genome-scale regulatory networks, called regulons, which coordinate nutrient catabolism with the larger cellular infrastructure. Therefore, nutrient metabolism cannot be considered as a hermetic process and engineering non-native carbon source utilisation should promote regulation processes together with metabolic pathway diversification and improvement for an efficient and dynamic cellular coordination. Indeed, recent studies demonstrated that targeting regulation of amino acid biosynthesis, an essential process to biomass formation and which is negatively impacted when *E. coli* grows on methanol, can improve methanol assimilation levels (Bennett et al., 2020a; Bennett et al., 2021).

In addition, synthetic biologists have left out the exploration of the spatial dimension for metabolic engineering in prokaryotes, as they do not have any particular subcellular organization. However, a study revealed that around 17% of bacteria contain a bacterial microcompartment (BMC) locus in their genome and that microcompartments are found in 23 different phyla (Axen et al., 2014) including *E. coli*. In addition, compartmentalization of the methanol assimilation pathway in peroxisomes has proven to be efficient in methylotrophic yeasts to protect themselves against toxicity of reactive intermediates (i.e. formaldehyde and H_2O_2) while improving the reaction efficiencies as enzymes are in close vicinity to their substrates and intermediates (van der Klei et al., 2006). Furthermore, compartmentalization allows to build an orthogonal network structure operating with minimal interactions to native metabolic and regulatory networks. Finally, in literature, tools to repurpose BMCs are described (Lau et al., 2018; Lee et al., 2019). For all these reasons, compartmentalizing the methanol assimilatory pathways seems to be an approach worthy of consideration to establish synthetic methylotrophy in prokaryotes.

Changing a lifestyle from heterotrophic to methylotrophic represents major engineering that have been partially achieved. In 4–10 years from now, it is likely that the next generation of synthetic methylotrophs will be further developed for the production of chemicals (e.g. bulk chemicals, fuels) from methanol, after which commercial production may be feasible. However, there is undoubtedly a long way ahead to achieve a synthetic strain able to use methanol for the production of both biomass and chemicals.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

CMV was supported by the PostAgreenskills Fund from INRAE.

REFERENCES

Axen, S. D., Erbilgin, O., and Kerfeld, C. A. (2014). A Taxonomy of Bacterial Microcompartment Loci Constructed by a Novel Scoring Method. *PLoS Comput. Biol.* 10, e1003898. doi:10.1371/journal.pcbi.1003898

Becker, J., and Wittmann, C. (2015). Advanced Biotechnology: Metabolically Engineered Cells for the Bio-Based Production of Chemicals and Fuels, Materials, and Health-Care Products. *Angew. Chem. Int. Ed.* 54, 3328–3350. doi:10.1002/anie.201409033

Bennett, R. K., Agee, A., Gerald Har, J. R., von Hagel, B., Siu, K.-H., Antoniewicz, M. R., et al. (2020a). Triggering the Stringent Response Enhances Synthetic Methanol Utilization in *Escherichia coli*. *Metab. Eng.* 61, 1–10. doi:10.1016/j.ymben.2020.04.007

Bennett, R. K., Dillon, M., Gerald Har, J. R., Agee, A., von Hagel, B., Rohlhill, J., et al. (2020b). Engineering *Escherichia coli* for Methanol-dependent Growth on Glucose for Metabolite Production. *Metab. Eng.* 60, 45–55. doi:10.1016/j.ymben.2020.03.003

Bennett, R. K., Gonzalez, J. E., Whitaker, W. B., Antoniewicz, M. R., and Papoutsakis, E. T. (2018). Expression of Heterologous Non-oxidative Pentose Phosphate Pathway from *Bacillus Methanolicus* and Phosphoglucose Isomerase Deletion Improves Methanol Assimilation and Metabolite Production by a Synthetic *Escherichia coli* Methylotroph. *Metab. Eng.* 45, 75–85. doi:10.1016/j.ymben.2017.11.016

Bogorad, I. W., Chen, C.-T., Theisen, M. K., Wu, T.-Y., Schlenz, A. R., Lam, A. T., et al. (2014). Building Carbon-Carbon Bonds Using a Biocatalytic Methanol Condensation Cycle. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15928–15933. doi:10.1073/pnas.1413470111

Brautaset, T., Jakobsen, Ø. M., Flickinger, M. C., Valla, S., and Ellingsen, T. E. (2004). Plasmid-dependent Methylotrophy in Thermotolerant Bacillus Methanolicus. *J. Bacteriol.* 186, 1229–1238. doi:10.1128/jb.186.5.1229-1238.2004

Burgener, S., Cortina, N. S., and Erb, T. J. (2020). Oxalyl-CoA Decarboxylase Enables Nucleophilic One-Carbon Extension of Aldehydes to Chiral α-Hydroxy Acids. *Angew. Chem. Int. Ed.* 59, 5526–5530. doi:10.1002/anie.201915155

Chen, C. T., Chen, F. Y., Bogorad, I. W., Wu, T. Y., Zhang, R., Lee, A. S., et al. (2018). Synthetic Methanol Auxotrophy of *Escherichia coli* for Methanol-dependent Growth and Production. *Metab. Eng.* 49, 257–266. doi:10.1016/j.ymben.2018.08.010

Chen, F. Y.-H., Jung, H.-W., Tsuei, C.-Y., and Liao, J. C. (2020). Converting *Escherichia coli* to a Synthetic Methylotroph Growing Solely on Methanol. *Cell.* 182, 933–946. doi:10.1016/j.cell.2020.07.010

Chou, A., Clomburg, J. M., Qian, S., and Gonzalez, R. (2019). 2-Hydroxyacyl-CoA Lyase Catalyzes Acyloin Condensation for One-Carbon Bioconversion. *Nat. Chem. Biol.* 15, 900–906. doi:10.1038/s41589-019-0328-0

Chou, A., Lee, S. H., Zhu, F., Clomburg, J. M., and Gonzalez, R. (2021). An Orthogonal Metabolic Framework for One-Carbon Utilization. *Nat. Metab.* 20213, 1385–1399. doi:10.1038/s42255-021-00453-0

Dai, Z., Gu, H., Zhang, S., Xin, F., Zhang, W., Dong, W., et al. (2017). Metabolic Construction Strategies for Direct Methanol Utilization in Saccharomyces cerevisiae, *Bioresour. Technol.*, 245, 1407–1412. doi:10.1016/j.biortech.2017.05.100

De Simone, A., Vicente, C. M., Peiro, C., Gales, L., Bellvert, F., Enjalbert, B., et al. (2020). Mixing and Matching Methylotrophic Enzymes to Design a Novel Methanol Utilization Pathway in *E. coli*. *Metab. Eng.* 61, 315–325. doi:10.1016/j.ymben.2020.07.005

Delépine, B., López, M. G., Carnicer, M., Vicente, C. M., Wendisch, V. F., and Heux, S. (2020). Charting the Metabolic Landscape of the Facultative Methylotroph Bacillus Methanolicus. *mSystems* 5, 5. doi:10.1128/mSystems.00745-20

Denby, K. J., Iwig, J., Bisson, C., Westwood, J., Rolfe, M. D., Sedelnikova, S. E., et al. (2016). The Mechanism of a Formaldehyde-Sensing Transcriptional Regulator. *Sci. Rep.* 6, 38879. doi:10.1038/srep38879

Dragosits, M., and Mattanovich, D. (2013). Adaptive Laboratory Evolution - Principles and Applications for Biotechnology. *Microb. Cell. Fact.* 12, 64. doi:10.1186/1475-2859-12-64

Dubey, A. A., and Jain, V. (2019). Mycofactocin Is Essential for the Establishment of Methylotrophy in Mycobacterium Smegmatis. *Biochem. Biophysical Res. Commun.* 516, 1073–1077. doi:10.1016/j.bbrc.2019.07.008

Erb, T. J., Jones, P. R., and Bar-Even, A. (2017). Synthetic Metabolism: Metabolic Engineering Meets Enzyme Design. *Curr. Opin. Chem. Biol.* 37, 56–62. doi:10.1016/j.cbpa.2016.12.023

Espinosa, M. I., Gonzalez-Garcia, R. A., Valgepea, K., Plan, M. R., Scott, C., Pretorius, I. S., et al. (2020). Adaptive Laboratory Evolution of Native Methanol Assimilation in *Saccharomyces cerevisiae*. *Nat. Commun.* 11, 5564. doi:10.1038/s41467-020-19390-9

Fan, L., Wang, Y., Qian, J., Gao, N., Zhang, Z., Ni, X., et al. (2021). Transcriptome Analysis Reveals the Roles of Nitrogen Metabolism and Sedoheptulose Bisphosphatase Pathway in Methanol-dependent Growth of Corynebacterium Glutamicum. *Microb. Biotechnol.* 14, 1797–1808. doi:10.1111/1751-7915.13863

Fan, L., Wang, Y., Tuyishime, P., Gao, N., Li, Q., Zheng, P., et al. (2018). Engineering Artificial Fusion Proteins for Enhanced Methanol Bioconversion. *ChemBioChem* 19, 2465–2471. doi:10.1002/cbic.201800424

Ganesh, I., Vidhya, S., Eom, G. T., and Hong, S. H. (2017). Construction of Methanol-Sensing *Escherichia coli* by the Introduction of a Paracoccus Denitrificans MxaY-Based Chimeric Two-Component System. *J. Microbiol. Biotechnol.* 27, 1106–1111. doi:10.4014/jmb.1611.11070

Gao, B., Zhao, N., Deng, J., Gu, Y., Jia, S., Hou, Y., et al. (2022). Constructing a Methanol-dependent Bacillus Subtilis by Engineering the Methanol Metabolism. *J. Biotechnol.* 343, 128–137. doi:10.1016/j.jbiotec.2021.12.005

He, H., Höper, R., Dodenhöft, M., Marlière, P., and Bar-Even, A. (2020). An Optimized Methanol Assimilation Pathway Relying on Promiscuous Formaldehyde-Condensing Aldolases in *E. coli*. *Metab. Eng.* 60, 1–13. doi:10.1016/j.ymben.2020.03.002

Hennig, G., Haupka, C., Brito, L. F., Rückert, C., Cahoreau, E., Heux, S., et al. (2020). Methanol-Essential Growth of Corynebacterium Glutamicum: Adaptive Laboratory Evolution Overcomes Limitation Due to Methanethiol Assimilation Pathway. *Ijms* 21, 3617. doi:10.3390/ijms21103617

Heux, S., Brautaset, T., Vorholt, J. A., Wendisch, V. F., and Portais, J. C. (2018). "Synthetic Methylotrophy: Past, Present, and Future," in *Methane Biocatalysis: Paving the Way to Sustainability* (Cham: Springer), 133–151. doi:10.1007/978-3-319-74866-5_9

Jakobsen, Ø. M., Benichou, A., Flickinger, M. C., Valla, S., Ellingsen, T. E., and Brautaset, T. (2006). Upregulated Transcription of Plasmid and Chromosomal Ribulose Monophosphate Pathway Genes Is Critical for Methanol Assimilation Rate and Methanol Tolerance in the Methylotrophic Bacterium Bacillus Methanolicus. *J. Bacteriol.* 188, 3063–3072. doi:10.1128/jb.188.8.3063-3072.2006

Jo, H.-J., Kim, J.-H., Kim, Y.-N., Seo, P.-W., Kim, C.-Y., Kim, J.-W., et al. (2022). Glyoxylate Carboligase-Based Whole-Cell Biotransformation of Formaldehyde into Ethylene Glycol via Glycolaldehyde. *Green Chem.* 24, 218–226. doi:10.1039/D1GC03549E

Keller, P., Noor, E., Meyer, F., Reiter, M. A., Anastassov, S., Kiefer, P., et al. (2020). Methanol-dependent *Escherichia coli* Strains with a Complete Ribulose Monophosphate Cycle. *Nat. Commun.* 11, 5403. doi:10.1038/s41467-020-19235-5

Keltjens, J. T., Pol, A., Reimann, J., and Op den Camp, H. J. M. (2014). PQQ-dependent Methanol Dehydrogenases: Rare-Earth Elements Make a Difference. *Appl. Microbiol. Biotechnol.* 98, 6163–6183. doi:10.1007/s00253-014-5766-8

Kiefer, P., Buchhaupt, M., Christen, P., Kaup, B., Schrader, J., and Vorholt, J. A. (2009). Metabolite Profiling Uncovers Plasmid-Induced Cobalt Limitation under Methylotrophic Growth Conditions. *PLoS One* 4, e7831. doi:10.1371/journal.pone.0007831

Kim, S., Lindner, S. N., Aslan, S., Yishai, O., Wenk, S., Schann, K., et al. (2020). Growth of *E. coli* on Formate and Methanol via the Reductive glycine Pathway. *Nat. Chem. Biol.* 16, 538–545. doi:10.1038/s41589-020-0473-5

Koopman, F. W., de Winde, J. H., and Ruijssenaars, H. J. (2009). C1 Compounds as Auxiliary Substrate for Engineered Pseudomonas Putida S12. *Appl. Microbiol. Biotechnol.* 83, 705–713. doi:10.1007/s00253-009-1922-y

Kyle Bennett, R., Agee, A., Har, J. R. G., Hagel, B., Antoniewicz, M. R., and Papoutsakis, E. T. (2021). Regulatory Interventions Improve the Biosynthesis of Limiting Amino Acids from Methanol Carbon to Improve Synthetic Methylotrophy in *Escherichia coli*. *Biotechnol. Bioeng.* 118, 43–57. doi:10.1002/bit.27549

Lau, Y. H., Giessen, T. W., Altenburg, W. J., and Silver, P. A. (2018). Prokaryotic Nanocompartments Form Synthetic Organelles in a Eukaryote. *Nat. Commun.* 9, 1311. doi:10.1038/s41467-018-03768-x

Lee, M. J., Palmer, D. J., and Warren, M. J. (2019). Biotechnological Advances in Bacterial Microcompartment Technology. *Trends Biotechnol.* 37, 325–336. doi:10.1016/j.tibtech.2018.08.006

Lessmeier, L., Pfeifenschneider, J., Carnicer, M., Heux, S., Portais, J.-C., and Wendisch, V. F. (2015). Production of Carbon-13-Labeled Cadaverine by Engineered Corynebacterium Glutamicum Using Carbon-13-Labeled Methanol as Co-substrate. *Appl. Microbiol. Biotechnol.* 99, 10163–10176. doi:10.1007/s00253-015-6906-5

Lu, X., Liu, Y., Yang, Y., Wang, S., Wang, Q., Wang, X., et al. (2019). Constructing a Synthetic Pathway for Acetyl-Coenzyme A from One-Carbon through Enzyme Design. *Nat. Commun.* 10, 1378. doi:10.1038/s41467-019-09095-z

Meyer, F., Keller, P., Hartl, J., Gröninger, O. G., Kiefer, P., and Vorholt, J. A. (2018). Methanol-essential Growth of *Escherichia coli*. *Nat. Commun.* 9, 1508. doi:10.1038/s41467-018-03937-y

Müller, J. E. N., Litsanov, B., Bortfeld-Miller, M., Trachsel, C., Grossmann, J., Brautaset, T., et al. (2014). Proteomic Analysis of the Thermophilic Methylotroph Bacillus Methanolicus MGA3. *Proteomics* 14, 725–737. doi:10.1002/pmic.201300515

Müller, J. E. N., Meyer, F., Litsanov, B., Kiefer, P., Potthoff, E., Heux, S., et al. (2015). Engineering *Escherichia coli* for Methanol Conversion. *Metab. Eng.* 28, 190–201. doi:10.1016/j.ymben.2014.12.008

Nattermann, M., Burgener, S., Pfister, P., Chou, A., Schulz, L., Lee, S. H., et al. (2021). Engineering a Highly Efficient Carboligase for Synthetic One-Carbon Metabolism. *ACS Catal.* 11, 5396–5404. doi:10.1021/acscatal.1c01237

Orita, I., Sakamoto, N., Kato, N., Yurimoto, H., and Sakai, Y. (2007). Bifunctional Enzyme Fusion of 3-Hexulose-6-Phosphate Synthase and 6-Phospho-3-Hexuloisomerase. *Appl. Microbiol. Biotechnol.* 76, 439–445. doi:10.1007/s00253-007-1023-8

Price, J. V., Chen, L., Whitaker, W. B., Papoutsakis, E., and Chen, W. (2016). Scaffoldless Engineered Enzyme Assembly for Enhanced Methanol Utilization. *Proc. Natl. Acad. Sci. U.S.A.* 113, 12691–12696. doi:10.1073/pnas.1601797113

Rohlhill, J., Gerald Har, J. R., Antoniewicz, M. R., and Papoutsakis, E. T. (2020). Improving Synthetic Methylotrophy via Dynamic Formaldehyde Regulation of Pentose Phosphate Pathway Genes and Redox Perturbation. *Metab. Eng.* 57, 247–255. doi:10.1016/j.ymben.2019.12.006

Rohlhill, J., Sandoval, N. R., and Papoutsakis, E. T. (2017). Sort-Seq Approach to Engineering a Formaldehyde-Inducible Promoter for Dynamically Regulated *Escherichia coli* Growth on Methanol. *ACS Synth. Biol.* 6, 1584–1595. doi:10.1021/acssynbio.7b00114

Roth, T. B., Woolston, B. M., Stephanopoulos, G., and Liu, D. R. (2019). Phage-Assisted Evolution of Bacillus Methanolicus Methanol Dehydrogenase 2. *ACS Synth. Biol.* 8, 796–806. doi:10.1021/acssynbio.8b00481

Sheehan, M. C., Bailey, C. J., Dowds, B. C. A., and McConnell, D. J. (1988). A New Alcohol Dehydrogenase, Reactive towards Methanol, from Bacillus Stearothermophilus. *Biochem. J.* 252, 661–666. doi:10.1042/bj2520661

Siegel, J. B., Smith, A. L., Poust, S., Wargacki, A. J., Bar-Even, A., Louw, C., et al. (2015). Computational Protein Design Enables a Novel One-Carbon Assimilation Pathway. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3704–3709. doi:10.1073/pnas.1500545112

Tashiro, Y., Hirano, S., Matson, M. M., Atsumi, S., and Kondo, A. (2018). Electrical-biological Hybrid System for CO2 Reduction. *Metab. Eng.* 47, 211–218. doi:10.1016/j.ymben.2018.03.015

Tomàs-Gamisans, M., Ferrer, P., and Albiol, J. (2018). Fine-tuning the P. Pastoris iMT1026 Genome-Scale Metabolic Model for Improved Prediction of Growth

on Methanol or Glycerol as Sole Carbon Sources. *Microb. Biotechnol.* 11, 224–237. doi:10.1111/1751-7915.12871

Tuyishime, P., and Sinumvayo, J. P. (2020). Novel Outlook in Engineering Synthetic Methylotrophs and Formatotrophs: a Course for Advancing C1-Based Chemicals Production. *World J. Microbiol. Biotechnol.* 36, 118. doi:10.1007/s11274-020-02899-y

Tuyishime, P., Wang, Y., Fan, L., Zhang, Q., Li, Q., Zheng, P., et al. (2018). Engineering Corynebacterium Glutamicum for Methanol-dependent Growth and Glutamate Production. *Metab. Eng.* 49, 220–231. doi:10.1016/j.ymben.2018.07.011

van der Klei, I. J., Yurimoto, H., Sakai, Y., and Veenhuis, M. (2006). The Significance of Peroxisomes in Methanol Metabolism in Methylotrophic Yeast. *Biochimica Biophysica Acta (BBA) - Mol. Cell. Res.* 1763, 1453–1462. doi:10.1016/j.bbamcr.2006.07.016

Vartiainen, E., Blomberg, P., Ilmén, M., Andberg, M., Toivari, M., and Penttilä, M. (2019). Evaluation of Synthetic Formaldehyde and Methanol Assimilation Pathways in Yarrowia Lipolytica. *Fungal Biol. Biotechnol.* 6, 27. doi:10.1186/s40694-019-0090-9

Wang, G., Olofsson-Dolk, M., Hansson, F. G., Donati, S., Li, X., Chang, H., et al. (2021). Engineering Yeast Yarrowia Lipolytica for Methanol Assimilation. *ACS Synth. Biol.* 10, 3537–3550. doi:10.1021/acssynbio.1c00464

Wang, X., Wang, Y., Liu, J., Li, Q., Zhang, Z., Zheng, P., et al. (2017). Biological Conversion of Methanol by Evolved *Escherichia coli* Carrying a Linear Methanol Assimilation Pathway. *Bioresour. Bioprocess.* 4, 41. doi:10.1186/s40643-017-0172-6

Wang, Y., Fan, L., Tuyishime, P., Liu, J., Zhang, K., Gao, N., et al. (2020a). Adaptive Laboratory Evolution Enhances Methanol Tolerance and Conversion in Engineered Corynebacterium Glutamicum. *Commun. Biol.* 3, 217. doi:10.1038/s42003-020-0954-9

Wang, Y., Fan, L., Tuyishime, P., Zheng, P., and Sun, J. (2020b). Synthetic Methylotrophy: A Practical Solution for Methanol-Based Biomanufacturing. *Trends Biotechnol.* 38, 650–666. doi:10.1016/j.tibtech.2019.12.013

Wendisch, V. F., Kosec, G., Heux, S., and Brautaset, T. (2021). Aerobic Utilization of Methanol for Microbial Growth and Production. *Adv. Biochem. Eng. Biotechnol. Nov.* 11. doi:10.1007/10_2021_177

Whitaker, W. B., Jones, J. A., Bennett, R. K., Gonzalez, J. E., Vernacchio, V. R., Collins, S. M., et al. (2017). Engineering the Biological Conversion of Methanol to Specialty Chemicals in *Escherichia coli*. *Metab. Eng.* 39, 49–59. doi:10.1016/j.ymben.2016.10.015

Witthoff, S., Schmitz, K., Niedenführ, S., Nöh, K., Noack, S., Bott, M., et al. (2015). Metabolic Engineering of Corynebacterium Glutamicum for Methanol Metabolism. *Appl. Environ. Microbiol.* 81, 2215–2225. doi:10.1128/AEM.03110-14

Woolston, B. M., King, J. R., Reiter, M., Van Hove, B., and Stephanopoulos, G. (2018). Improving Formaldehyde Consumption Drives Methanol Assimilation in Engineered *E. coli*. *Nat. Commun.* 9, 2387. doi:10.1038/s41467-018-04795-4

Wu, T.-Y., Chen, C.-T., Liu, J. T.-J., Bogorad, I. W., Damoiseaux, R., and Liao, J. C. (2016). Characterization and Evolution of an Activator-independent Methanol Dehydrogenase from Cupriavidus Necator N-1. *Appl. Microbiol. Biotechnol.* 100, 4969–4983. doi:10.1007/s00253-016-7320-3

Yang, X., Yuan, Q., Luo, H., Li, F., Mao, Y., Zhao, X., et al. (2019). Systematic Design and *In Vitro* Validation of Novel One-Carbon Assimilation Pathways. *Metab. Eng.* 56, 142–153. doi:10.1016/j.ymben.2019.09.001

Yasueda, H., Kawahara, Y., and Sugimoto, S.-i. (1999). Bacillus Subtilis yckG and yckF Encode Two Key Enzymes of the Ribulose Monophosphate Pathway Used by Methylotrophs, and yckH Is Required for Their Expression. *J. Bacteriol.* 181, 7154–7160. doi:10.1128/jb.181.23.7154-7160.1999

Yu, H., and Liao, J. C. (2018). A Modified Serine Cycle in *Escherichia coli* Coverts Methanol and CO2 to Two-Carbon Compounds. *Nat. Commun.* 9 (9), 3992. doi:10.1038/s41467-018-06496-4

Yurimoto, H., Kato, N., and Sakai, Y. (2009). Genomic Organization and Biochemistry of the Ribulose Monophosphate Pathway and its Application in Biotechnology. *Appl. Microbiol. Biotechnol.* 84, 407–416. doi:10.1007/s00253-009-2120-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 08 August 2022
doi: 10.3389/fbioe.2022.869111

Nanocellulose Composites as Smart Devices With Chassis, Light-Directed DNA Storage, Engineered Electronic Properties, and Chip Integration

*Elena Bencurova[1†], Sergey Shityakov[2†], Dominik Schaack[1], Martin Kaltdorf[1], Edita Sarukhanyan[1], Alexander Hilgarth[3], Christin Rath[4], Sergio Montenegro[3], Günter Roth[4,5], Daniel Lopez[6] and Thomas Dandekar[1,7]**

[1]Functional Genomics and Systems Biology Group, Department of Bioinformatics, Biocenter, University of Würzburg, Würzburg, Germany, [2]Laboratory of Chemoinformatics, Infochemistry Scientific Center, ITMO University, Saint Petersburg, Russia, [3]Aerospace Information Technology, University of Würzburg, Würzburg, Germany, [4]Laboratory for Microarray Copying, Center for Biological Systems Analysis (ZBSA), University of Freiburg, Freiburg, Germany, [5]BioCopy GmbH, Emmendingen, Germany, [6]Centro Nacional de Biotecnologia CNB, Universidad Autonoma de Madrid, Madrid, Spain, [7]Structural and Computational Biology, European Molecular Biology Laboratory, Heidelberg, Germany

OPEN ACCESS

Edited by:
Jean Marie François, Institut Biotechnologique de Toulouse (INSA), France

Reviewed by:
Javier Macia, Pompeu Fabra University, Spain
Wen Ma, University of California, San Diego, United States

***Correspondence:**
Thomas Dandekar
dandekar@biozentrum.uni-wuerzburg.de

[†]These authors share first authorship

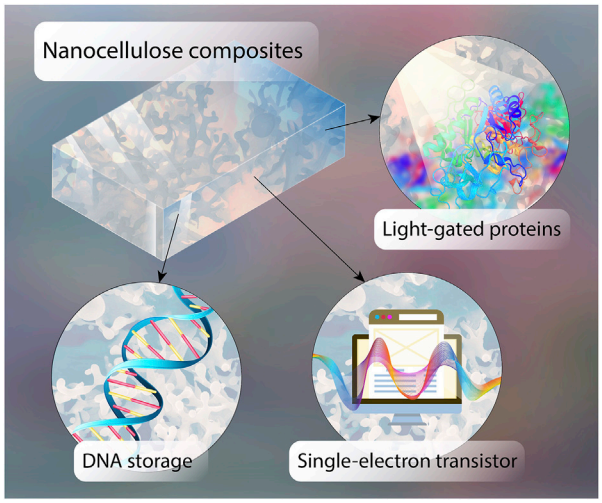
Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *03 February 2022*
Accepted: *24 June 2022*
Published: *08 August 2022*

Citation:
Bencurova E, Shityakov S, Schaack D, Kaltdorf M, Sarukhanyan E, Hilgarth A, Rath C, Montenegro S, Roth G, Lopez D and Dandekar T (2022) Nanocellulose Composites as Smart Devices With Chassis, Light-Directed DNA Storage, Engineered Electronic Properties, and Chip Integration. Front. Bioeng. Biotechnol. 10:869111. doi: 10.3389/fbioe.2022.869111

The rapid development of green and sustainable materials opens up new possibilities in the field of applied research. Such materials include nanocellulose composites that can integrate many components into composites and provide a good chassis for smart devices. In our study, we evaluate four approaches for turning a nanocellulose composite into an information storage or processing device: 1) nanocellulose can be a suitable carrier material and protect information stored in DNA. 2) Nucleotide-processing enzymes (polymerase and exonuclease) can be controlled by light after fusing them with light-gating domains; nucleotide substrate specificity can be changed by mutation or pH change (read-in and read-out of the information). 3) Semiconductors and electronic capabilities can be achieved: we show that nanocellulose is rendered electronic by iodine treatment replacing silicon including microstructures. Nanocellulose semiconductor properties are measured, and the resulting potential including single-electron transistors (SET) and their properties are modeled. Electric current can also be transported by DNA through G-quadruplex DNA molecules; these as well as classical silicon semiconductors can easily be integrated into the nanocellulose composite. 4) To elaborate upon miniaturization and integration for a smart nanocellulose chip device, we demonstrate pH-sensitive dyes in nanocellulose, nanopore creation, and kinase micropatterning on bacterial membranes as well as digital PCR micro-wells. Future application potential includes nano-3D printing and fast molecular processors (e.g., SETs) integrated with DNA storage and conventional electronics. This would also lead to environment-friendly nanocellulose chips for information processing as well as smart nanocellulose composites for biomedical applications and nano-factories.

Keywords: nanocellulose, DNA storage, light-gated proteins, single-electron transistors, protein chip



GRAPHICAL ABSTRACT |

INTRODUCTION

Potential applications using the interaction of nanocellulose with DNA have been investigated for several years. Nanocellulose is a versatile material with several features, such as optical transparency, conductivity, and flexibility. It has various applications, such as packaging material, drug delivery, tissue scaffold, printed electronics, and reinforced polymer composites (Razaq et al., 2011; Thomas et al., 2018; Tao et al., 2020; Jiao et al., 2021) For instance, one approach took advantage of both DNA's structural compatibility with nanocellulose and its inherent ability for molecular recognition *via* base pairing. By attaching ssDNA oligomers to nanocellulose crystals, it is possible for complementary sequences from oligonucleotides to bind to separate cellulose nanocrystals. They pair with each other creating a nanocellulose/DNA hybrid nanomaterial (Mangalam et al., 2009; Habibi, 2014). In addition to that, nanocellulose-based matrices have been successfully used as an ion-exchange membrane for temporary storage of DNA oligonucleotides (Razaq et al., 2011). However, gradually, nanocellulose composites became ever more attractive and environmentally friendly as multipurpose materials used in medicine, food industry, biotechnology, and engineering (Hoeng et al., 2016; Azeredo et al., 2017; Osorio et al., 2019; de Amorim et al., 2020).

In our study, we explore nanocellulose composite as a smart material. This could be, for instance, a chassis for an information storage device applied preferably to natural, fully degradable components. DNA storage has begun to show large storage potential (Church et al., 2012) for preserving various kinds of data (Goldman et al., 2013) over the course of thousands of years (Grass et al., 2015). Recent major developments had been published recently including the DNA fountain (Erlich and Zielinski, 2017), "DNA-of-things" storage architecture (Koch et al., 2020), and image-based DNA storage systems (Cao et al., 2021). On the other hand, the extraction and decoding time is still challenging. It requires 1 to 3 days, depending on the sequencing technique. Hence, apart from clinical applications such as human genetics/patient samples, it has not yet gained such popularity compared to electronic storage. Here, we evaluate previous concepts (Dandekar et al., 2019) in practice: 1) nanocellulose as a chassis with support, protection, and integration for such a smart device and its components. Nanocellulose was chosen because of its sustainability, it is easy to scale up the production, and it has neither negative nor positive effect on the DNA. 2) We introduce light-gated nucleotide-processing enzymes so that the DNA storage can easily be read and retrieved, accessed, and the information content changed (processed). 3) We show that different electronic properties can be achieved in nanocellulose including semiconductivity and single-electron transistors. 4) The key to unlocking the full potential is to achieve synergies such as light-gated synthesis of DNA wires. This permits adaptive changes in the chip layout. We demonstrate that several useful techniques, including nano-structuring, are readily applied to the nanocellulose composite improving its function. Nevertheless, for high performance, there is still a long way of development to go.

TABLE 1 | Translation table for DNA storage.

Letter	Code	Letter	Code	Letter	Code	Letter	Code
A	ACT	H	CGT	O	TGT	V	GTA
B	CAT	I	CTG	P	GAG	W	ATG
C	TCA	J	TGC	Q	TAT	X	AGT
D	TAC	K	TCG	R	CAC	Y	GAC
E	CTA	L	ATC	S	TGA	Z	GCA
F	GCT	M	ACA	T	TAG	space	AGC
G	GTC	N	CTC	U	GAT	.	ACG

MATERIALS AND METHODS

Preparation of Nanocellulose

As the source of the bacterial nanocellulose, a Kombucha membrane (symbiotic culture of bacteria and yeast) was used (a kind gift of Carmen Aquilar, University of Würzburg). The culture was grown in tea infusion comprising 10 g of black tea (Schwarztee Misching, EDEKA, Germany) and 10% of sucrose (AppliChem, Germany) infused for 10 min into 1 L of boiling water. After cooling to room temperature, 50 ml of media from the previous culture was added to reach the favorable pH for the symbiotic culture. In addition, Kombucha was added. The glass bottle was covered with lightweight paper, and the culture was fermented at room temperature for 30 days. The extraction of nanocellulose was performed using alkaline treatment, acid hydrolysis, and blending.

For the alkaline treatment, the brownish culture was washed 15 times for 10 min in the 1 M NaOH solution using an ultrasonic bath (Sonorex Super, Bandelin, Germany). It was neutralized by washing 10 times for 5 min in distilled water until the membrane reached pH 7.5. After the alkaline treatment, the nanocellulose was ground in a mixer (Princess, 1,000 Wat, 23,000 U/min) for 5 × 2 min and homogenized using glass beads. Dried nanocellulose was obtained by keeping a small amount of the nanocellulose in the desiccator for 48 h.

For the pH-sensitive experiment, 10 µl of malachite green (aqueous solution) was added to 0.5 g of nanocellulose and pH was modified by HCl/NaOH.

DNA Storage Experiments

For the DNA storage investigations, the text "University of Wuerzburg: Light-gated polymerase" (**Supplementary Figure S1**) and "Wuerzburg" were encoded into 123 and 26 nt DNA using a DNA writer (Urbano, 2013) and synthesized (Eurofins Genomics, Germany). The encoding schema is given in **Table 1**. Moreover, 500 picomol of oligonucleotide was subjected to dried nanocellulose and after drying was placed in a sterile closed Petri dish. It was left at room temperature for 2 weeks/2 months/2 years. After these time points, samples were rehydrated in sterile water, and DNA was subcloned to TOPO 2.1 vector (Thermo Fisher Scientific, Germany) and amplified by sequencing primers. Amplicons were sequenced by Sanger sequencing (Eurofins Genomics, Germany). Sequencing was performed in triplicates.

Preparation and Analysis of BLUF–GFP Constructs

The *E. coli* strain DH-5α was transformed with a pPK-CMV-F1 vector containing the GFP-encoding gene in the C-terminus (ProKine, Germany) with inserted BLUF domain amplified from the genomic DNA of *E. coli* (forward primer: AAAAAA CTGCAGAGATCTATGCTTACCACCCTTATTTATCGTAGC, reverse primer: TTTTTTGAGCTCTTCGAAGCGAGACAG TAGTATTCAATCGACTTT). The transformed *E. coli* was inoculated to 5 ml of the LB medium with ampicillin and cultured overnight at 37 C at 250 rpm. A volume of 1 ml of overnight culture was transferred to two 500 ml flasks with 50 ml of fresh LB medium containing ampicillin and cultures were grown at 37°C to $O.D_{600}$ 0.5. The first bottle was wrapped completely with aluminum foil, and the second was kept under daylight (as a source of blue light). Immediately, 1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG, Sigma-Aldrich, Germany) was added and incubated for 4 h at 30°C. After expression, the cells were analyzed by fluorescent microscopy. Apart from the BLUF–GFP construct, we also constructed other light-activated nucleotide-processing enzymes for the functional assays: LOV–Taq polymerase, BLUF–adenyltransferase, BLUF–T4 kinase, BLUF–Cid I polymerase (including mutated and wild type version and two different BLUF domain constructs), and LOV 2–adenylate kinase (see **Supplementary Table S1** for available constructs). Sequences are listed in the **Supplementary Material**.

Exonuclease Production and Activity Assay

The BLUF–exonuclease construct was ordered as a synthetic gene (Eurofins Genomics, Germany) and cloned into an expression vector (pQE-30-UA-mCherry-GFP, in-house modified plasmid from Qiagen, United States). The protein was expressed as described above. Purification was performed using the Ni-NTA resin (Merck, Germany), following the manufacturer's instructions. For the activity assay, 1 µM of Cy-5 hexamer (AAAAAA) was mixed with 10 U of exonuclease I (NEB, Germany) as the positive control, or 1 µl of BLUF–exonuclease I. The mixture was incubated for 30 min at 37°C and then heat-inactivated. Samples were mixed with TBE-Urea sample buffer (Thermo Fisher Scientific, Germany) and resolved on 10% TBE-Urea gel. The gel was scanned by Odyssey (LI-COR, Germany).

CidI Polymerase Docking

The CidI molecule (PDB: 4FH5 and 4FHX) and ligands (adenine and uracil) were energy-minimized before docking with the help of the Molecular Operating Environment (MOE) software [Molecular Operating Environment (MOE), 2016]. This was carried out with the MMFF94 (Merck Molecular) force field. Protein structure refinement, as well as ligand library preparation, was carried out with the tools of the same software. Molecular docking simulations were performed using GOLD (Jones et al., 1997), MOE (Molecular Operating Environment (MOE), 2016), and AutoDock (Morris et al., 2009).

The AutoDock 4.2.6 (Morris, et al., 2009) software was obtained from the site of "The Scripps Research Institute" (http://autodock.scripps.edu/downloads/autodock-registration/autodock-4-2-download-page/) to perform molecular docking simulations.

The structure of polymerase was protonated, and the rotatable bonds for the ligands were clearly defined. The dimension of the grid box was set as 50 Å × 50 Å × 50 Å for all the docking simulations along with spacing of 0.375 Å. The center of the grid box was placed so that it involved the residues of the active site of the protein and coincided with the center of ligand in the active site.

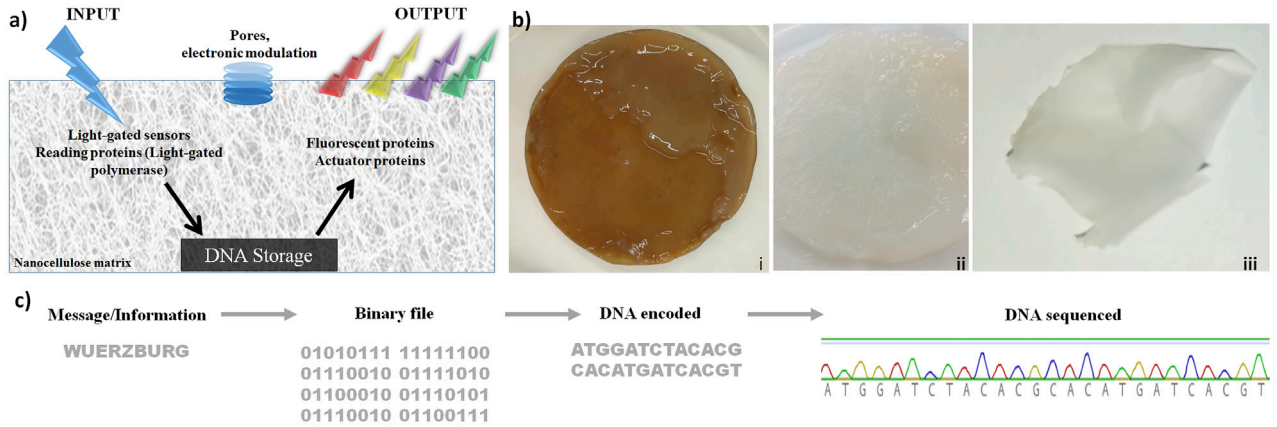


FIGURE 1 | Nanocellulose (NC) as support, chassis, and protection for DNA storage. **(A)** DNA chip concept: after light stimulation, proteins are activated and allowed to read the information stored on the nanocellulose. Subsequently, the actuator proteins read and analyze the processed DNA. **(B)** Production of bacterial nanocellulose **(i)** not purified bacterial nanocellulose from SCOBY (symbiotic culture of bacteria and yeast), **(ii)**: NC purified, and **(iii)** dried. **(C)** DNA storage on NC: the word "Wuerzburg" was encoded by an easy encoding system, where three nucleotides were used to encode one letter. The resulting DNA strand was synthesized and stored on a dried NC film at room temperature. After 2 weeks, 2 months, and 2 years, the DNA was extracted from the NC and sequenced.

A Lamarckian genetic algorithm (GA) was applied for all the docking simulations. The orientation, torsions, and position of the drug molecule were set randomly. A total of 50 runs GA were performed for each docking. The final analysis of ligand conformations as well as their interaction profile with a target protein was performed using Chimera software version 1.14 (Pettersen, et al., 2004).

The scoring function is represented by ΔG_{bind}—the binding energy difference between the bound and unbound states both for macromolecule and ligand.

$$\Delta G_{bind} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors},$$

where ΔG_{vdW} gives the difference in free energy due to van der Waals interactions in the bound and unbound states. ΔG_{elec} is the energy difference due to electrostatic interactions in the bound and unbound states, ΔG_{hbond} describes the difference in free energies due to hydrogen bond formation between the unbound and bound states, ΔG_{desolv} represents the desolvation free energy change in the unbound and bound states, and ΔG_{tors} is the energy difference of torsional entropy. The standard error for ΔG is around 2 kcal/mol (Forli, et al., 2016).

Sequenase Activity

The pH sensitivity of enzymes was demonstrated on the Sequenase ver. 2.0 (Affymetrix, United States). As the template, the DNA from bacteriophage M13mp18 was used. The reactions were performed as per manufacturer's instructions. The pH of the reaction was modified by HCl and NaOH.

T4 Kinase Phosphorylation Activity Assay

This assay followed the methods of Song and Zhao (2009) facilitating optical monitoring of the phosphorylation increase and conformational change of the substrate oligonucleotide. Using the light-gated T4 kinase construct, we observed light-controlled phosphate transfer. The assay for the T4 kinase constructs used the processed fluorescent oligonucleotides (Song and Zhao, 2009), for monitoring their activity; the calculations for the constructs considered cooperative changes following Halabi et al. (2009) and Lee et al. (2008).

Modeling of the Properties of a Nanocellulose-Based Single-Electron Transistor

The nanocellulose monomer was retrieved from the PubChem database as a 2D structure and converted to the 3D model by the ChemAxon software (Costache et al., 2018; Kim et al., 2021). The nanocellulose (NCL)-based SETs with and without iodine modification (I, I_2, and I_3 atoms), consisted of gold (111) nano-electrodes and a subunit of nanocellulose molecule as a central element. These were carried out with the combination NCL-SET, NCL(I)-SET, NCL(I_2)-SET, and NCL(I_3)-SET as molecular junctions were constructed by the Atomistix ToolKit suite and simulated using the DFT model available in the ATK software (Smidstrup et al., 2017). To analyze the HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) levels of NCL in the presence of the surrounding electrodes, the Molecular Projected Self-Consistent Hamiltonian (MPSH) was performed to obtain the MPSH states by diagonalizing the molecular part of the full self-consistent Hamiltonian (Stokbro et al., 2003).

Iodine Doping of Nanocellulose

A never-dried bacterial nanocellulose hydrogel was immersed in aqueous iodine solutions of different concentrations at room temperature for 24 h with constant shaking (150 rpm). Iodinated nanocellulose was rinsed with distilled water five

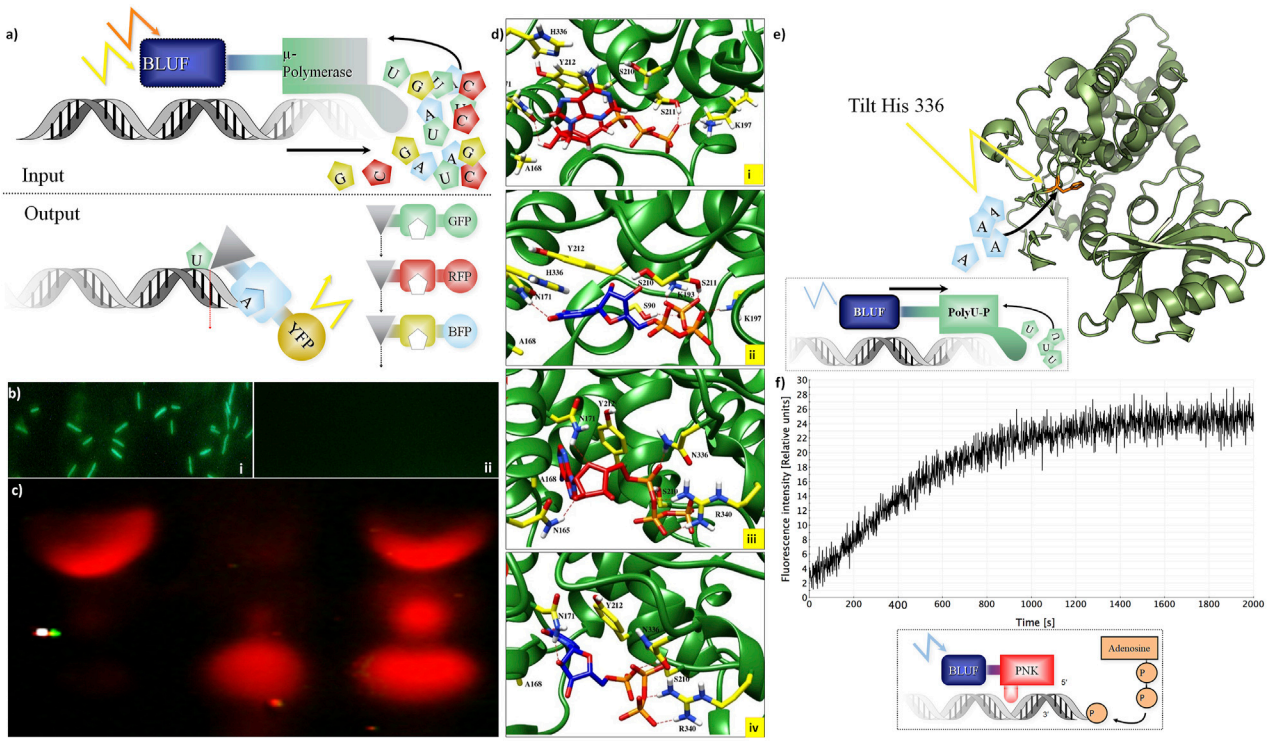


FIGURE 2 | Integration of DNA storage for easy read-in and read-out. **(A)** Concept of light-gated protein. Input (top): μ-DNA polymerase is used to achieve light-gated (BLUF domain fused to μ-DNA polymerase constructs for each nucleotide) and template-free DNA synthesis. Output constructs **(**bottom): light-gated exonuclease constructs (triangles) are fused to specific nucleotide-binding domains (squares) and trigger different fluorescent proteins for read-out. **(B)** Light-gated protein expression *in vivo*. BLUF–GFP construct transfected by bacteria with different expressions of GFP with (Panel **(i)**) and without (Panel **(ii)**) presence of blue light. **(C)** Light-gated exonuclease activity: first line: positive control (only cyanine-5–labeled hexamer loaded on gel), second line: BLUF–exonuclease activity in the dark, third line: BLUF–exonuclease activity with the presence of blue light. **(D)** Modeling: ATP (red) and UTP (blue) at binding site of the Cid1 polymerase. Panel **(i)**: docking of ATP in the wt Cid1 binding pocket. Panel **(ii)** depicted the UTP-Cid1 for the docking to the binding pocket of the wt Cid1 (PDB: 4FH5). In panel **(iii)** (ATP-Cid1 complex) and **(iv)** (UTP-Cid1 complex), the complexes are shown for the case of the H336N mutation (PDB: 4FHX). The histidine 336 does not form any hydrogen bonds with the ligands, whereas N336 is involved in H-bond formation with ATP. Residues S211 and K197 from the binding cavity of the wt. of Cid 1 polymerase (H336) are participating in hydrogen bond formation with ATP, CTP, GTP and UTP. K193 participates in hydrogen bond formation with all triphosphates, except ATP. In the latter N336 and R340 play roles in H-bond formations with ATP, GTP and UTP. N171 forms H-bonds with ATP, CTP, and UTP, but it does not form H-bond with GTP. Moreover, the number of residual contacts for the ATP as well as for the UTP docked to the binding cavity of H336N polymerase is higher compared to the ATP and UTP docked to the native structure. We may assume that mutation of H to asparagine at position 336 could influence the binding mode of ATP as well as UTP. Key amino acid residue composition inside the binding site involved in hydrogen bond formations are highlighted in yellow. Red dashed lines indicate the H-bonds. **(E)** Histidine tilting of PolyU polymerase (structure modeling). Top: a histidine in the PolyU polymerase domain (PDB file shown: 4FH3) determines if ATP or UTP will be accepted for the elongation of the RNA strand. The histidine 336 could be tilted by light to achieve rapid changes in substrate specificity according to user-specified sequences of ATP and UTP. **(F)** T4 Polynucleotide kinase as an example of DNA-processing enzyme. Light-controlled phosphate transfer measurement: assay for T4 kinase DNA elongation constructs using processed fluorescent oligonucleotides for monitoring their activity; construct calculations to predict joined cooperative changes after (Lee et al., 2008; Halabi et al., 2009). Bottom: a sketch of the light-gated protein construct. A BLUF domain (BLUF, blue) is directed by light (blue flash) and stops or activates the T4 polynucleotide kinase (PNK, red) which then transfers a phosphate group from ATP to the end of a polynucleotide **(**example: DNA, shown in grey shape).

times for 5 min to remove the remaining iodine. Measurements of conductivity were performed in triplicate by Laqua Twin (Horiba, Japan).

Engineered Patterns in a Real Biofilm

Key sensor histidine kinase genes from *B. subtilis* bacteria were artificially deleted (*kinC* and *kinD*). Another experiment utilized spontaneous mutations in the strong biofilm repressor *sinR*, which reinitiate tight interactions and achieve patterning of colonies with biofilm-forming and non-forming regions (right colony). For large-scale active DNA storage, the light-gated and monitoring constructs have to be introduced.

Membrane Dyes to Monitor Membrane Damage

Nile red-stained bacterial membrane and accumulated in areas where damage has been made.

RESULTS

Nanocellulose as Support, Chassis, and Protection for DNA Storage

These results are shown in **Figure 1**. The concept is shown in Figure 1A: nanocellulose as chassis for DNA storage, including efficient DNA protection, while processing enzymes as well as

electronic modulation are later steps. Production and purification of bacterial nanocellulose (NC) are depicted in Figure 1B. To test storage capabilities, we used different DNA oligonucleotides. **Figure 1C** shows how the word "Wuerzburg" is encoded with a simple coding schema using three nucleotides per letter. The resulting DNA was synthesized and stored on dried NC at room temperature (**Figure 1Biii**). After 2 weeks, 2 months, and 2 years, the DNA was retrieved and sequenced by Sanger sequencing (**Figure 1C**). Sequenced DNA encoded the longer segment ("University of Wuerzburg light-gated polymerase") stored for 2 years is depicted in **Supplementary Figure S1**. We noted some change/mutation in the nucleotide sequence but every nucleotide was well preserved. Looking at the long-term storage properties of nanocellulose, we estimated that under ambient conditions the DNA will be stable for at least 10 years. Under cold preservation conditions (freezer at −20 or even −80), this time would be much longer. Though this is only based on extrapolation, it suggests for nanocellulose storage in DNA good preservation for many years without any nucleotide errors.

In addition to these concrete results regarding protection, there are also general advantages: 1) the additional protection of DNA against UV radiation (Zhang et al., 2019; Klochko et al., 2021) is a clear advantage of nanocellulose. Whether this leads to as good preservation as in bones (Allentoft et al., 2012) or by vitrification (Grass et al., 2015) remains to be tested. Moreover, several other advantages of nanocellulose have become evident, in particular, 2) nanocellulose provides a versatile composite. Moreover, 3) it can be made transparent, allowing transparent coverage or can protect a display. 4) It has been shown to integrate well and efficiently functional electronic parts, and 5) it is in general an ideal composite host entity, integrating many different materials. It is also worth noting that nanocellulose is more ecologically friendly and easier to produce under standard laboratory conditions than a filter paper (made from normal cellulose). For performant DNA storage besides direct protection of DNA and the excellent properties of nanocellulose as a host material are important. Nanocellulose as a versatile composite is critical so that, for instance, conventional electronics can be directly integrated. Furthermore, a read-out of the storage can be displayed using transparent nanocellulose.

Integration of DNA Storage for Easy Read-In and Read-Out

For DNA storage allowing active read-in and read-out, we investigated nucleotide-processing enzymes (**Figure 2**). The concept (**Figure 2A**) relies on light-gated proteins, allowing them to control their activity by the light of specific wavelengths (e.g., blue light using LOV and BLUF domain as a sensor). "Read-in": the four DNA nucleotides are incorporated only if the DNA polymerase investigated here is fused to a blue light–harvesting domain (BLUF I/II domain). This activates a polymerase (**Figure 2A**). The LOV–Taq polymerase (**Supplementary Table S1**) works quite efficiently, but it can also be a template-free polymerase such as μ-DNA polymerase (Uchiyama et al., 2009). This activation by a cooperative structure change happens only if blue light hits the BLUF protein domain linked to the polymerase.

TABLE 2 | Cid1 polymerase substrate binding comparing wild type and H336N mutant.[a]

	ATP	CTP	GTP	UTP
H336	−6.81	−9.12	−8.83	−9.35
H336N	−8.63	−7.78	−7.65	−7.87

[a]*The free energies of binding ΔG (kcal/mol) for the ATP, CTP, GTP, and UTP, from the dockings to the binding cavity of the native Cid1 polymerase as well as to the one with H336N mutation. Detailed methods and standard error estimates are given in* Materials and Methods.

This concept was tested in practice first in a simpler construct controlling GFP activity by fusing it to a BLUF domain *in vivo* (**Figure 2B**). This then allows switching on green protein fluorescence only if the construct is exposed to blue light before (top), otherwise not (bottom). However, more interesting is the control of nucleotide-processing activity. "Read-out": we show light-gated exonuclease activity in **Figure 2C**. The exonuclease is operating on a Cy-5–labeled DNA hexamer. This demonstrates the concept of read-out from the DNA storage controlled by light. Similarly, we constructed several light-gated constructs (**Supplementary Table S1**) and different oligonucleotides for each application (**Supplementary Table S2**). All experiments on read-in and read-out for the different constructs were done each at least in triplicates. We can state that these light-gated constructs (exonuclease, polymerases, and nucleotide processing) work well in an efficient manner. We have thus clear "read-in" and "read-out" of information into the DNA storage.

We next wanted to study whether this is also possible for RNA, and here the Cid1 polymerase has experimentally proven polyU addition activity (Lunde et al., 2012). We tested, in addition, whether the light-gated constructs work on Cid1 polymerase and observed clear dependence of activity if the light was given; there was not any RNA synthesis if no light was present.

Next, we wanted to understand how substrate change can be effected using Cid1 polymerase. For this, specific mutations that change substrate preference from uracil to adenine are known (Lunde et al., 2012). We hence created a light-gated version of the H336N mutated version of the Cid1 polymerase. Substrate specificity can be changed by mutating H336 to A or N. The H336N mutation of Cid1 polymerase was shown by data from experiments to have a preference for ATP instead of UTP (Lunde et al., 2012). Hence, using wild type as well as the mutated version of Cid1 polymerase (made as described in *Materials and Methods*), we can either preferentially add UTP or ATP to the RNA template. In such a PCR-like process of changing the enzyme activated or used, one can, in principle, "write" with RNA. It first adds adenine, then uracil, next adenine, and so on.

The structure details of this change in substrate specificity were further investigated *in silico*.

We investigated different structures and models, thus besides Cid1 polymerase (with several available structures) there is also a PDB file (2Q66) of a yeast poly(A) polymerase with ATP and oligo(A). In result, we compare and show here one PDB structure

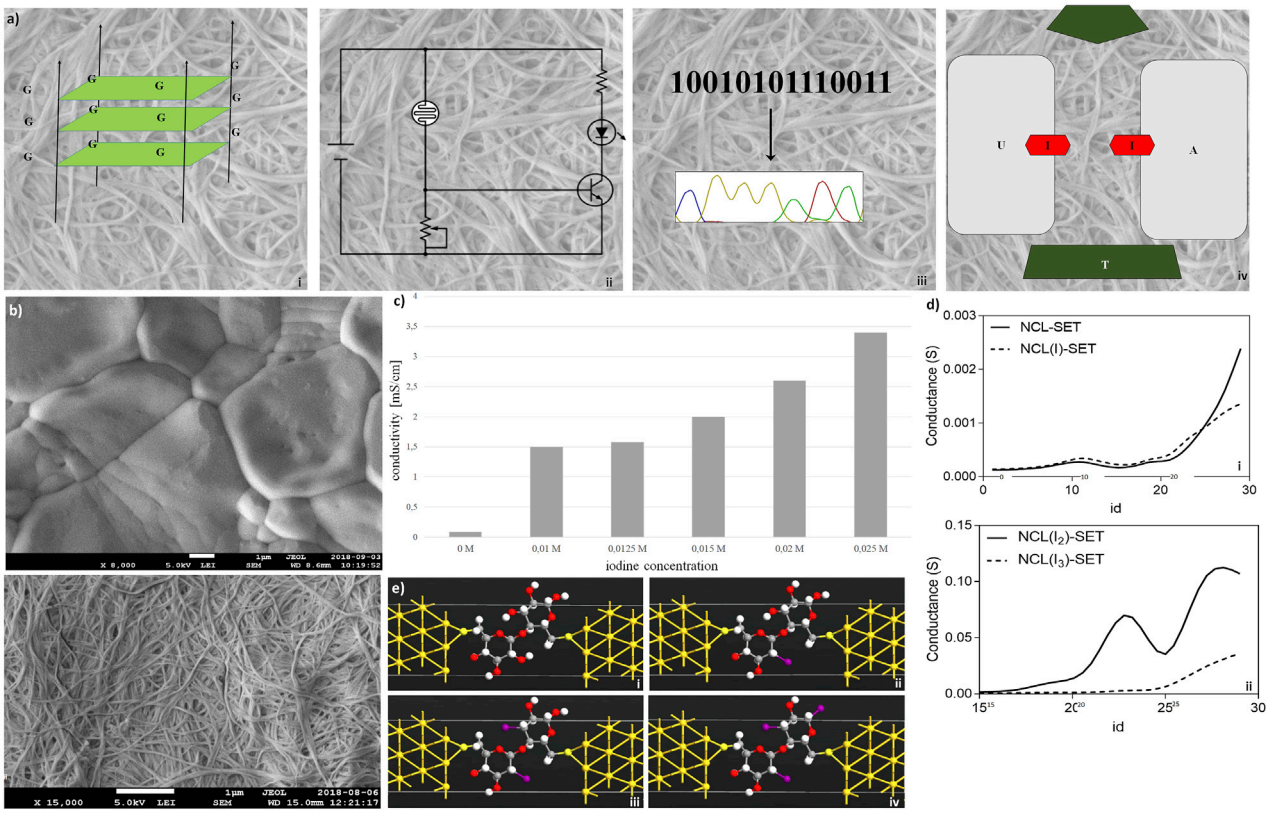


FIGURE 3 | Electronic features of nanocellulose. **(A)** Concept: nanocellulose can replace silicon, for example, introducing electronics and modifying enzymes: **(i)** G-wire quartets, allowing conduction of electric current; **(ii)** condenser/resistor, different printed electronics can be applied; **(iii)** sequencing result, different enzymes and DNA can be attached on the NC; and **(iv)** represents the introducing of SET. **(B)** Nanocellulose becomes conductive with iodine doping (0.025 M iodine) (Panel **(i)**), untreated control is depicted in Panel **(ii)**. **(C)** Measured conductivities for the iodine treated nanocellulose. **(D)** Conductance curves **(i,ii)** calculated for the nanocellulose (NCL)-based SETs with and without iodine modification (I, I_2, and I_3 atoms), consisting of gold (111) nano-electrodes and a subunit of nanocellulose molecule as a central element: NCL-SET, NCL(I)-SET, NCL(I_2)-SET, and NCL(I_3)-SET as molecular junctions. **(E)** *In silico* molecular illustration of nanocellulose (NCL)-based SETs with and without iodine modification (I, I_2, and I_3 atoms): consisting of gold (111) nano-electrodes and a subunit of nanocellulose molecule as a central element. The nanocellulose-based SETs: NCL-SET **(i)**, NCL(I)-SET **(ii)**, NCL(I_2)-SET **(iii)**, and NCL(I_3)-SET **(iv)** as molecular junctions were designed using the Virtual NanoLab software.

for the native (4FH5) and one for the mutated Cid1 (4FHX) so that we can complement the experimental data we have on Cid1 polymerase by exactly the corresponding PDB structures. For our analysis, we do not show the RNA substrate in full, but just the nucleotides. Our study was designed to exactly match the already published work of Lunde et al. (2012), in which the authors showed by experiments with nucleotide triphosphates (ATP, UTP, etc) that the H336 in the Cid1 polymerase is the crucial and essential binding site for the ATP binding and recognition. Using the structure of Cid1 polymerase (PDB file 4FH5 for wild type and PDB 4FHX for mutated version) for *in silico* modeling, the cleft of the mutated nucleotide-processing enzyme is compared to the wild type in **Figure 2D**. Cid1 is compared in a state in which the template and product RNA strands are bound and the active site is open. The conformations of ATPs (**Figures 2Di,iii**) inside the pocket differ from those of UTPs (**Figures 2Dii,iv**) when bound to the native structure of Cid1 polymerase versus the mutated H336N (detailed binding interactions in legend). In **Table 2**, the binding energies for nucleotide triphosphates are compared, in wild-type Cid1 UTP binds best with −9.35 kcal/mol, and mutated H336N prefers ATP (−8.63 kJ/mol) over UTP. H336 is hence critical for substrate specificity. These calculated binding energies give support to what has been seen by the experimental data and are in accordance with the well-known specificity of such polymerases.

The histidine 336 is pivotal for specificity (see also experimental data in Yates et al., 2012), and it was recognized as essential ATP binding site (Lunde et al., 2012). Could the histidine be changed in its conformation by light in a similar way as the light-gated domains controlling the enzyme activity? By this, we would change substrate specificity "on the fly" just by a light pulse and without a mutation and hence fast for many nucleotide synthesis steps instead of the slow PCR-type way we indicated above. In principle, this seems to be possible, but as these are challenging experiments to actually achieve this, this was again only investigated *in silico*.

Thus, irradiating the histidine at the binding pocket with its absorption maximum wavelength at 220 nm (as determined by

TABLE 3 | Nanocellulose single-electron transistor properties.[a]

Nanodevice	HOMO[a]	LUMO[a]	Gap
NCL-SET	−1.18	4.97	6.15
NCL(I)-SET	−1.32	2.77	4.09
NCL(I_2)-SET	−1.31	2.77	4.08
NCL(I_3)-SET	−1.18	1.31	2.49

[a]*HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) calculations for the NCL-based SETs, with and without iodine modification (I, I_2, and I_3 atoms), consisting of gold (111) nano-electrodes and a subunit of nanocellulose molecule as central element (all units are measured in eV).*

M. O. Iwunze, 2007), such a light pulse, if energetic enough should interfere with the binding pocket. For comparison, Wei et al. (2007) identified a single tryptophan (Trp) residue responsible for loss of binding and biological activity testing this by UV light irradiation in a humanized monoclonal antibody (MAb) against respiratory syncytial virus (RSV). Changing the histidine 336 conformation by UV light (using its 220 nm optimum for histidine light absorption), it would thus allow in principle rapid change of the substrate cleft around the histidine (**Figure 2E**; structure modeling cartoon). However, whether this can allow rapid change of incorporation of the nucleotides during active polymerization, for example, changing adenine for uracil remains to be determined in future experiments.

A digital addition of individual nucleotides can also be achieved by using nucleotidyltransferases. Here, the light-gated constructs for different wavelengths allow us again to control the activity of the nucleotidyltransferases (e.g., BLUF and LOV, see *Materials and Methods*, or using instead halorhodopsin). Each nucleotide is added and then the transferase is halted. However, this was not yet studied in experiments here.

Finally, we monitored directly activity switches of DNA-processing enzymes after light-gating activation using a BLUF domain. This is shown for T4 polynucleotide kinase, allowing nucleotide processing by phosphorylation modification if activated by blue light (**Figure 2F**). Moreover, ATP and energized nucleotides are easily integrated into the composite. They can be regenerated using either current or, for example, pH gradient–powered ATP synthetases and/or adding adenylate kinase to buffer the concentrations of all four nucleotides against each other.

Electronic features of nanocellulose are investigated in **Figure 3**. The concept is shown in **Figure 3A**: nanocellulose replaces silicon and becomes electronic. First of all, nanocellulose can well incorporate or attach nucleic acids such as DNA or RNA after chemical treatment or UV crosslinking. One possibility is to integrate electronic features and electric current. This can be accomplished (**Figure 3Ai**) through the incorporation of DNA wires such as G-wire quartets (Livshits et al., 2014). However, by direct treatment, nanocellulose may become electronic conductive and act as a capacitor or resistor (**Figure 3Aii**). Complex treatment is not necessary. We typically used a simple protocol where the aqueous suspension of DNA was directly subjected to the dried nanocellulose and the suspension was subsequently left for drying. As we used double-stranded DNA, no further procedure was necessary such as UV-immobilization or acid treatment. However, activation by acid or other pH change is another method we tested to improve DNA attachment and achieve covalent linkage.

Hence, a sequencing result may either be stored by DNA or even electronically in our nanocellulose composite. For optimal information processing, both types of storage may be used and combined (**Figure 3Aiii**). For actual electronic parts made from nanocellulose, we suggest and investigated a SET (single-electron transistor) as an attractive molecular electronics device (**Figure 3Aiv**). The experimental results regarding nanocellulose are shown in the following: iodine treatment makes the nanocellulose conductive. The treated nanocellulose surface is shown by electron microscopy (**Figure 3Bi**), and the untreated nanocellulose is shown for comparison in **Figure 3Bii**. The measured conductance in nanocellulose for different concentrations is given in **Figure 3C**. Nanostructured bridges and connections are readily visible in the electron micrographs of the nanocellulose (iodine treated and untreated samples) and are a natural part of the fine structure of the nanocellulose. However, actual electronic parts would require ultrafine structuring of the nanocellulose surface. This could be efficiently achieved with coverage by photo-paint and UV etching as in conventional semiconductor manufacturing. To estimate the miniaturization potential in this direction for electronic nanocellulose we calculated *in silico* conductance curves for the nanocellulose (NCL)-based SETs with and without iodine modification (I, I_2, and I_3 atoms; **Figure 3D** panel i for 0 and 1 iodine atoms and panel ii for 2 and 3 iodine atoms). **Figure 3E** shows in detail the dimensions of a SET, exploring maximal miniaturization for nanocellulose-based semiconductors: each SET consists of gold (111) nano-electrodes and a subunit of nanocellulose molecule as a central element: NCL-SET, NCL(I)-SET, NCL(I_2)-SET, and NCL(I_3)-SET as molecular junction. The predicted good performance properties of such a nanodevice are furthermore supported by orbital calculations on a nanocellulose SET (**Table 3**) and supported by the actual conductance data for the iodine-treated nanocellulose.

Synergy From Full Integration of All Components in the Nanocellulose Composite

If all components investigated in **Figures 1–3** are integrated into one nanocellulose composite, several synergies arise (summarized in **Table 4**): This can furthermore be exploited for a 3D printer and nanofactory (**Figure 4A**) or a performant nanocellulose composite usable as a computer chip (a "CellChip"; **Figure 4B**).

Table 4 gives more data and resulting optimal performance estimates on 1) individual enzyme performance, 2) high-density storage possibilities for DNA storage, and 3) calculation speed.

A first demonstrator using nanocellulose as chassis is summarized in **Figure 4C**. Although all major components are available and ready for use and integration, it will take several years of development for full DNA storage performance to be reliably achieved, and full synergy between all components will take even longer. However, for much easier and more reliable

TABLE 4 | Optimal performance characteristics.

Individual enzyme	Light-gated polymerase: After switching on this is as fast as normal Klenow polymerase, that is, 40–87 nucleotides per second (Schwartz and Quake, 2009). Using many molecules in parallel is easily possible using micro-wells (**Figure 4F**). Enzyme signaling can be very fast, 4 nanoseconds after inhibitor binding for tyrosine phosphatases (Chirgadze et al., 2021) and in light-harvesting reaction centers even within picoseconds (Dods et al., 2021). However, to have this much faster speed requires molecular integration of all components as explored in **Figure 3**. Molecular components we provide for the nanocellulose composite include SETs as well as self-assembling oligos (Saoji and Paukstelis, 2015) and the electronic current-carrying DNA wires (Livshits et al., 2014; Saoji and Paukstelis, 2015; Dvorkin et al., 2018)
Storage	Storage: usage in parallel allows Exabyte (Church et al., 2012); storage of millions of years (Grass et al., 2015); latest actual DNA storage achieved: 250 Terabytes per gram with correction bits; theoretical optimum: 1 Exabyte per kilogram, so 10^{18} bit per kg (Extance, 2016)
Calculation speed	Calculation speed: Use many molecules in parallel (10 million); or fast and direct information processing by light (femtoseconds; references) → Teraflops and very fast processing (up to Petahertz) processing Concrete application: Using digital picoliter PCR allows parallelization of up to 1 billion reactions in parallel, each operating on another DNA molecule in a separate well (Wohrle et al., 2020) allowing not only fast library screens but also direct access to different storage items (up to 1 billion) as RAM (random access memory) Integration of conventional electronics: gigahertz (Ghz) processors, fast light processing using LED and light-sensitive diodes (Jung et al., 2016). Single-electron transistors are biodegradable and have ultralow energy consumption, dimensions are approximately 10 nm (Xu et al., 2020)
Synergy achieves performant devices	(i) Smart DNA storage devices providing fast read-in and out by light-gated enzymes, permanent high-density DNA storage without external energy source and rapid calculations integrating standard electronics; (ii) Higher miniaturization exploiting photolithographic techniques would allow full chip integration of nanocellulose SETs and semiconductors, standard electronic components for fast processing, as well as high-density DNA storage, fast read-in and read-out by light-gated processing enzymes, synthesizing and being accessed and modulated by DNA wire networks. (iii) Nanofactory or smart (light-gated control) conveyor belt on nanocellulose composite chassis and adapts specificity according to electric current. (iv) 3D printer: self-assembling oligos plus DNA wire-grid, at cross pixels directed by pH or voltage selectively opens nanopores for ink (**Figure 4H** shows the pores we tested). (v) Artificial ribosome/translator combining different light-gated enzymes with specific peptide- or nucleotide-specificities.

output, normal electronics can also be integrated on nanocellulose paper with good performance results (Jung et al., 2015). Moreover, pH change, including pH differences caused by currents, can change the substrate specificity (**Figure 4D**, left). This is shown for the enzyme sequenase and changed nucleotide specificity for all four nucleotides (**Figure 4D**, right). Operating DNA storage in this way by fast pH change and incorporating then different substrates from the enzyme according to the applied current is an attractive option. Similarly, either pH change or light activation of polymerase allows, in principle, polymerase activity either bound to the template (copy polymerase) or without the template (*de novo* synthesizing polymerase).

There are numerous further possible refinements of the nanocellulose chassis (**Figure 4**):

We studied micro-surface patterning (**Figure 4Ei**) as well as piercing micro-holes (**Figure 4Eii**) using a *B. subtilis* biofilm with kinase mutations for patterning (see *Materials and Methods*) and compounds such as Nile Red for generation of holes. With similar strategies, one can also fix proteins or other components at specific places of the nanocellulose composite. Another application would be turning the nanocellulose composite into a nanoprinter (**Figure 4E**, bottom), either to allow the dropping of ink through tiny holes for 3D printing (bacterial membranes of *B. subtilis* stained with Nile Red, the red spots representing cellular damage, in this case, pores) or for connecting active pores (for instance light-gated) between different bacterial membranes or nanocellulose composites.

Finally, we show that we already possess a decent technical alternative to operate our DNA storage by using PCR mini wells (Wöhrle et al., 2020) allowing access to different specific storage contents (**Figure 4F**).

DISCUSSION

We show here only proof-of-principle experimental results covering many aspects of a smart nanocellulose composite. Promising for real applications are:

(1) Electronic capabilities including measured semiconductor properties and miniaturization potential shown for the composite in micrographs and *in silico* calculations for a single-electron transistor from nanocellulose.
(2) DNA storage potential includes energy-saving permanent storage with reliable, enzyme-specific controlled read-in and read-out. However, speed is currently still slow while storage density and potential total storage capacity is high (**Table 4**). Moreover, we show that nanocellulose protects DNA well and has a number of attractive properties as a composite for DNA storage.
(3) Regarding integration of enzymes into nanocellulose composites, this has pre-runners, for example, pH sensors; however, we show here for the first time an active DNA storage, that is, read-in and read-out by light-gated nucleotide-processing enzymes which remain intact and can be used for many cycles.

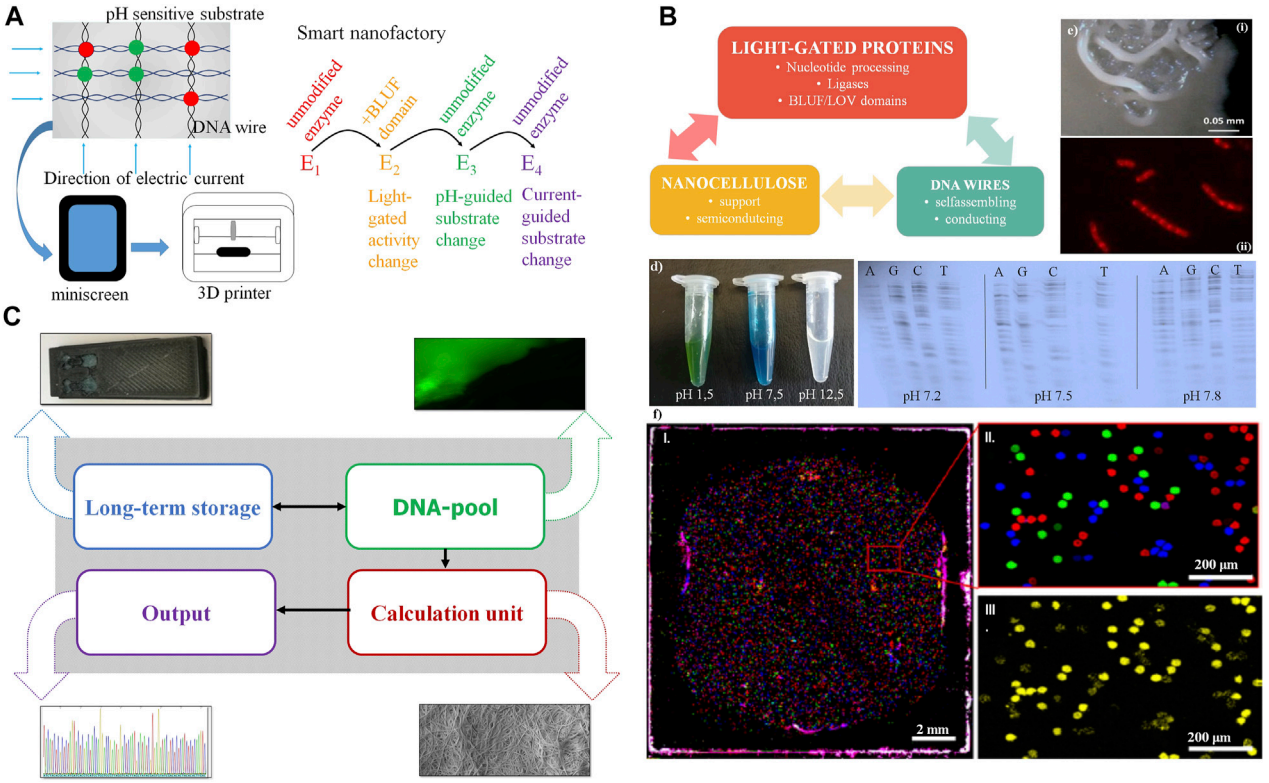


FIGURE 4 | Synergy from full integration of all components in the nanocellulose composite. **(A)** Further nanocellulose composite extensions: Output devices using the nanocellulose chassis as printer or mini screen; light-gated, current- or pH-guided enzymes have accurately controlled activity and substrate specificity by this and complement each other in nanocellulose composite cavities ("nanofactory") or use the mini wells shown in **(F)**. **(B)** Smart card (or ultimately computer chip) combining nanocellulose, light-gated proteins and DNA wires. **(C)** Demonstrator: we integrate **(i)** DNA for storage (long-term storage); **(ii)** modify the DNA pool by fluorescent nucleotides for read-in or read-out (DNA pool); **(iii)** do efficient calculations combining DNA-processing enzymes, nanocellulose semiconductors, SETs (inset), and classical electronics integrated using the nanocellulose chassis (Jung et al., 2015) (calculation unit), and **(iv)** get efficient sequence read-out (results) by fast sequencing (e.g., next-generation sequencing, output). **(D)** Left: pH-sensitive dye (malachite green) on nanocellulose; right: polymerase (sequenase) specificity are different pH levels. **(E)** Composite biofilm modification: Top: patterning of *B. subtilis* biofilm into spatially defined phenotypic changes by the accumulation of spontaneous *sinR* mutations in a subpopulation of Δ*kinCD* mutant; bottom: cells stained with Nile red. The dye accumulates in damaged regions of the membrane thus membrane-damaged areas are visualized as spotted Nile red areas in bacterial membranes. **(F)** PCR mini wells (Woehrle et al., 2020) with pico-cavities allow specific access to DNA storage molecules. **(i)**: normal size; **(ii)**: enlargement, different DNA molecule species (color); and **(iii)**: specific read-out for one species (yellow).

(4) True molecular integration (**Figure 4**) shows the highest performance potential, but needs a lot of additional development; even more nanotechnology is required for highest performance, for example, direct molecular integration and synergy. However, with this the potential for a competitor to current electronics would clearly be there, particular regarding energy-saving permanent storage.

Potential of Nanocellulose Composites

Nanocellulose and nanocellulose composites have several advantageous properties (Pagliaro et al., 2021). Also, classical electronics can easily be integrated (Jung et al., 2015). Similarly, the theoretical advantages of DNA as storage medium have been shown previously including sophisticated storing and encoding schemes (Chen et al., 2020; Chen et al., 2021; Lim et al., 2021). Based on both points, we explore to what extent nanocellulose composites may provide a basis for smart devices: we demonstrate long-lasting DNA storage in nanocellulose, enzymes allow active DNA storage by controlled read-in and writing DNA and read-out or reading DNA using light-gated nucleotide-processing enzymes. Furthermore, the combination of nanocellulose with other substances such as cinnamoyl chloride or copper iodide can provide long-term UV protection of stored DNA, thus preventing its degradation (Zhang et al., 2019; Klochko et al., 2021). Electronic properties of nanocellulose, DNA wires and single-electron transistors (SETs) give the composite even more attractive properties. We show that there are many attractive options for surface structuring of the composite for improved DNA storage: We tested DNA micro-wells and pores, electric and pH-mediated substrate change and pH sensitivity dies. Finally, we examine the strong potential for synergies from all these components.

Light-activating proteins have traditionally been used for optogenetic control and monitoring electrical and biochemical parameters (Paz et al., 2013; Hartsough et al., 2020; Ochoa-Fernandez et al., 2020), but their potential is much higher, for example, we show here that they can be used to control nucleotide

activities processing DNA enzymes. As an alternative chemical approach, the recent work of Kesici et al. used two types of photocleavable linkers that were covalently attached to various enzyme types such as polymerase, restriction enzyme and exonuclease for the reversible and controlled activation of proteins (Kesici et al., 2022). Their method is different, using DNA-processing enzymes activated by UV using a specific photocleavable linker. Instead, our approach uses a natural light-gated domain allowing, again and again, activation of the processing enzymes. In our approach, there is also no chemistry involved which means that these light-gated proteins can be synthesized in normal bacteria or eukaryotic cells using the constructs we provide. They even can act inside a cell after activation by light (not investigated here by us). Moreover, the light-gated domains allow specific activation by different wavelengths of different enzymes, so that for instance a general polymerase, a U-specific polymerase, a template-free polymerase or a template-bound (copy) polymerase can each be activated after the other and reused after some time in a new cycle. Each of these constructs (see also **Supplementary Table S2** and detailed sequences in **Supplementary Material**) has been tested in many experiments by us and we compared also different light-gated domains.

The molecular basis how the light-gated protein domains signal and regulate the activity of the subsequent protein has been investigated in several studies. In particular, regarding the BLUF domain molecular details are known (Fujisawa and Masuda, 2018): X-ray crystallography and mutagenesis disclosed that a rearrangement of the hydrogen bond network involving a specific Tyr, Gln, and the FAD cofactor of the BLUF domain are recognized as essential for formation of a BLUF protein signaling state by photoactivation. The hydrogen bond structural change in the active site is propagating in the protein, alters the conformation and transmits the light-induced allosteric signal. Investigating the BLUF domain containing protein PixD from cyanobacterium *Synechocystis* sp. PCC6803, it was found that conformations of the C-terminal helices in the BLUF domain are particularly relevant to signal transduction. The effected intramolecular conformational changes may in some BLUF domain containing proteins even control the interprotein association or dissociation and by this the activity of the protein controlled by the BLUF domain. latest advances in spectroscopy and computation allow now to reveal even more details of the molecular mechanism involved including tackling for example the complex structure of the AppA BLUF protein, which controls photosynthesis and suitable gene expression in the purple bacterium *Rhodobacter sphaeroides* (Hashem et al., 2021).

Nevertheless, Hashem et al. (2021) and Fujisawa and Masuda (2018) are descriptive studies: They nail down specific involved residues of the BLUF domain but do not look at the full protein structure. Moreover, this varies according to the specific protein structure examined. More detail is not known currently according to experimental data.

To get more insight, we did molecular dynamics simulation investigating the BLUF domain (**Supplementary Figure S2**; Methods described in **Supplementary Material**): The BLUF–POL interaction hypothesis holds that the photo-activated BLUF domain (its activated conformation) has a higher affinity to the DNA-POL, when it is inactive the affinity decreases. There is a photo-activated BLUF domain available as PBD file (6W72). We show that its conformational energy is high enough to achieve this due to its photoactivation. We used the BLUF domain of BlsA at the ground (green) and photo-activated states (cyan). As you can see, there is a conformational shift of the flexible loop structure upon the BLUF activation (RMSD = 1.76 A) in the residues 110–122 (with a gap: 113 and 114 aa). These residues were predicted as a protein–protein interface to be most likely involved in the interaction with polymerase or T4 polynucleotide kinase. An elevated energy level was also detected starting from −2,185.69 kcal/mol for the ground state to −2,051.68 kcal/mol for the activated state. This mechanism might explain how a BLUF domain interacts with the polymerase or activate it. Similarly, this helps to clarify how a BLUF domain stops or activates the T4 polynucleotide kinase.

The modification potential of our approach is high as demonstrated regarding engineering enzyme specificity for example by detailed engineering of the photo-active enzyme fatty acid photodecarboxylase in its binding pocket, substrate specificity and reaction speed (Amer et al., 2020; Gil et al., 2020). Regarding the control element advocated here, the light-activating domains BLUF or LOV allow for controllable switching on from a few seconds to several tens of minutes (Mathes and Gotze, 2015). Accurate deactivation can be achieved by adding another reporter with a reverse function, such as Opn7b or Dronpa (Karapinar et al., 2021).

Switching the substrate specificity in a controlled way allows writing or reading other nucleotide letters and, even, to change enzyme specificity. Classical but time consuming and irreversible is site-directed mutagenesis. We use such a mutation and achieve a mutated CidI polymerase with a higher preference for adenine. However, a targeted change of substrate and catalytic specificity by electrical current or pH as shown here for the DNA polymerase sequenase is particularly promising as reversible, fast and applicable to any enzyme of choice. Similarly, DNA can be switched, for instance, a pH modifying dye was used to achieve a light-driven conformational switch in i-motif DNA (Liu et al., 2007). This allows use of nucleotide-processing enzymes in several modes, for example, as template-bound copy polymerases (Klenow polymerase, sequenase) and, after a switch by current, pH or light as template-free polymerase to incorporate the nucleotide of choice. We stress that we show here only proof-of-principle for all these nucleotide-processing enzyme modifications. Accurate and fast performance will require a lot of developmental work.

Nanocellulose Composite Electronic Properties

Nanocellulose can be rendered conducting by the addition of metal ions (Silva et al., 2020) or graphene (Yang et al., 2017). We show here that it can be rendered semi-conductive by iodine doping at different concentrations. Moreover, we show the fine

structure of nanocellulose allows using miniature gates and junctions, which presents in principle an attractive alternative to silicon. To further explore this, we present simulation data on how a SET device made of nanocellulose would work, including molecular orbital calculations. By decreasing the gap between the orbitals we can increase the conductivity and vice versa. Indeed, the energy gap between HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) determines if the material is a conductor, insulator or semiconductor. Therefore, if the energy gap is small (e.g., in metals) the electrons can jump easily from the HOMO to LUMO orbitals thus the material is a conductor. If the energy gap is very high, the material is an insulator and if it is between conductor and insulator, the material is a semiconductor (useful for a transistor). In our case, the nanocellulose is probably a weak insulator (or not very good semiconductor) *per se* and it becomes the semiconductor after a doping process, which was precisely modeled *in silico* using DFT calculations on a SET device using well established techniques (Shityakov et al., 2017) (**Table 3**). As a sensitive molecular transistor, we show this here for the first time and stress the high potential of nanocellulose in this respect. However, photolithographic techniques for nano-structuring nanocellulose required for such devices to get appropriate connections for the transistor device were not attempted.

The electronic properties in nanocellulose are a new and interesting observation. This was confirmed by us by repeated measurements (at least triplicates). Moreover, Jung et al. (2015) showed that nanocellulose is a well-suited host material for conventional electronic parts. We show now as a new property that nanocellulose itself can be titrated from full conducting to semiconducting and not conducting, this indicates by the experimental data here a new area for applications of nanocellulose.

We found that the fine structure of the nanocellulose shows tiny bridges and bifurcations and that these remain also after iodine treatment (**Figure 3B**). However, to target these structures is quite a challenge: to enable tiny conducting wires transporting current on the gate and then measure the current through the other two contacts as required for a transistor is a major project needing years in a fine-structure lab. So instead, we simulated this only *in silico* including detailed molecular calculations. The theoretical calculation looked at the properties a single-electron transistor made from nanocellulose would have. They confirm the electronic properties we found in the experiment looking after the iodine doping. The detailed molecular structure of the SET is also shown (**Figure 3E**). We show by calculation of HOMO LUMO orbitals and gap energy that such a transistor made from nanocellulose is calculated to work efficiently as a SET. The results are comparable to those for two other SETs, an Indigo and a Tyrian Purple Single-Electron Nano-Transistor, respectively (Shityakov et al., 2017). Hence, the potential for powerful electronic capabilities by nanocellulose is there, though full experimental proof of transistor properties is beyond our current capabilities, instead we just calculate the properties of such a device and show that nanocellulose has a suitable fine structure to allow such usage.

Nanocellulose Smart Card Synergies

The combination of a nanocellulose chassis, nucleotide-processing enzymes and electronic properties is a major new contribution with powerful synergies (**Figure 4**; **Table 4**). For instance, we can use DNA wires on the nanocellulose and change the DNA wires by synthesis from the light-controlled nucleotide-processing enzymes. Similarly, the current transported by the DNA wires can be used to change the substrate specificity of the nucleotide-processing enzymes. Moreover, for rapid calculations nanocellulose is an optimal chassis for classical electronic parts. Furthermore, LEDs can be used for efficient operation of the nucleotide-processing enzymes of the DNA storage while the computing is done by the latest silicon chips. Our nanocellulose composite was further modified by applying nanotechnologies such as micro-pores, nanopatterning and picoliter DNA assay wells. **Table 4** illustrates the large potential of the individual components as well as some of the synergies possible.

Regarding optimized production of a nanocellulose DNA storage chip device, there is an efficient scale-up possible regarding energy and speed using digital PCR (Wöhrle et al., 2020). To obtain a functional device one could use 3D printing combined with standard photolithographic techniques to generate computer chips. We are currently testing our constructs in such a 3D printer setting. Together, this illustrates that there is already now a good potential for this approach for miniaturization and efficient printing and several working methods are at hand. However, there is still a considerable road to go regarding development and optimization to achieve efficient and cheap high-throughput production.

CONCLUSION

In this article, we show the huge potential of nanocellulose composites for information storage and smart card devices. While DNA-based long-term high-density storage capabilities are well known and undisputed (Extance, 2016), we demonstrate here the potential to tackle open challenges. Encompassed in such a device including DNA storage read-in and read-out using light-gated enzymes, it is possible to achieve fast processing by using light. This has very high potential, a proof of concept was explored but there is still a long road to get high performance. Electronic capabilities of nanocellulose are demonstrated by us including miniaturization and applying DNA wires (available, tested by us in place). In addition, direct integration of electronic components as well as many nanotechnology techniques improve this, such as pH-sensitive colors, integrating patterned bacterial membranes, nanopores, micro-wells, and digital PCR. This allows steady refinement of the nanocellulose composite. Hence, we are confident that the nanocellulose chip is an attractive device with high potential. However, for competitive applications compared to typical electronic devices, longer development is still needed. It will then achieve a cheap, environmentally friendly long-term DNA storage. Future fast processing capacity will profit from direct integration of classical electronic components (transistors,

LEDs) with novel components explored here such as light-gated nucleotide-processing enzymes, semiconducting nanocellulose with SETs, DNA wires, and oligonucleotides.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

TD conceived the idea and supervised the project. EB, DS, DL, CR, GR, and MK made molecular biology work. AH made electronic measurements, supervised by SM, and ES performed the docking study. SS created and calculated the SET model and analyzed results. TD and EB wrote the draft. All authors gave comments, provided edits, and approved the final version of the manuscript.

FUNDING

We thank Land Bavaria for its support (including its contribution to DFG Project number 324392634—TRR 221/INF). This publication was funded by the University of Würzburg in the funding program Open Access Publishing.

ACKNOWLEDGMENTS

Sincere thanks to Sara Giddins and Todd Johnsen for their diligent native speaker proofreading of this research article and to Anastasia Nenashkina for the graphical abstract preparation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.869111/full#supplementary-material

REFERENCES

Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., et al. (2012). The Half-Life of DNA in Bone: Measuring Decay Kinetics in 158 Dated Fossils. *Proc. R. Soc. B* 279 (1748), 4724–4733. doi:10.1098/rspb.2012.1745

Amer, M., Wojcik, E. Z., Sun, C., Hoeven, R., Hughes, J. M. X., Faulkner, M., et al. (2020). Low Carbon Strategies for Sustainable Bio-Alkane Gas Production and Renewable Energy. *Energy Environ. Sci.* 13 (6), 1818–1831. doi:10.1039/d0ee00095g

Azeredo, H. M. C., Rosa, M. F., and Mattoso, L. H. C. (2017). Nanocellulose in Bio-Based Food Packaging Applications. *Industrial Crops Prod.* 97, 664–671. doi:10.1016/j.indcrop.2016.03.013

Cao, B., Zhang, X., Wu, J., Wang, B., Zhang, Q., and Wei, X. (2021). Minimum Free Energy Coding for DNA Storage. *IEEE Trans.on Nanobioscience* 20 (2), 212–222. doi:10.1109/tnb.2021.3056351

Chen, K., Zhu, J., Bošković, F., and Keyser, U. F. (2020). Nanopore-based DNA Hard Drives for Rewritable and Secure Data Storage. *Nano Lett.* 20 (5), 3754–3760. doi:10.1021/acs.nanolett.0c00755

Chen, W., Han, M., Zhou, J., Ge, Q., Wang, P., Zhang, X., et al. (2021). An Artificial Chromosome for Data Storage. *Natl. Sci. Rev.* 8 (5), nwab028. doi:10.1093/nsr/nwab028

Chirgadze, Y. N., Battaile, K. P., Likhachev, I. V., Balabaev, N. K., Gordon, R. D., Romanov, V., et al. (2021). Signal Transfer in Human Protein Tyrosine Phosphatase PTP1B from Allosteric Inhibitor P00058. *J. Biomol. Struct. Dyn.*, 1–10. doi:10.1080/07391102.2021.1994879

Church, G. M., Gao, Y., and Kosuri, S. (2012). Next-generation Digital Information Storage in DNA. *Science* 337 (6102), 1628. doi:10.1126/science.1226355

Costache, A., Biro, E., Hoffmann, E., and Szakacs, P. (2018). An Application of ChemAxon's Platform for Education. *Abstr. Pap. Am. Chem. Soc.* 256, 435.

Dandekar, T., Bencurova, E., and Shityakov, S. (2019). *Nanocellulose Composite with Electronic Properties and DNA Traces Germany Patent Application DE102019000074A1*.

de Amorim, J. D. P., de Souza, K. C., Duarte, C. R., da Silva Duarte, I., de Assis Sales Ribeiro, F., Silva, G. S., et al. (2020). Plant and Bacterial Nanocellulose: Production, Properties and Applications in Medicine, Food, Cosmetics, Electronics and Engineering. A Review. *Environ. Chem. Lett.* 18 (3), 851–869. doi:10.1007/s10311-020-00989-9

Dods, R., Båth, P., Morozov, D., Gagnér, V. A., Arnlund, D., Luk, H. L., et al. (2021). Ultrafast Structural Changes within a Photosynthetic Reaction Centre. *Nature* 589 (7841), 310–314. doi:10.1038/s41586-020-3000-7

Dvorkin, S. A., Karsisiotis, A. I., and Webba da Silva, M. (2018). Encoding Canonical DNA Quadruplex Structure. *Sci. Adv.* 4 (8), eaat3007. doi:10.1126/sciadv.aat3007

Erlich, Y., and Zielinski, D. (2017). DNA Fountain Enables a Robust and Efficient Storage Architecture. *Science* 355 (6328), 950–954. doi:10.1126/science.aaj2038

Extance, A. (2016). How DNA Could Store All the World's Data. *Nature* 537 (7618), 22–24. doi:10.1038/537022a

Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., and Olson, A. J. (2016). Computational Protein-Ligand Docking and Virtual Drug Screening with the AutoDock Suite. *Nat. Protoc.* 11 (5), 905–919. doi:10.1038/nprot.2016.051

Fujisawa, T., and Masuda, S. (2018). Light-induced Chromophore and Protein Responses and Mechanical Signal Transduction of BLUF Proteins. *Biophys. Rev.* 10 (2), 327–337. doi:10.1007/s12551-017-0355-6

Gil, A. A., Carrasco-López, C., Zhu, L., Zhao, E. M., Ravindran, P. T., Wilson, M. Z., et al. (2020). Optogenetic Control of Protein Binding Using Light-Switchable Nanobodies. *Nat. Commun.* 11 (1). 4044. doi:10.1038/s41467-020-17836-8

Goldman, N., Bertone, P., Chen, S., Dessimoz, C., LeProust, E. M., Sipos, B., et al. (2013). Towards Practical, High-Capacity, Low-Maintenance Information Storage in Synthesized DNA. *Nature* 494 (7435), 77–80. doi:10.1038/nature11875

Grass, R. N., Heckel, R., Puddu, M., Paunescu, D., and Stark, W. J. (2015). Robust Chemical Preservation of Digital Information on DNA in Silica with Error-Correcting Codes. *Angew. Chem. Int. Ed.* 54 (8), 2552–2555. doi:10.1002/anie.201411378

Habibi, Y. (2014). Key Advances in the Chemical Modification of Nanocelluloses. *Chem. Soc. Rev.* 43 (5), 1519–1542. doi:10.1039/c3cs60204d

Halabi, N., Rivoire, O., Leibler, S., and Ranganathan, R. (2009). Protein Sectors: Evolutionary Units of Three-Dimensional Structure. *Cell* 138 (4), 774–786. doi:10.1016/j.cell.2009.07.038

Hartsough, L. A., Park, M., Kotlajich, M. V., Lazar, J. T., Han, B., Lin, C. J., et al. (2020). Optogenetic Control of Gut Bacterial Metabolism to Promote Longevity. *Elife* 9. e56849. doi:10.7554/eLife.56849

Hashem, S., Macaluso, V., Nottoli, M., Lipparini, F., Cupellini, L., and Mennucci, B. (2021). From Crystallographic Data to the Solution Structure of Photoreceptors: the Case of the AppA BLUF Domain. *Chem. Sci.* 12 (40), 13331–13342. doi:10.1039/d1sc03000k

Hoeng, F., Denneulin, A., and Bras, J. (2016). Use of Nanocellulose in Printed Electronics: a Review. *Nanoscale* 8 (27), 13131–13154. doi:10.1039/c6nr03054h

Iwunze, M. O. (2007). The Characterization of the Fluorescence of L-Histidine in Simulated Body Fluid. *J. Photochem. Photobiol. A Chem.* 186 (2-3), 283–289. doi:10.1016/j.jphotochem.2006.05.034

Jiao, D., Lossada, F., Guo, J., Skarsetz, O., Hoenders, D., Liu, J., et al. (2021). Electrical Switching of High-Performance Bioinspired Nanocellulose Nanocomposites. *Nat. Commun.* 12 (1), 1–10. doi:10.1038/s41467-021-21599-1

Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997). Development and Validation of a Genetic Algorithm for Flexible Docking 1 1Edited by F. E. Cohen. *J. Mol. Biol.* 267 (3), 727–748. doi:10.1006/jmbi.1996.0897

Jung, Y. H., Chang, T. H., Zhang, H., Yao, C., Zheng, Q., Yang, V. W., et al. (2015). High-performance Green Flexible Electronics Based on Biodegradable Cellulose Nanofibril Paper. *Nat. Commun.* 6, 7170. doi:10.1038/ncomms8170

Karapinar, R., Schwitalla, J. C., Eickelbeck, D., Pakusch, J., Mücher, B., Grömmke, M., et al. (2021). Reverse Optogenetics of G Protein Signaling by Zebrafish Non-visual Opsin Opn7b for Synchronization of Neuronal Networks. *Nat. Commun.* 12 (1). 4488. doi:10.1038/s41467-021-24718-0

Kesici, M.-Z., Tinnefeld, P., and Vera, A. M. (2022). A Simple and General Approach to Generate Photoactivatable DNA Processing Enzymes. *Nucleic acids Res.* 50 (6), e31. doi:10.1093/nar/gkab1212

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2021). PubChem in 2021: New Data Content and Improved Web Interfaces. *Nucleic Acids Res.* 49 (D1), D1388–D1395. doi:10.1093/nar/gkaa971

Klochko, N. P., Barbash, V. A., Klepikova, K. S., Kopach, V. R., Tyukhov, I. I., Yashchenko, O. V., Zhadan, D. O., Petrushenko, S. I., Dukarov, S. V., Sukhov, V. M., and Khrypunova, A. L. (2021). Biodegradable Flexible Transparent Films with Copper Iodide and Biomass-Derived Nanocellulose for Ultraviolet and High-Energy Visible Light Protection. *Sol. Energy* 220, 852–863. doi:10.1016/j.solener.2021.04.014

Koch, J., Gantenbein, S., Masania, K., Stark, W. J., Erlich, Y., and Grass, R. N. (2020). A DNA-Of-Things Storage Architecture to Create Materials with Embedded Memory. *Nat. Biotechnol.* 38 (1), 39–43. doi:10.1038/s41587-019-0356-z

Lee, J., Natarajan, M., Nashine, V. C., Socolich, M., Vo, T., Russ, W. P., et al. (2008). Surface Sites for Engineering Allosteric Control in Proteins. *Science* 322 (5900), 438–442. doi:10.1126/science.1159052

Lim, C. K., Nirantar, S., Yew, W. S., and Poh, C. L. (2021). Novel Modalities in DNA Data Storage. *Trends Biotechnol.* 39 (10), 990–1003. doi:10.1016/j.tibtech.2020.12.008

Liu, H., Xu, Y., Li, F., Yang, Y., Wang, W., Song, Y., et al. (2007). Light-driven Conformational Switch of I-Motif DNA. *Angew. Chem. Int. Ed.* 46 (14), 2515–2517. doi:10.1002/anie.200604589

Livshits, G. I., Stern, A., Rotem, D., Borovok, N., Eidelshtein, G., Migliore, A., et al. (2014). Long-range Charge Transport in Single G-Quadruplex DNA Molecules. *Nat. Nanotech* 9 (12), 1040–1046. doi:10.1038/nnano.2014.246

Lunde, B. M., Magler, I., and Meinhart, A. (2012). Crystal Structures of the Cid1 Poly (U) Polymerase Reveal the Mechanism for UTP Selectivity. *Nucleic Acids Res.* 40 (19), 9815–9824. doi:10.1093/nar/gks740

Mangalam, A. P., Simonsen, J., and Benight, A. S. (2009). Cellulose/DNA Hybrid Nanomaterials. *Biomacromolecules* 10 (3), 497–504. doi:10.1021/bm800925x

Mathes, T., and Götze, J. P. (2015). A Proposal for a Dipole-Generated BLUF Domain Mechanism. *Front. Mol. Biosci.* 2, 62. doi:10.3389/fmolb.2015.00062

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., et al. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* 30 (16), 2785–2791. doi:10.1002/jcc.21256

Ochoa-Fernandez, R., Abel, N. B., Wieland, F.-G., Schlegel, J., Koch, L.-A., Miller, J. B., et al. (2020). Optogenetic Control of Gene Expression in Plants in the Presence of Ambient White Light. *Nat. Methods* 17(7), 717, 725. doi:10.1038/s41592-020-0868-y

Osorio, M., Fernández-Morales, P., Gañán, P., Zuluaga, R., Kerguelen, H., Ortiz, I., et al. (2019). Development of Novel Three-Dimensional Scaffolds Based on Bacterial Nanocellulose for Tissue Engineering and Regenerative Medicine: Effect of Processing Methods, Pore Size, and Surface Area. *J. Biomed. Mat. Res.* 107 (2), 348–359. doi:10.1002/jbm.a.36532

Pagliaro, M., Ciriminna, R., Yusuf, M., Eskandarinezhad, S., Ahmad Wani, I., Ghahremani, M., et al. (2021). Application of Nanocellulose Composites in the Environmental Engineering: A Review. *Jcc* 3 (7), 114–128. doi:10.52547/jcc.3.2.5

Paz, J. T., Davidson, T. J., Frechette, E. S., Delord, B., Parada, I., Peng, K., et al. (2013). Closed-loop Optogenetic Control of Thalamus as a Tool for Interrupting Seizures after Cortical Injury. *Nat. Neurosci.* 16 (1), 64–70. doi:10.1038/nn.3269

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera?A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* 25, 1605–1612. doi:10.1002/jcc.20084

Razaq, A., Nyström, G., Strømme, M., Mihranyan, A., and Nyholm, L. (2011). High-Capacity Conductive Nanocellulose Paper Sheets for Electrochemically Controlled Extraction of DNA Oligomers. *Plos One* 6 (12), e29243. doi:10.1371/journal.pone.0029243

Saoji, M., and Paukstelis, P. J. (2015). Sequence-dependent Structural Changes in a Self-Assembling DNA Oligonucleotide. *Acta Cryst. D. Biol. Crystallogr.* 71, 2471–2478. doi:10.1107/S1399004715019598

Schwartz, J. J., and Quake, S. R. (2009). Single Molecule Measurement of the "speed Limit" of DNA Polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 106 (48), 20294–20299. doi:10.1073/pnas.0907404106

Shityakov, S., Roewer, N., Förster, C., and Broscheit, J.-A. (2017). In Silico Modeling of Indigo and Tyrian Purple Single-Electron Nano-Transistors Using Density Functional Theory Approach. *Nanoscale Res. Lett.* 12 (1), 439. doi:10.1186/s11671-017-2193-7

Silva, R. R., Raymundo-Pereira, P. A., Campos, A. M., Wilson, D., Otoni, C. G., Barud, H. S., et al. (2020). Microbial Nanocellulose Adherent to Human Skin Used in Electrochemical Sensors to Detect Metal Ions and Biomarkers in Sweat. *Talanta* 218, 121153. doi:10.1016/j.talanta.2020.121153

Smidstrup, S., Stradi, D., Wellendorff, J., Khomyakov, P. A., Vej-Hansen, U. G., Lee, M. E., et al. (2017). First-principles Green's-function Method for Surface Calculations: A Pseudopotential Localized Basis Set Approach. *Phys. Rev. B* 96 (19). 195309. doi:10.1103/physrevb.96.195309

Stokbro, K., Taylor, J., Brandbyge, M., Mozos, J. L., and Ordejon, P. (2003). Theoretical Study of the Nonlinear Conductance of Di-thiol Benzene Coupled to Au(111) Surfaces via Thiol and Thiolate Bonds. *Comput. Mater. Sci.* 27 (1-2), 151–160. doi:10.1016/S0927-0256(02)00439-1

Tao, J., Wang, R., Yu, H., Chen, L., Fang, D., Tian, Y., Xie, J., Jia, D., Liu, H., Wang, J., Tang, F., Song, L., and Li, H. (2020). Highly Transparent, Highly Thermally Stable Nanocellulose/Polymer Hybrid Substrates for Flexible OLED Devices. *ACS Appl. Mat. Interfaces* 12 (8), 9701–9709. doi:10.1021/acsami.0c01048

Thomas, B., Raj, M. C., B, A. K., H, R. M., Joy, J., Moores, A., et al. (2018). Nanocellulose, a Versatile Green Platform: from Biosources to Materials and Their Applications. *Chem. Rev.* 118 (24), 11575–11625. doi:10.1021/acs.chemrev.7b00627

Uchiyama, Y., Takeuchi, R., Kodera, H., and Sakaguchi, K. (2009). Distribution and Roles of X-Family DNA Polymerases in Eukaryotes. *Biochimie* 91 (2), 165–170. doi:10.1016/j.biochi.2008.07.005

Urbano, A. L. (2021). *DNA Writer: Storing Information in DNA Exercise. Montessori Muddle*. Available at: https://montessorimuddle.org/2013/02/02/dna-writer/ Accessed June 22, 2018.

Wei, Z., Feng, J., Lin, H.-Y., Mullapudi, S., Bishop, E., and Tous, G. I. (2007). Identification of a Single Tryptophan Residue as Critical for Binding Activity in a Humanized Monoclonal Antibody against Respiratory Syncytial Virus. *Anal. Chem.* 79 (7), 2797–2805. doi:10.1021/ac062311j

Wöhrle, J., Krämer, S. D., Meyer, P. A., Rath, C., Hügle, M., Urban, G. A., et al. (2020). Digital DNA Microarray Generation on Glass Substrates. *Sci. Rep.* 10 (1), 5770. doi:10.1038/s41598-020-62404-1

Xu, X., Heinig, K.-H., Möller, W., Engelmann, H.-J., Klingner, N., Gharbi, A., et al. (2020). Morphology Modification of Si Nanopillars under Ion Irradiation at Elevated Temperatures: Plastic Deformation and Controlled Thinning to 10 Nm. *Semicond. Sci. Technol.* 35. 015021. doi:10.1088/1361-6641/ab57ba

Yang, W., Zhang, Y., Liu, T., Huang, R., Chai, S., Chen, F., et al. (2017). Completely Green Approach for the Preparation of Strong and Highly Conductive Graphene Composite Film by Using Nanocellulose as Dispersing Agent and Mechanical Compression. *ACS Sustain. Chem. Eng.* 5 (10), 9102–9113. doi:10.1021/acssuschemeng.7b02012

Yates, L. A., Fleurdépine, S., Rissland, O. S., De Colibus, L., Harlos, K., Norbury, C. J., et al. (2012). Structural Basis for the Activity of a Cytoplasmic RNA Terminal Uridylyl Transferase. *Nat. Struct. Mol. Biol.* 19, 782–787. doi:10.1038/nsmb.2329

Zhang, Z., Zhang, B., Grishkewich, N., Berry, R., and Tam, K. C. (2019). Cinnamate-Functionalized Cellulose Nanocrystals as UV-Shielding Nanofillers in Sunscreen and Transparent Polymer Films. *Adv. Sustain. Syst.* 3 (4), 1800156. doi:10.1002/adsu.201800156

Conflict of Interest: Author GR was employed by BioCopy GmbH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Review
PUBLISHED 10 August 2022
DOI 10.3389/fbioe.2022.909591

OPEN ACCESS

EDITED BY
Jean Marie François,
Institut Biotechnologique de Toulouse (INSA), France

REVIEWED BY
Aleš Berlec,
Institut Jožef Stefan (IJS), Slovenia
Felix Broecker,
Idorsia Pharmaceuticals Ltd, Switzerland

*CORRESPONDENCE
Jingyu Chen,
chenjy@cau.edu.cn

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 31 March 2022
ACCEPTED 15 July 2022
PUBLISHED 10 August 2022

CITATION
Dong Y, Xu T, Xiao G, Hu Z and Chen J (2022), Opportunities and challenges for synthetic biology in the therapy of inflammatory bowel disease. *Front. Bioeng. Biotechnol.* 10:909591. doi: 10.3389/fbioe.2022.909591

Opportunities and challenges for synthetic biology in the therapy of inflammatory bowel disease

Yumeng Dong[1,2], Tiangang Xu[1], Guozheng Xiao[1], Ziyan Hu[1] and Jingyu Chen[1]*

[1]College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, China, [2]Suzhou U-Synbio Co., Ltd., Suzhou, China

Inflammatory bowel disease (IBD) is a complex, chronic intestinal inflammatory disorder that primarily includes Crohn's disease (CD) and ulcerative colitis (UC). Although traditional antibiotics and immunosuppressants are known as the most effective and commonly used treatments, some limitations may be expected, such as limited efficacy in a small number of patients and gut flora disruption. A great many research studies have been done with respect to the etiology of IBD, while the composition of the gut microbiota is suggested as one of the most influential factors. Along with the development of synthetic biology and the continuing clarification of IBD etiology, broader prospects for novel approaches to IBD therapy could be obtained. This study presents an overview of the currently existing treatment options and possible therapeutic targets at the preclinical stage with respect to microbial synthesis technology in biological therapy. This study is highly correlated to the following topics: microbiota-derived metabolites, microRNAs, cell therapy, calreticulin, live biotherapeutic products (LBP), fecal microbiota transplantation (FMT), bacteriophages, engineered bacteria, and their functional secreted synthetic products for IBD medical implementation. Considering microorganisms as the main therapeutic component, as a result, the related clinical trial stability, effectiveness, and safety analysis may be the major challenges for upcoming research. This article strives to provide pharmaceutical researchers and developers with the most up-to-date information for adjuvant medicinal therapies based on synthetic biology.

KEYWORDS

inflammatory bowel disease, MicroRNA, mesenchymal stem cell, live biotherapeutic products, fecal microbiota transplant, synthetic biology

1 Introduction

Inflammatory bowel disease (IBD), mainly including Crohn's disease (CD) and ulcerative colitis (UC) is a complex, chronic intestinal autoimmune disease. Similar pathological changes and clinical symptoms are usually found in such diseases, while the targeting sites of inflammation may be varied in patients' digestive systems. Inflammation caused by UC is mainly found in the colon or rectum, whereas inflammation caused by

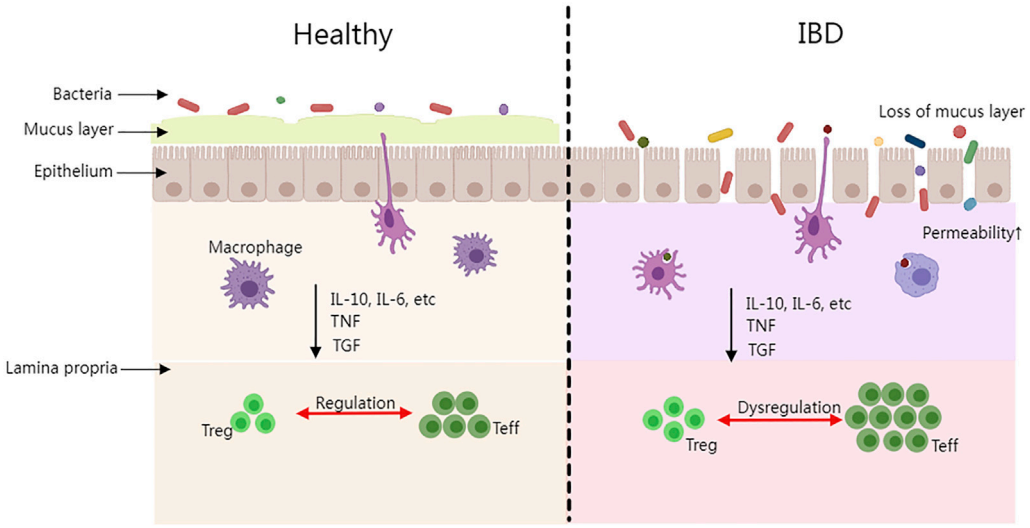


FIGURE 1
The intestinal difference between healthy people and IBD patients. (Treg, regulatory T cell; Teff, effector T cell; IL, interleukin; TNF, tumor necrosis factor; TGF, and transforming growth factor).

CD may be expected throughout the digestive tract, from the mouth to the anus (Zeng et al., 2019). In Western countries, the incidence of IBD stabilized in the 20th century, while the incidence rate in developing countries remains increased in recent years. For instance, the number of IBD patients is expected to reach 1.5 million by 2025 in China based on the current epidemiological data (Kaplan, 2015), which could place a significant strain on public health care systems worldwide. Aside from intervening in the development of IBD, investment in the investigation of IBD's infectious causes and related novel treatment strategies could be the most efficient and economically friendly way to lessen the burden.

The pathogenesis of IBD is still under research, however, the dysbiosis of the gut microbiota, the host gene, the immune system, and non-inheritable factors are already known as the most prevalent consensus aspects. Firstly, the majority of IBD patients have imbalanced gut microbiota compositions. Intestinal species diversity and stability have dropped dramatically (mainly Firmicutes), while potentially dangerous microbes have increased (primarily Proteobacteria such as Enterobacteriaceae, Bilophila) and potentially protective anti-inflammatory microbes have declined (Lee and Chang, 2020). It is commonly acknowledged that the Firmicutes/Bacteroidetes (F/B) ratio plays a significant role in preserving a healthy intestinal homeostasis. Dysbiosis is characterized as a decreased F/B ratio, which is seen as a typical phenomenon for IBD (Stojanov, Berlec, and Štrukelj, 2020). In addition, Firmicutes and Bacteroidetes support the general adjustment of the intestinal mucus barrier (*in vivo*) by changing goblet cells and mucin glycosylation to maintain colonic epithelial homeostasis (Wrzosek et al., 2013). Secondly, polymorphisms and mutations of the host genome are essential factors in IBD development. Approximately 240 risk loci associated with IBD have been identified through genome-wide association studies, in which IBD is more susceptible among these groups of individuals (Lange et al., 2017). Thirdly, in IBD patients, the homeostatic balance of the immune system may be disrupted which could result in the induction of an inflammatory response. The complex regulatory process involves macrophages, dendritic cells (DC), helper T cells (Th), regulatory T cells (Treg), effector T cells (Teff), and other immune cells as well as cytokines, interleukin (IL), tumor necrosis factor (TNF), transforming growth factor (TGF), and so on secreted by immune cells (Rogler and Andus, 1998). IBD progresses and tissue damage is caused by the imbalance of pro-inflammatory and anti-inflammatory cytokines (Guan and Zhang, 2017). Figure 1 depicts the differences between healthy people with a balanced immune system and IBD patients in terms of epithelial tissue shape, microorganisms, and immunological components. Microbial antigens in the intestinal cavity migrate into the lamina propria when the intestinal epithelial barrier is breached. An acute mucosal inflammatory response is formed when immune cells in the intestinal lamina propria (such as macrophages and Teff) display a high immunological response and release a large number of cytokines (such as TNF, IL-10, IL-6, and TGF). The acute inflammatory response could stimulate the immune cells and help them eliminate germs and pathogens from the patient's body. If immune cells continue to activate or the pathogens continue to stimulate the immune system while the activity of regulatory cells is suppressed during this process, chronic enteritis will develop over time, and persistent

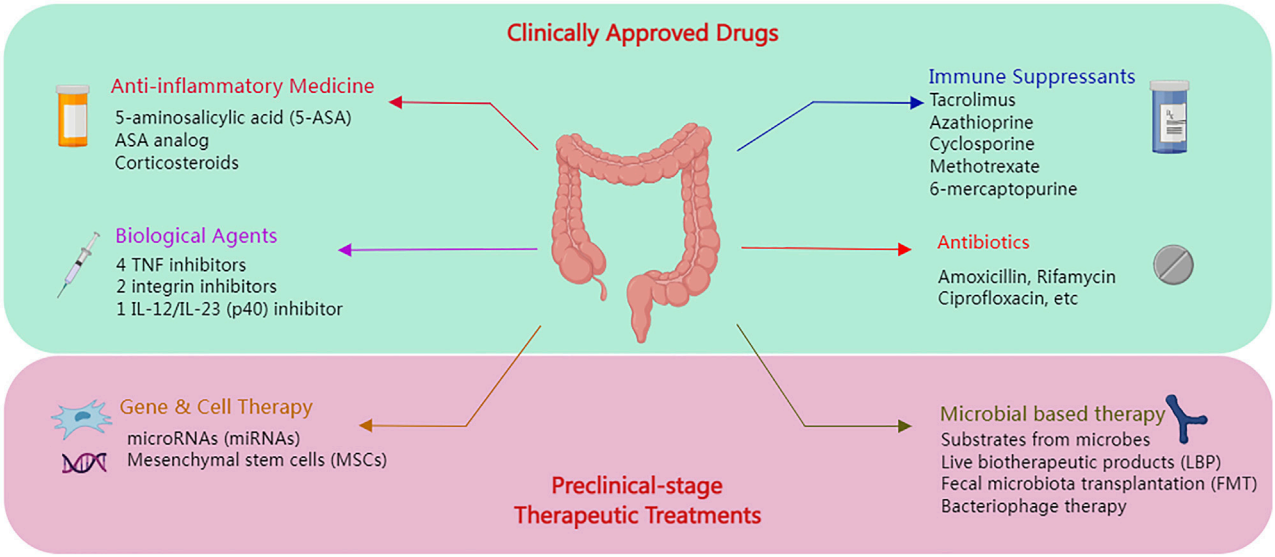


FIGURE 2
The overview of the IBD therapy treatment.

inflammation will lead to IBD disease. Finally, it is suggested that non-heritable factors such as environmental pollution factors, unbalanced daily diets, smoking, antibiotic abuse, etc., could play a role in IBD pathogenesis. Having a family history of IBD, nursing, eczema, and drinking tap water were all identified as risk factors in the incidence case investigations. The longer the breastfeeding period is, the larger the preventive effect it could have on the infant, thus lowering the risk of IBD. (Kane et al., 2000; Baron et al., 2005; Ananthakrishnan et al., 2018). Although, the aforementioned elements are all confirmed as related factors to the progression of IBD, it is known that none of the elements could be solely sufficient for IBD development.

The therapeutic approach has also been developed at the same time as IBD pathological research, which suggested its clinically approved anti-inflammatory medications, antibiotics, corticosteroids, immune suppressants, and biological treatments. Meanwhile, an ongoing study also suggested further preclinical-stage therapeutic alternatives in genetic and cellular therapies, live biotherapeutic products (LBP), fecal microbiota transplantation (FMT), etc. (Oka and Sartor, 2020). Figure 2 displays the overview.

In the early 1940s, many drugs were developed to treat IBDs of various severity, such as mesalazine, olsalazine, and balsalazide disodium (Williams, 1994). Sulfasalazine (SAS) was one of the most significant and effective anti-inflammatory drugs; 5-aminosalicylic acid (5-ASA) is the active principle of SAS. Prednisone, budesonide, methylprednisolone, and hydrocortisone are among the corticosteroids used to treat UC (Bar-Meir et al., 1998; Schauer et al., 2021). Glucocorticoids have a long history in IBD treatment, which could help restore the intestinal barrier function as well as reduce inflammation (Riccardi et al., 2008; Greenhill, 2014; Marcin et al., 2016). Furthermore, previous literature also suggested evidence of glucocorticoid addiction, while some patients could be resistant to glucocorticoids. Moreover, IBD could also be treated with antibiotics such as amoxicillin, rifamycin, ciprofloxacin, ethambutol, and fosfomycin. Contrarily, antibiotics could result in the risk of decreasing the overall bacterial diversity while stifling beneficial bacteria, which might trigger an imbalance of the gut microbiota (Lewis et al., 2015). The immune suppressants, such as azathioprine, 6-mercaptopurine, methotrexate, cyclosporine, and tacrolimus, could be another type of treatment for IBD (Ardizzone, Cassinotti, and de Franchis, 2012). Other than that, biological agents could also be another alternative for IBD treatment. Regarding the treatment of inflammatory bowel disease (IBD), the Food and Drug Administration (FDA) has approved seven biologics in the current state of the art: four TNF inhibitors (infliximab, adalimumab, certolizumab, and golimumab); two integrin inhibitors (natalizumab and vedolizumab); and one inhibitor of IL-12/IL-23 (p40) (ustekinumab) (Ardizzone et al., 2012; Danese, Vuitton, and Peyrin-Biroulet, 2015). The pharmaceuticals nominated were approved clinically and authorized for sale. However, with the advancement of synthetic biology, efforts are being made to develop and deploy cutting-edge medical treatments.

Synthetic biology is a novel technology that combines computer science, molecular biology, system biology, bioengineering, and other interdisciplinary fields. It enables the creation of new biological substances or the restructuring

of existing biosystems by modifying a genetic code or a critical metabolic pathway. Synthetic biology is progressing at a breakneck pace these days, due to the decline in the cost of biomolecular synthesis technology, the advanced progress of genetic engineering technology, and a profound grasp of genomic databases. The design and transformation of enzymes, metabolic pathways and networks, and biological chassis are at the core of synthetic biology technology. Currently, efficient enzyme design platforms based on synthetic biology have been built. Currently, effective platforms for creating enzymes have been built using synthetic biology. The rapid and targeted artificial evolution of enzymes has been carried out, the targeted enzyme mutants have been generated, and the enzyme libraries of various commonly used enzymes in the industry have successfully been constructed using high-throughput screening and testing methods. Simultaneously, the gene-editing platforms have been established to redesign the metabolic network of the chassis organisms, and a complete chassis cell bank including high-performance strains such as *Escherichia coli*, yeast, and lactic acid bacteria was established. Synthetic biology has had a significant impact on a variety of domains such as cell therapy (Yin et al., 2019), environmental pollution detection (Xinyi et al., 2019), the biosynthesis of unnatural compounds (Luo et al., 2019), and so on. Synthetic biology could also contribute greatly to the IBD treatment regarding various aspects, such as microbiota-derived metabolites, microRNAs (miRNAs), mesenchymal stem cells (MSCs), calreticulin, and microbe-related methods such as LBP, FMT, bacteriophage therapy, engineered bacteria, and their functional secreted synthetic products for IBD medical implementation.

2 Role of intestinal microbiota-derived metabolites in the therapy of IBD

As previously stated, the pathophysiology of IBD has not been fully researched. The relationship between intestinal microbes and innate immunity has been implicated in the etiology of IBD in several studies. However, the precise mechanism of action is still being investigated. Among the various pathogenesis mechanisms, the metabolites from the gut microbiota are considered to be one of the primary modes. The metabolites derived from different dietary substrates could influence the immune system and the permeability of the mucosal epithelium. The changes in the diversity and amount of the intestinal microbiota could lead to the fluctuation of the composition and concentration of metabolites. Among these metabolites, the bile acids, short-chain fatty acids (SCFAs), tryptophan, and succinic acid have been studied in the pathogenesis of IBD.

The bile acids produced in the liver are the end product of cholesterol metabolism. In addition to participating in the digestion of dietary lipids and fat-soluble vitamins, bile acids could also act as a signal regulator, exerting metabolic and immune effects (Marilidia et al., 2018; Albillos, Gottardi, and Rescigno, 2019). The bile acids are a group of molecules synthesized in the liver which can be further metabolized by the gut microbiota in the intestine. Multiple nuclear receptors are involved in the regulation of bile acid metabolism, such as the farnesoid X receptor (FXR), fibroblast growth factor 19 (FGF19), and G protein-coupled BAs receptor 5 (TGR5) (Browning et al., 2019; Q. Zhai et al., 2019a). The TGR5 regulates the macrophages by the nuclear factor- κB (NF-κB) and releases the cytokines IL-1, IL-6, and TNF (Calmus et al., 2010; Keitel et al., 2008; Y. D. Wang, et al., 2011). Through the regulation of the intestinal microbial composition and the gene expression of the corresponding regulatory factors, bile acid metabolism could be regulated, which implies that the bile acids could be a potential target in future IBD treatment.

SCFAs, which consist of acetate, propionate, and butyrate, are known as beneficial dietary metabolites generated from microbiota-accessible carbohydrates with different proportions depending on an individual's dietary habit. The SCFAs could pass through the epithelium and trigger the transformation of Treg to Teff (Lavelle and Sokol, 2020). SCFA deficiency has been suggested by a previous research study as a risky influential factor in IBD development. Butyrate was suggested to have a boost effect on the butyrate transporter MCT-1 and reduce inflammation in UC patients *via* the inhibition of NF-kB activation (Vanhoutvin et al., 2009). The SCFAs are thought to be involved in the treatment of *Akkermansia muciniphila* (*A. muciniphila*) and *Clostridium* cocktails, which will be discussed in depth in the following part.

Tryptophan, an essential human aromatic amino acid, is obtained through everyday foods such as poultry and fish. The gut bacteria can convert tryptophan to aryl hydrocarbon receptor (AhR) ligands. T cell immunity is mediated by AhR, a transcription factor that is activated by IL-22 (Zenewicz et al., 2008). The AhR ligands were lowered and inflammation was reduced in mice after they were inoculated with three *Lactobacillus* strains that are capable of tryptophan metabolism (Lamas et al., 2016).

Aside from that, succinic acid, a tricarboxylic acid cycle intermediate, is becoming a hot topic in the treatment of IBD. Succinic acid has been shown to regulate macrophages *via* IL-1 (Mills et al., 2016).

In conclusion, research into the link between intestinal microbiota-derived metabolites and IBD has a promising future. Furthermore, future studies will focus on the regulation of microbiota-derived metabolites using a combination of metagenomics, host reporter assays, synthetic biology, and bioinformatics technology.

TABLE 1 miRNAs and the targeted genes/signaling pathways/transcripts in IBD patients.

miRNA	Expression	Targeted gene/signaling pathway/transcript associated with IBD	References
miR-7	Increased	NF-κB signaling pathway; RNF183; IκBα	Qiao et al. (2016)
miR-10a	Decreased	MyD88 pathway; Toll-like receptor (TLR); Interleukin (IL)-12/IL-23p40	Wu et al. (2015)
miR-16	Decreased/ Increased	NF-κB signaling pathway; denosine A2a receptor; TNF-α and IL-12p40	Huang et al. (2015); Tian et al. (2016)
miR-21	Increased	NF-κB signaling pathway; TNF-α; PTEN/PI3K/AKT pathway; RhoB; The programmed cell death 4 (PDCD4) tumor suppressor gene	Yang et al. (2013); Ludwig et al. (2013); Yang et al. (2013); Zhang, et al. (2015)
miR-31	Increased	WNT and Hippo signaling pathway; IL-13; IL-25; Hypoxia inducible factor 1	Markus et al. (2018), Olaru et al. (2015), Shi et al. (2017), Tian et al. (2019)
miR-23a	Increased	Tumor necrosis factor alpha inhibitor protein 3 (TNFαIP3); NF-κB signaling pathway; TNF-α	Felwick, et al. (2019)
miR-122	—	The gene nucleotide-binding oligomerization domain 2 (NOD2)	Chen et al. (2013)
miR-155	Increased	FOXO3a; IL13RA1; Est-1; IL-23/17/6	Min et al. (2014); Hou et al. (2017); Markus et al. (2018)
miR-185-3p	Increased	Colon cancer–associated transcript-1 (CCAT1); MLCK signaling pathway	Ma et al. (2019)
miR-192	Decreased	Macrophage Inflammatory Peptide-2α; NOD2	Chuang et al. (2014)
miR-214	Increased	NF-κB/IL-6 pathway; Phosphatase and tensin homolog; PDZ and LIM domain protein 2	Christos et al. (2015); Liu et al. (2019)
miR-141	Decreased	CXCL12β	Huang et al. (2014)
miR-494-3p	Decreased	IKKβ/NF-κB	Song et al. (2021)
miR-511-3p	Decreased/ Increased	Toll-like receptor-4 (TLR4)	Heinsbroek et al. (2015)

3 Role of miRNAs in the therapy of IBD

In the early 1990s, the non-coding single-stranded miRNA with a length of roughly 21–25 nucleotides was discovered. A great number of similar studies have been found on the critical role that miRNAs play in the onset and progression of IBD. Table 1 lists the recently discovered, abnormally expressed miRNAs in IBD patients, as well as their probable regulatory locations and signaling pathways. miRNA has been shown to affect key cytokines that are involved in the pathogenesis of IBD. Cytokines have a significant impact on regulating the immune response and maintaining physiological homeostasis. Immune illnesses like IBD could be caused by a cytokine imbalance. An in-depth study of the regulation process of cytokines is of great significance for the pathogenesis of IBD. In the future, predicting miRNAs that can regulate certain cytokines and better clarification of the regulatory mechanism of miRNAs against cytokines at the molecular level will be a new therapeutic drug development idea based on the biological roles of cytokines. Concerning the current research, miRNAs produced by host cells could infiltrate intestinal bacteria and thus could regulate the expression of intestinal bacterial genes, which enables a better management between the host and the intestinal bacteria. Gut microorganisms, on the other hand, could control the host gene expression by altering the level of expression of their miRNAs (Filip et al., 2016). In animal testing with a mouse model, inhibiting the expression of one of the miR-425 targets—Foxol was found to be able to interfere with the differentiation of T cell into Th17 (Xue et al., 2018). Blocking miRNA *in vivo* could be an appropriate therapeutic method for the treatment of IBD. Overall, sufficient literature studies have indicated the critical role miRNA plays in the diagnosis, prevention, and therapy of IBD.

The miRNA treatment method is not without its drawbacks. First and foremost, the efficient creation of miRNA differential expression profiles is required. MiRNA chip technology, gene set enrichment analysis, and bioinformatics analysis are currently being implemented to gradually settle the subject (Ma et al., 2019). Second, suitable targeted drug delivery vectors for miRNA nucleic acid medicines to intestinal immune cells must be explored and designed. Third, the future research plan is to evaluate the effect of the reproduction and blocking of the miRNA function on the abnormal response of the IBD immune system through cell and animal experiments. In summary, miRNA is expected to be an important gene target and a potential nucleic acid drug in the future.

4 Role of MSCs in the therapy of IBD

MSCs are known as one of the most commonly used stem cells in cell therapy, which have a promising and novel therapeutic future for the IBD treatment approach. MSCs' excellent tissue regeneration and immune regulatory abilities

provide a foundation for their broad application in IBD treatment. The efficacy and safety of low-dose MSC injections for CD therapy have been verified in recent trials (Molendijk et al., 2013). T cells, neutrophils, and macrophages are among the immune cells that MSCs can regulate (Wang et al., 2019). MSCs originating from various tissues, such as placenta (P-MSCs), umbilical cord (C-MSCs), bone marrow (bm-MSCs), adipose tissue (at-MSCs), and gingiva (g-MSCs) exhibit varying levels of proliferation, differentiation, and migration. MSCs from the placenta and adipose tissues have better immunoregulatory properties than the others (Talwadekar, et al., 2015). Furthermore, studies in mice showed that MSCs decrease IBD by changing the redox balance. In the MSC-injected animals, the levels of reactive oxygen species (ROS) and inflammation-related markers (TNF-, IL-4, and CD8) were reduced (Jung et al., 2020). Furthermore, a new MSC-coated approach shows that antibody-coated MSCs could be transported more efficiently to inflammatory colon regions, improving therapeutic efficacy. The survival rates of mice have improved considerably (Ko et al., 2010). In addition to immunosuppression and tissue repair ability, MSCs also have a strong effect on restoring the diversity and richness of normal gut flora, as well as on intestinal flora regulation. As a result, in the case of IBD, combining MSCs with microbial therapy could result in a more effective clinical therapeutic outcome. The *Lactobacillus* rhamnosus culture supernatant combined with bm-MSCs improves the intestinal barrier function and affects autophagy and lymphocyte function (Rui-Cong et al., 2016; Cai et al., 2019). However, there are still obstacles in the way of MSC therapeutic development: 1) the chemotactic mechanism of MSCs, as well as differentially expressed genes and pathways; 2) the efficient migration of MSCs to the targeted organs or tissues; 3) the *in vivo* residence period; and 4) the optimum source, dose, and infusion mode.

5 Role of calreticulin in the therapy of IBD

Calreticulin is a calcium-binding chaperone that has a function in integrin subunit activation (ITGAs). The suppression of calreticulin binding to ITGAs could reduce neutrophil and T cell adhesiveness, alleviating IBD symptoms. The interaction of the calreticulin and ITGAs on the pathogenesis of IBD, on the other hand, is currently being studied. ER-464195-01, a small oral chemical, was developed to prevent calreticulin from binding to ITGAs. Pro-inflammatory genes were downregulated and IBD's severity was reduced in the mice models according to the transcriptome analysis (Ohkuro, Kim, Kuboi, Hayashi, and Fukamizu, 2018). Another study found that the mean level of anti-calreticulin antibodies was considerably higher in patients with UC than in people with healthy gut microbiota (Watanabe et al., 2006). Furthermore, Mendlovic et al. (2017) identified, cloned, and expressed the Taenia solium calreticulin. The experimental colitis mice were orally administered with the calreticulin. The calreticulin significantly reduced the inflammatory parameters, including TNF-α and IL-6, and thus prevented the experimental intestinal inflammation. Calreticulin has the potential to be used as a treatment for IBD related to immune suppressants and microbial-based medicine. The difficulties with this strategy are that it requires more research on the therapy mechanism. Synthetic biology could be used in the methodological development of transcriptomics or bioinformatics to find gene regulatory locations to control or block the calreticulin binding to its site of action.

6 Role of microbiota in the therapy of IBD

Intestinal microbial diversity and stability are essential variables in IBD. Furthermore, the loss of the mucus layer in IBD patients could increase the permeability of the epithelium to microorganisms, which could contribute to immunological activation and thereby induce an inflammatory response (Shan, Lee, and Chang, 2022). Based on microbial regulation, various therapeutic options may be available.

6.1 Live biotherapeutic products

Recently, the role of gut microbiota in the development, progression, and remission of IBD has caught much attention from pharmaceutical researchers and clinical product developers. Live bacterial species that may be able to survive and grow in the gastrointestinal tract and provide a health benefit to the host by modifying the microbiota are known as probiotics (FAO/WHO,2002). The LBP is regarded as the next-generation probiotic. LBP was defined by the FDA in 2016 as a biological product that: 1) contains live organisms, such as bacteria; 2) is used to prevent, treat, or cure a disease or a condition in humans; and 3) is not a vaccine. There has been preclinical research on the effectiveness of LBP in the treatment of IBD.

6.1.1 *Lactococcus lactis* strains

L. lactis is a non-pathogenic, non-colonizing bacterium that has a long history of usage in fermented foods, which is classified as a "generally regarded as safe" (GRAS) microorganism by the FDA. *L. lactis* was genetically modified to produce biologically active compounds that could be administered directly to the mucosa. *L. lactis* was genetically modified to secrete the anti-inflammatory cytokine IL-10 by Steidler et al. (2000). In animal experiments, the daily administration of *L. lactis* expressing IL-10 resulted in a reduction in IBD symptoms. Another study suggested that *L. lactis* was modified to release elafin, which

has anti-inflammatory characteristics. Elafin as a natural protease inhibitor is expressed in the healthy intestinal mucosa. In a mouse IBD model, the oral treatment of elafin-expressing *L. lactis* reduced inflammation and restored gut homeostasis (Motta et al., 2012). There was also an investigation into the ability of *L. lactis* I-1631 which is a non-engineered *L. lactis* isolated from fermented milk products to carry the bacterial enzyme superoxide dismutase (SodA) (Ballal et al., 2015). SodA has the ability to detoxify superoxide anions and show anti-oxidative characteristics. Because of the enzyme's short half-life, SodA delivery by *L. lactis* would be more effective than SodA as an individual (Weber, 2015). Furthermore, a dairy *L. lactis* NZ9000 strain (NZ9000/IL-35) was modified to express murine IL-35. And an oral treatment of NZ9000/IL-35 inhibited the dextran sodium sulfate (DSS)-induced colitis progression. Furthermore, associated cytokines such as IL-6, IL-17, IFN-, and TNF- were proved to be modulated which suggested that NZ9000/IL-35 could be a good candidate for preventing IBD development (J. Wang et al., 2019). These investigations give an excellent foundation for the potential efficacy of engineered *L. lactis* as LBP in the treatment of IBD.

6.1.2 *Saccharomyces* yeast strains

Since the 1950s, *Saccharomyces* yeast species have been authorized to be sold as probiotics and were categorized as safe strains (Mcfarland, 2010). Sylvester et al. (2012) conducted a study that showed that giving *Saccharomyces boulardii (S. boulardii)* to infants could be an effective treatment for necrotizing enterocolitis. However, another study suggested a contradictory outcome, which stated that an engineered *S. boulardii* producing IL-10 had no significant medical effect on IBD in mice as compared to the control group (Svenja et al., 2013). With the advancement of biotechnology, a yeast strain gene library was successfully constructed, which is useful for further designed editing. Furthermore, due to its increased efficacy, broader universality, and stability, the CRISPR-Cas9 system is gradually advancing in eukaryotic systems (Sen and Mansell, 2020). Scott et al. (2021) used a CRISPR–Cas9-based technique to create an altered *Saccharomyces cerevisiae (S.cerevisiae)* that expresses the human P2Y2 purinergic receptor and the ATP-degrading enzyme apyrase as a self-tunable probiotic yeast. The modified human P2Y2 receptor binds to eATP, which is produced by activated immune cells and commensal bacteria, with a 1,000-fold affinity. The eATP is thought to be an inflammation signal that promotes IBD progression by increasing the production of pro-inflammatory cytokines, inhibiting Treg activation, and raising the production of pro-inflammatory cytokines. In the meantime, the apyrase can hydrolyze eATP, reducing the inflammatory reaction. An oral administration of modified *S. cerevisiae* to mice reduced experimental intestinal inflammation (Scott et al., 2021).

6.1.3 *Akkermansia muciniphila* strains

A. muciniphila is a next-generation probiotic that was first isolated from healthy human feces and has shown promise in the treatment of inflammatory bowel disease (IBD) (Derrien, 2004; Ting and Zhang, 2019). However, contradictory research results were obtained regarding whether *A. muciniphila* promotes or inhibits inflammation in IBD patients (T. Zhang et al., 2021). The positive results revealed: 1) the *A. muciniphila* probably relieves IBD by increasing SCFA production, improving the diversity of gut microbiota, and thus promoting Treg differentiation (Zhai R. et al., 2019); 2) the *A. muciniphila* probably relieves IBD by increasing SCFA production, improving the diversity of gut microbiota, and thus promoting Treg differentiation (Zhai et al., 2019); 3) In mouse research, *A. muciniphila* can successfully inhibit macrophage infiltration, hence blunting IBD (L. Wang et al., 2020). The negative results revealed: 1) *A. muciniphila* could allow microorganisms to enter the epithelium through mucus layer degradation (T. Zhang et al., 2021); 2) The *A. muciniphila* aggravated the symptoms of IBD in IL-10-deficient mice (Seregin et al., 2017), the mechanism is possible due to the lipopolysaccharides (LPS) of *A. muciniphila* which can cause higher levels of cytokine production including IL-1, IL-4, IL-6, TNF-α, etc. (Meng and Lowell, 1997; Singh and Jiang, 2003).

A. muciniphila, a type of gut commensal bacterium that could be utilized as LBP in the future, still has several limitations. The first step is to determine whether *A. muciniphila* has a pro-inflammatory or anti-inflammatory effect in IBD patients. Secondly, because *A. muciniphila* is particularly sensitive to oxygen, consideration must be given to its isolation, purification, cultivation, and storage (Ouwerkerk et al., 2016). Third, the mucin in the specific cultivation medium is an animal-derived protein that may cause an allergic response. Later, the use of *A. muciniphila* as an LBP could be concentrated on the aforementioned aspects.

6.1.4 *Escherichia coli* strains

Since the first world war, *E. coli* Nissle 1917 (EcN) has been used in the treatment of numerous gastrointestinal disorders and is one of the best-studied non-pathogenic Gram-negative probiotic strains. EcN appeared to be as effective as mesalazine in the treatment of IBD, according to existing research and data obtained (Schultz and Butt, 2010). EcN was known as a popular carrier for the application of genetically engineered biosynthesis with respect to its qualities of safety and effectiveness. EcN was found to have a strong upregulation ability in the tight junction protein ZO-1 expression in murine intestinal epithelial cells in previous works. A higher ZO-1 expression protects mucosal permeability, and hence has the potential to be used as a treatment for IBD (Ukena et al., 2007).

EcN was also modified to release colicins such as E1 and E9, which have been found to kill adherent-invasive *E. coli* (AIEC)

strains *via* an allelic exchange technique. The AIEC is thought to be a key pathogenic element in the development of IBD (Kotlowski, 2016).

6.1.5 *Clostridium* cocktail

The *Clostridium* species could be classified into 19 clusters (I to XIX) (Collins et al., 1994). Based on previous studies, the *Clostridium* clusters IV, XIVa, and XVIII were decreased in IBD patients (Frank et al., 2007). Because different strains have varied metabolic and immunological activities, combining them could result in a more comprehensive IBD treatment effect. Atarashi et al. (2014) identified 17 *Clostridium* strains from healthy human feces and created a *Clostridium* cocktail to test the impact on IBD in mice. The results demonstrated that the cocktail can successfully prevent the intestinal inflammation caused by the DSS. The mechanism can be described by the following factors: 1) the synthesis of SCFAs, which elicits the Treg; 2) the conversion of indole from tryptophan by *Clostridium*, which has been shown to improve the epithelial barrier (Yosuke et al., 2013); and 3) a rise in gut microbial diversity.

6.2 Products from engineered microbes

6.2.1 *Salmonella* effector protein AvrA

The soluble effector protein in intestinal probiotics transforms into the cytoplasm of the target cells and suppresses the inflammatory and immunoregulatory pathways, alleviating the IBD inflammation response. AvrA, a *Salmonella* acetyltransferase, inhibits the activation of a number of inflammatory effector genes. However, due to the pathogenicity of *Salmonella*, it is not appropriate to administer it. AvrA and other virulent proteins will be given together, potentially posing a health risk. As a result, using modern bioengineering technology, researchers discovered a strategy to solely distribute the naturally occurring immunomodulatory protein AvrA in the absence of *Salmonella*. The functional proteins were produced and purified after the ArvA genes were cloned in *E. coli*. Purified AvrA was then turned into a cross-linked protein nanoparticle that might be used to deliver drugs (Herrera Estrada et al., 2017). The anti-inflammatory efficacy was demonstrated *in vitro* and in murine colitis models, indicating that it has clinical promise for IBD treatment.

As molecular and immunology technologies advance, more molecular active compounds could be investigated for the treatment of IBD. Together with biotechnology, the intestinal effector pattern could lead to an effective therapeutic method in future IBD treatments.

6.3 Fecal microbiota transplantation

FMT has been used for more than 50 years since the discovery of the gut microbiota's role in IBD. By engrafting the microbiota from a healthy donor, FMT aims to re-establish the gut microbial population in the recipient. After filtering, the donor's feces were administered through enema, colonoscopy, nasoduodenal, or nasogastric infusion (Borody and Khoruts, 2012). Anderson et al. (2012) conducted a systemic analysis of FMT in IBD and found that 69% of patients with IBD were able to achieve remission. Greenhil. (2014) reviewed 31 studies that used FMT to treat IBD in which 71% of IBD patients reported a decrease in symptoms. The research available is minimal, and further work is needed to ensure that FMT is a valid strategy. The commercialization of the FMT approach faces a number of challenges, including an assurance of safety, contamination risk, donor stability, and public acceptance. Furthermore, the enteric virome must be considered because it has a significant impact on the host's physiology. As a result, when performing the FMT treatment, it is vital to evaluate the potential risks posed by a change in the enteric virome (Norman et al., 2015). Furthermore, since the strains are long-term colonizers in the human gut, the FMT could be a trustworthy source of species for LBP development. The species isolated from the FMT preparations could also be considered safe cytokine delivery vehicles through microbial synthesis technology. The strains could help restructure the microbiome imbalances.

6.4 Bacteriophage therapy

Recently, the drug development research on bacteriophage therapy has returned to the fore again because of antibiotic resistance issues. According to the study, the number and abundance of bacteriophages on the surface of the intestinal mucosa have increased in IBD patients, implying that bacteriophages might display an undiscovered function in the progression of IBD. According to researchers, bacteriophages may kill probiotics in the colon, leading to a preponderance of "harmful bacteria" and an inflammatory reaction (Duerkop et al., 2018). To be more specific, the AIEC is thought to be a key pathogenic component in the development of IBD. Three phages targeting AIEC have been identified from wastewater and have been shown to diminish AIEC colonization in the intestine. In mice with DSS-induced colitis, the decrease of AIEC induced a laxative effect (Matthieu et al., 2017). Bacteriophages are a key component of the mucosal barrier's defense against bacteria. Studies conducted *in vitro* have demonstrated that phages could stick to the mucus layer, reducing microbial colonization and disease (Barr et al., 2013). Additionally, research has shown that the development of metabolic diseases could be triggered by bacterial translocation from the gut to tissues, which would also cause inflammation. By directly eradicating fragile bacteria, phages could prevent bacterial translocation and, in return, the gut inflammation can be brought on by bacterial translocation (Qv et al., 2021).

The intricacy of intestinal bacteriophages and viruses, as well as their relationships, has increasingly been demonstrated using

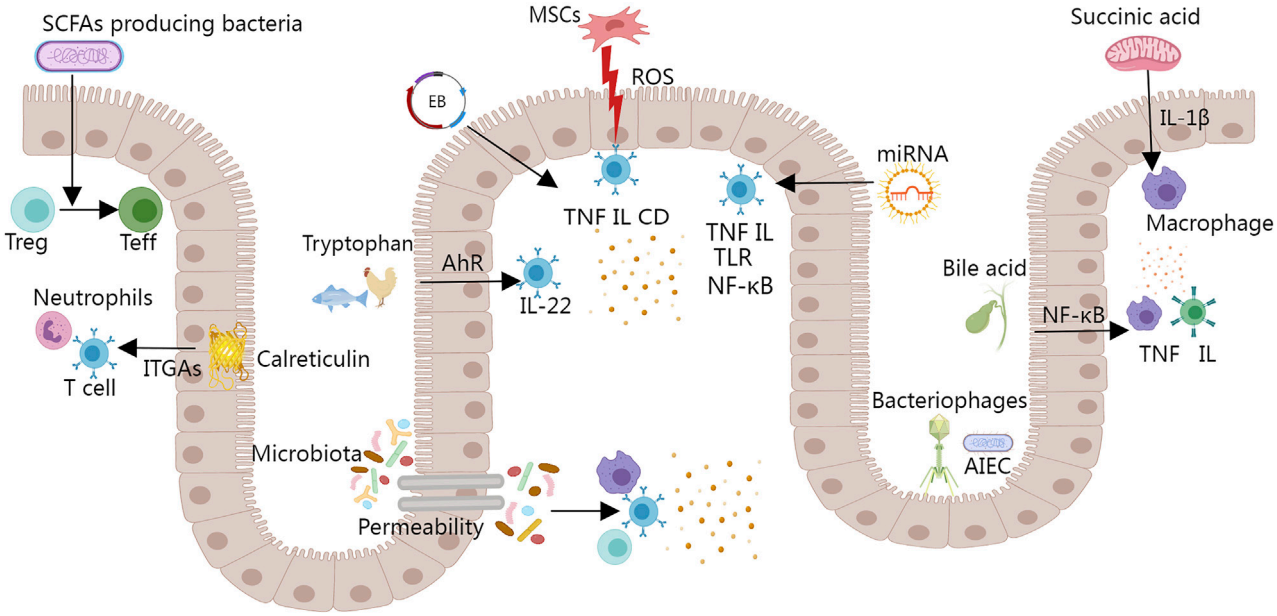


FIGURE 3
The current IBD treatment methods and potential targeted regulatory sites. (Treg, regulatory T cell; Teff, effector T cell; ITGAs, integrin α subunits; MSCs, mesenchymal stem cells; ROS, reactive oxygen species; EB, engineered bacteria; AhR, aryl hydrocarbon receptor; IL, interleukin; TNF, tumor necrosis factor; CD, cluster of differentiation; TLR, Toll-like receptor; NF-κB, nuclear factor- κB).

high-throughput metagenomic sequencing, transcriptomic, and proteomic techniques (Zwa et al., 2021). There are still many hurdles to overcome before bacteriophage therapy is approved for broad-scale clinical use: 1) scientific research and practical application need to be verified; 2) the virus database is not complete enough; 3) individual differences such as host age, sex, and diet; 4) the alteration of the bacteriophage community during IBD progression and the potential influential mechanism. In clinical experiments, bacteriophages could be utilized to target and destroy the bacteria that cause inflammation, perhaps slowing or even preventing the progression of IBD. Bacteriophage therapy could also be tailored to increase probiotic growth. Despite the fact that there are still many unknowns to be discovered, the promise of phage therapy is exciting. In the future, bacteriophage therapy could be considered to be used in conjunction with other microbial synthesis technologies to improve the specificity of virus pathogenicity and efficiently increase immunogenicity.

The existing and proposed therapy techniques based on microbial regulation are covered in the aforementioned text. Last but not least, the safety of genetically modified organisms (GMOs) is a serious challenge associated with the therapy approach based on genetically engineered bacteria. The implementation must be governed by strict regulations: 1) genetically engineered strains must be safe for human consumption; 2) the discharge and treatment of genetically engineered bacteria must be strictly regulated to avoid unpredictable gene variation, leakage, drift, and pollution; and 3) clinical applications must be subjected to a thorough scientific evaluation and strict government approval. In the current state of the art, the FDA's attitude toward gene intervention therapy methods is to issue related regulations and guidelines (Jensen, Gtzsche, and Woldbye, 2021). In a nutshell, the benefits and drawbacks of microbial-based IBD therapy coexist.

To date, IBD therapy remains a topic that requires more studies. It is still not exactly clear how the IBD pathogenesis is connected with gut microbiota, intestinal microbiota-derived metabolites, the immune system, etc. To elucidate the disease triggers in IBD, more *in vivo* and *in vitro* studies are required. With the finding of more therapeutic targets, further optimization of synthetic biology approaches may be needed in the future. The current popular therapeutic techniques, prospective targeted regulatory locations, and proposed mechanisms are outlined in Figure 3.

7 Conclusion

IBD, including CD and UC, is an autoimmune illness. With today's fast-paced lifestyle, the number of IBD patients will continue to rise in the future due to unbalanced diets, work, and rest. The etiology of IBD is complex and the disease

is easy to relapse, which brings a serious economic burden to patients and increases the pressure on society. Therefore, the ultimate purpose of the treatment of IBD is to reduce the number of relapses and hospitalizations, preserve long-term disease remission, and improve the long-term quality of life. Targeted antibiotic therapy is a sensible method, but the risk of resistant bacteria and the resulting gut flora imbalance make it unsuitable for long-term use. There is evidence showing that the imbalance of gut microbiota homeostasis is regarded as one of the essential initiating factors of IBD. Subsequently, the advancement of synthetic biotechnology provides technical assistance for new drug development.

The following procedures could offer suggestions for the clinical studies of brand-new medications: Firstly, the abnormal expression of genes, receptors, proteins, and other biomolecules in IBD patients was first evaluated using transcriptomics or protein analysis methods, as well as other biologic approaches. Secondly, pharmaceutical researchers and developers could create therapeutic strategies that are specifically targeted at the molecules with aberrant expressions. Thirdly, it is crucial to pick appropriate carriers for regulatory components. The options include engineering strains, encapsulation-coated techniques, MSCs, etc. Finally, to provide a theoretical basis for therapeutic application, research at the cellular and animal levels could also be conducted.

The FMT is a growing and appealing treatment, yet it still has a lot of flaws. Individual differences between donors and recipients, for example, are unknown, and there are no universal donors who can give consistent efficacy. Furthermore, there is no standardization of FMT donor selection, fecal sample preparation, or transplantation modality. And after the FMT therapy, the bacteriophages in parts of the patients increase. The increase of phages has the potential to intensify the inflammatory responses based on the animal study. But the stains isolated from the FMT could be further identified and applied to other biological treatment schemes.

In contrast, the LBP, as a next-generation product, is more flexible and easier to moderate. The LBP could be designed for specific patients based on the certainty of the microbial community of individuals. The specific gene could be operated by novel biotechnology and the microbes can be used as effective carriers. However, there are only therapeutic effects of the LBP in animal models and a finite proportion of IBD patients. More clear and meaningful research has not been found. This may be due to the genetic complexity associated with IBD and other environmental factors. In the future, with the advancement of clinical trials, LBP is believed to have great opportunities in the treatment of IBD.

Future drug development may consider multiple regulatory points and therapeutic approaches based on synthetic biology. To sum up, there are both opportunities and challenges for synthetic biology in the therapy of IBD. The challenges for the synthetic biotechnology-based therapeutic approaches are: 1) the deficiency in the study of the etiology mechanism of IBD; 2) the restriction of biotechnological implementation such as the completion of the related gene database; 3) the safety evaluation of live bacteria; 4) the efficacy and stability for long-term use including the passage cultivation analysis which is also an important index to evaluate whether the strain meets the needs of subsequent industrialization; and 5) the regulation of the novel drugs. There will still be a long way to keep on moving.

Author contributions

YD and TX conceived of the idea and wrote the main part of the manuscript. GX and ZH participated in revising the manuscript and formatting the references. They wrote and revised the manuscript together under the guidance of JC. All authors contributed to the manuscript and approved the submitted version.

Funding

This work was supported by the Beijing Natural Science Foundation (L202045).

Conflict of interest

Author YD is employed by Suzhou U-Synbio Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Albillos, A., Gottardi, A. D., and Rescigno, M. (2019). The gut-liver axis in liver disease: Pathophysiological basis for therapy. *J. Hepatology* 72 (3), 558–577. doi:10.1016/j.jhep.2019.10.003

Ananthakrishnan, A. N., Bernstein, C. N., Iliopoulos, D., Macpherson, A., Neurath, M. F., Ali, R. A., et al. (2018). Environmental triggers in IBD: A review of progress and evidence. *Nat. Rev. Gastroenterol. Hepatol.* 15 (1), 39–49. doi:10.1038/nrgastro.2017.136

Anderson, J. L., Edney, R. J., and Whelan, K. (2012). Systematic review: Faecal microbiota transplantation in the management of inflammatory bowel disease. *Aliment. Pharmacol. Ther.* 36 (6), 503–516. doi:10.1111/j.1365-2036.2012.05220.x

Ardizzone, S., Cassinotti, A., and de Franchis, R. (2012). Immunosuppressive and biologic therapy for ulcerative colitis. *Expert Opin. Emerg. Drugs* 17 (4), 449–467. doi:10.1517/14728214.2012.744820

Atarashi, K., Tanoue, T., Oshima, K., Suda, W., Honda, K., Nishikawa, H., et al. (2014). Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* 500 (7461), 232–236. doi:10.1038/nature12331

Ballal, S. A., Veiga, P., Fenn, K., Michaud, M., Kim, J. H., Gallini, C. A., et al. (2015). Host lysozyme-mediated lysis of Lactococcus lactis facilitates delivery of colitis-attenuating superoxide dismutase to inflamed colons. *Proc. Natl. Acad. Sci. U. S. A.* 112 (25), 7803–7808. doi:10.1073/pnas.1501897112

Bar-Meir, S., Chowers, Y., Lavy, A., Abramovitch, D., Sternberg, A., Leichtmann, G., et al. (1998). Budesonide versus prednisone in the treatment of active Crohn's disease. *Gastroenterology* 115 (4), 835–840. doi:10.1016/s0016-5085(98)70254-9

Baron, S., Turck, D., Leplat, C., Merle, V., Gower-Rousseau, C., Marti, R., et al. (2005). Environmental risk factors in paediatric inflammatory bowel diseases: A population based case control study. *Gut* 54 (3), 357–363. doi:10.1136/gut.2004.054353

Barr, J. J., Auro, R., Furlan, M., Whiteson, K. L., Erb, M. L., Pogliano, J., et al. (2013). Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc. Natl. Acad. Sci. U. S. A.* 110 (26), 10771–10776. doi:10.1073/pnas.1305923110

Borody, T. J., and Khoruts, A. (2012). Fecal microbiota transplantation and emerging applications. *Nat. Rev. Gastroenterol. Hepatol.* 9 (2), 88–96. doi:10.1038/nrgastro.2011.244

Browning, M. G., Pessoa, B. M., Khoraki, J., and Campos, G. M. (2019). Changes in bile acid metabolism, transport, and signaling as central drivers for metabolic improvements after bariatric surgery. *Curr. Obes. Rep.* 8, 175–184. doi:10.1007/s13679-019-00334-4

Cai, C., Chen, D. Z., Ge, L. C., Chen, W. K., Ye, S., Ye, W. W., et al. (2019). Synergistic effects of Lactobacillus rhamnosus culture supernatant and bone marrow mesenchymal stem cells on the development of alcoholic steatohepatitis in mice. *Am. J. Transl. Res.* 11 (9), 5703–5715.

Calmus, Y., Guechot, J., Podevin, P., Bonnefis, M. T., Giboudeau, J., Poupon, R., et al. (2010). Differential effects of chenodeoxycholic and ursodeoxycholic acids on interleukin 1, interleukin 6 and tumor necrosis factor-alpha production by monocytes. *Hepatology* 16 (3), 719–723. doi:10.1002/hep.1840160317

Chen, Y., Wang, C., Liu, Y., Tang, L., Zheng, M., Xu, C., et al. (2013). miR-122 targets NOD2 to decrease intestinal epithelial cell injury in Crohn's disease. *Biochem. Biophys. Res. Commun.* 438 (1), 133–139. doi:10.1016/j.bbrc.2013.07.040

Christos, P., Daniel, W., and Hommes, T. (2015). MicroRNA214 is associated with progression of ulcerative colitis, and inhibition reduces development of colitis and colitis-associated cancer in mice - ScienceDirect. *Gastroenterology* 149 (4), 981–992. e911. doi:10.1053/j.gastro.2015.05.057

Chuang, A. Y., Chuang, J. C., Zhai, Z., Wu, F., and Kwon, J. H. (2014). NOD2 expression is regulated by microRNAs in colonic epithelial HCT116 cells. *Inflamm. Bowel Dis.* 1, 126–135. doi:10.1097/01.mib.0000436954.70596.9b

Collins, M., Lawson, P. A., Willems, A., Cordoba, J. J., Fe Rnandez-Garayzabal, J., Garcia, P., et al. (1994). The phylogeny of the genus Clostridium: Proposal of five new genera and eleven new species combinations. *Int. J. Syst. Bacteriol.* 44 (4), 812–826. doi:10.1099/00207713-44-4-812

Danese, S., Vuitton, L., and Peyrin-Biroulet, L. (2015). Biologic agents for IBD: Practical insights. *Nat. Rev. Gastroenterol. Hepatol.* 12 (9), 537–545. doi:10.1038/nrgastro.2015.135

Derrien, M., Vaughan, E. E., Plugge, C. M., and de Vos, W. M. (2004). Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* 54, 1469–1476. doi:10.1099/ijs.0.02873-0

Duerkop, B. A., Kleiner, M., Paez-Espino, D., Zhu, W., Bushnell, B., Hassell, B., et al. (2018). Murine colitis reveals a disease-associated bacteriophage community. *Nat. Microbiol.* 3, 1023–1031. doi:10.1038/s41564-018-0210-y

FAO/WHO (2002). *Guidelines for the evaluation of probiotics in food. Report of a joint FAO/WHO working group on drafting guidelines for the evaluation of probiotics in food*. London (Ontario, Canada): World Health Organization.

FDA (2016). *Early clinical trials with live biotherapeutic products: Chemistry, manufacturing, and control information: Guidance for industry*. Maryland: FDA.

Felwick, R. K., Dingley, G., Martinez-Nunez, R., Sanchez-Elsner, T., and Collins, J. E. (2019). MicroRNA23a overexpression in Crohn's disease targets tumour necrosis factor Alpha inhibitor protein 3, increasing sensitivity to TNF and modifying the epithelial barrier. *J. Crohn's. Colitis* 14 (3), 381–392. doi:10.1093/ecco-jcc/jjz145

Feng, Wu, Michelle, Z., and Themistocles, D. (2008)., 135. Gastroenterology, 1624. doi:10.1053/j.gastro.2008.07.068MicroRNAs are differentially expressed in ulcerative colitis and alter expression of macrophage inflammatory peptide-2α

Filip, A. T., Balacescu, O., Marian, C., and Anghel, A. (2016). Microbiota small RNAs in inflammatory bowel disease. *J. Gastrointestin. Liver Dis.* 25 (4), 509–516. doi:10.15403/jgld.2014.1121.254.lip

Frank, D. N., Amand, A., Feldman, R. A., Boedeker, E. C., Harpaz, N., Pace, N. R., et al. (2007). Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. U. S. A.* 104 (34), 13780–13785. doi:10.1073/pnas.0706625104

Greenhill, C. (2014). Ibd: Glucocorticoids revealed to augment intestinal epithelial barrier function. *Nat. Rev. Gastroenterol. Hepatol.* 11 (2), 75. doi:10.1038/nrgastro.2013.255

Guan, Q., and Zhang, J. (2017). Recent advances: The imbalance of cytokines in the pathogenesis of inflammatory bowel disease. *Mediat. Inflamm.* 2017, 1–8. doi:10.1155/2017/4810258

Heinsbroek, S., Squadrito, M. L., Schilderink, R., Hilbers, F., and Jonge, W. D. (2015). miR-511-3p, embedded in the macrophage mannose receptor gene, contributes to intestinal inflammation. *Gastroenterology* 9, 960. doi:10.1038/mi.2015.113

Herrera Estrada, L., Wu, H., Ling, K., Zhang, G., Sumagin, R., Parkos, C. A., et al. (2017). Bioengineering bacterially derived immunomodulants: A therapeutic approach to inflammatory bowel disease. *ACS Nano* 11 (10), 9650–9662. doi:10.1021/acsnano.7b03239

Hou, J., Chen, B., Hu, X., Chen, X., Zhao, L., Chen, Z., et al. (2017)., 213. S0344033817303588, 1289.miR-155 targets Est-1 and induces ulcerative colitis via the IL-23/17/6-mediated Th17 pathway*Pathology Res. Pract.*

Huang, Z., Ma, J., Chen, M., Jiang, H., Fu, Y., Gan, J., et al. (2015). Dual TNF-α/IL-12p40 interference as a strategy to protect against colitis based on miR-16 precursors with macrophage targeting vectors. *Mol. Ther.* 23, 1611–1621. doi:10.1038/mt.2015.111

Huang, Z., Shi, T., Zhou, Q., Shi, S., Zhao, R., Shi, H., et al. (2014). miR-141 Regulates colonic leukocytic trafficking by targeting CXCL12β during murine colitis and human Crohn's disease. *Gut* 63 (8), 1247–1257. doi:10.1136/gutjnl-2012-304213

Ianiro, G., Bibbò, S., Scaldaferri, F., Gasbarrini, A., and Cammarota, G. (2014). Fecal microbiota transplantation in inflammatory bowel disease. *Medicine* 93 (19), e97. doi:10.1097/md.0000000000000097

Jensen, T. L., Gøtzsche, C. R., and Woldbye, D. P. D. (2021). Current and future prospects for gene therapy for rare genetic diseases affecting the brain and spinal cord. *Front. Mol. Neurosci.* 14, 695937. doi:10.3389/fnmol.2021.695937

Jung, K. J., Lee, G. W., Park, C. H., Lee, T. J., Kim, J. Y., Sung, E. G., et al. (2020). Mesenchymal stem cells decrease oxidative stress in the bowels of interleukin-10 knockout mice. *Gut Liver* 14, 100–107. doi:10.5009/gnl18438

Kane, S. V., Corradino, S., Cho, J. H., and Hanauer, S. B. (2000). Smoking is associated with disease type in families with both Crohn's disease (CD) and ulcerative colitis (UC). *Gastroenterology* 118 (4), A333. doi:10.1016/s0016-5085(00)83428-9

Kaplan, G. G. (2015). The global burden of IBD: From 2015 to 2025. *Nat. Rev. Gastroenterol. Hepatol.* 12, 720–727. doi:10.1038/nrgastro.2015.150

Keitel, V., Donner, M., Winandy, S., Kubitz, R., and H?Ussinger, D. (2008). Expression and function of the bile acid receptor TGR5 in Kupffer cells. *Biochem. Biophys. Res. Commun.* 372 (1), 78–84. doi:10.1016/j.bbrc.2008.04.171

Ko, I. K., Kim, B. G., Awadallah, A., Mikulan, J., Lin, P., Letterio, J. J., et al. (2010). Targeting improves MSC treatment of inflammatory bowel disease. *Mol. Ther.* 18 (7), 1365–1372. doi:10.1038/mt.2010.54

Kotlowski, R. (2016). Use of *Escherichia coli* Nissle 1917 producing recombinant colicins for treatment of IBD patients. *Med. Hypotheses* 93, 8–10. doi:10.1016/j.mehy.2016.05.002

Lamas, B., Richard, M. L., Leducq, V., Pham, H. P., Michel, M. L., Costa, G. D., et al. (2016). CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nat. Med.* 22, 598–605. doi:10.1038/nm.4102

Lange, K., Moutsianas, L., Lee, J. C., Lamb, C. A., Barrett, J. C., Kennedy, N. A., et al. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat. Genet.* 49 (2), 256–261. doi:10.1038/ng.3760

Lavelle, A., and Sokol, H. (2020). Gut microbiota-derived metabolites as key actors in inflammatory bowel disease. *Nat. Rev. Gastroenterology Hepatology* 17, 223. doi:10.1038/s41575-019-0258-z

Lee, M., and Chang, E. (2020). Inflammatory bowel diseases (IBD) and the microbiome-searching the crime scene for clues. Gastroenterology 160, 524. doi:10.1053/j.gastro.2020.09.056

Lewis, J., hen, E. C., Baldassano, R., Otley, A., Griffiths, A., Lee, D., et al. (2015). Inflammation, antibiotics, and diet as environmental stressors of the gut microbiome in pediatric Crohn's disease. *Cell Host Microbe* 18 (4), 489–500. doi:10.1016/j.chom.2015.09.008

Liu, S., Zhang, S., Lv, X., Lu, J., Chen, Y., Zeng, Z., et al. (2019). Limonin ameliorates ulcerative colitis by regulating STAT3/miR-214 signaling pathway. *Int. Immunopharmacol.* 75, 105768. doi:10.1016/j.intimp.2019.105768

Ludwig, K., Fassan, M., Mescoli, C., Pizzi, M., Balistreri, M., Albertoni, L., et al. (2013). PDCD4/miR-21 dysregulation in inflammatory bowel disease-associated carcinogenesis. *Virchows Arch.* 462, 57–63. doi:10.1007/s00428-012-1345-5

Luo, X., Reiter, M. A., d'Espaux, L., Wong, J., Denby, C. M., Lechner, A., et al. (2019). Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* 567 (7746), 123–126. doi:10.1038/s41586-019-0978-9

Ma, D., Cao, Y., Wang, Z., He, J., Chen, H., Xiong, H., et al. (2019). CCAT1 lncRNA promotes inflammatory bowel disease malignancy by destroying intestinal barrier via downregulating miR-185-3p. *Inflamm. Bowel Dis.* 25, 862. doi:10.1093/ibd/izy381

Marcin, G., Marzena, S-Z., Marcin, A., and Dobrowolska, A. (2016). The impact of genetic factors on response to glucocorticoids therapy in IBD. *Scand. J. Gastroenterology* 51 (6), 654–665. doi:10.3109/00365521.2015.1132336

Marilidia, A., Piglionica, M., and Cariello, M. (2018). The gut-liver axis in hepatocarcinoma: A focus on the nuclear receptor FXR and the enterokine FGF19. *Curr. Opin. Pharmacol.* 43, 93–98. doi:10.1016/j.coph.2018.08.005

Markus, G., Rocio, M. N., Simon, W., Victor, B., Andy, C., Jane, C., et al. (2018). MicroRNA-31 and MicroRNA-155 are overexpressed in ulcerative colitis and regulate IL-13 signaling by targeting interleukin 13 receptor α-1. *Genes* 9 (2), 85. doi:10.3390/genes9020085

Matthieu, G., De, S. L., Adeline, S., Amélie, d. V., Damien, M., Christel, N., et al. (2017). Bacteriophages targeting adherent invasive *Escherichia coli* strains as a promising new treatment for Crohn's disease. *J. Crohns Colitis* 11 (7), 840–847. doi:10.1093/ecco-jcc/jjw224

Mcfarland, L. V. (2010). Systematic review and meta-analysis of Saccharomyces boulardii in adult patients. *World J. Gastroenterol.* 18, 2202. doi:10.3748/wjg.v16.i18.2202

Mendlovic, F., Cruz-Rivera, M., Diaz-Gandarilla, J. A., Flores-Torres, M. A., Avila, G., Perfiliev, M., et al. (2017). Orally administered Taenia solium Calreticulin prevents experimental intestinal inflammation and is associated with a type 2 immune response. *Plos One* 12 (10), e0186510. (Accession No. 29036211). doi:10.1371/journal.pone.0186510

Meng, F., and Lowell, C. A. (1997). Lipopolysaccharide (LPS)-induced macrophage activation and signal transduction in the absence of Src-family kinases Hck, Fgr, and Lyn. *J. Exp. Med.* 185 (9), 1661–1670. doi:10.1084/jem.185.9.1661

Mills, E., Kelly, B., Logan, A., Costa, A. H., Varma, M., Bryant, C., et al. (2016). Succinate dehydrogenase supports metabolic repurposing of mitochondria to drive inflammatory macrophages. *Cell* 167, 457–470.e13. doi:10.1016/j.cell.2016.08.064

Min, M., Peng, L., Yang, Y., Guo, M., Wang, W., Gang, S., et al. (2014). MicroRNA-155 is involved in the pathogenesis of ulcerative colitis by targeting FOXO3a. *Inflamm. Bowel Dis.* 20 (4), 652–659. doi:10.1097/mib.0000000000000009

Molendijk, I., Hommes, D. W., and Duijvestein, M. (2013). *Mesenchymal stromal cell therapy in Crohn's disease*. New Jersey: Humana Press.

Motta, J. P., Bermudez-Humaran, L. G., Deraison, C., Martin, L., Rolland, C., Rousset, P., et al. (2012). Food-grade bacteria expressing elafin protect against inflammation and restore colon homeostasis. *Sci. Transl. Med.* 4 (158), 158ra144. doi:10.1126/scitranslmed.3004212

Norman, J. M., Handley, S. A., Baldridge, M. T., Droit, L., Liu, C. Y., Keller, B. C., et al. (2015). Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* 160, 447–460. doi:10.1016/j.cell.2015.01.002

Ohkuro, M., Kim, J. D., Kuboi, Y., Hayashi, Y., Fukamizu, A., Kobayashi-Kuramochi, H., et al. (2018). Calreticulin and integrin alpha dissociation induces anti-inflammatory programming in animal models of inflammatory bowel disease. *Nat. Commun.* 9 (1), 1982. doi:10.1038/s41467-018-04420-4

Oka, A., and Sartor, R. B. (2020). Microbial-based and microbial-targeted therapies for inflammatory bowel diseases. *Dig. Dis. Sci.* 65, 757–788. doi:10.1007/s10620-020-06090-z

Olaru, A. V., Selaru, F. M., Yuriko, M., Christine, V., Stefan, D., Bogdan, P., et al. (2015). Dynamic changes in the expression of MicroRNA-31 during inflammatory bowel disease-associated neoplastic transformation. *Inflamm. Bowel Dis.* 17 (1), 221–231. doi:10.1002/ibd.21359

Ouwerkerk, J. P., Van, D., Davids, M., Claassens, N. J., Finestra, T. R., Vos, W. D., et al. (2016). Adaptation of Akkermansia muciniphila to the oxic-anoxic interface of the mucus layer. *Appl. Environ. Microbiol.* 82 (23), 6983–6993. doi:10.1128/aem.01641-16

Qiao, Y., Zhang, S., Kang, C., Rui, F., Chen, M., Li, M., et al. (2016). E3 ubiquitin ligase RNF183 is a novel regulator in inflammatory bowel disease. *J. Crohn's Colitis* 10 (6), 713–725. doi:10.1093/ecco-jcc/jjw023

Qv, L., Mao, S., Li, Y., Zhang, J., and Li, L. (2021). Roles of gut bacteriophages in the pathogenesis and treatment of inflammatory bowel disease. *Front. Cell. Infect. Microbiol.* 11, 755650. doi:10.3389/fcimb.2021.755650

Riccardi, L., Mazzon, E., Bruscoli, S., Esposito, E., Cuzzocrea, S., Di Paola, R., et al. (2008). Peroxisome proliferator-activated receptor-α modulates the anti-inflammatory effect of glucocorticoids in a model of inflammatory bowel disease in mice. *Shock (Augusta, Ga.)* 31 (3), 308–316. doi:10.1097/shk.0b013e31818339e7

Rogler, G., and Andus, T. (1998). Cytokines in inflammatory bowel disease. *World J. Surg.* 22 (4), 382–389. doi:10.1007/s002689900401

Rui-Cong, C., Lan-Man, X., Shan, J., Wu, D., Huang, S. S., Wu, H., et al. (2016). Lactobacillus rhamnosus GG supernatant promotes intestinal barrier function, balances Treg and TH17 cells and ameliorates hepatic injury in a mouse model of chronic-binge alcohol feeding. *Toxicol. Lett.* 241, 103–110. doi:10.1016/j.toxlet.2015.11.019

Schauer, C., Avery, V., Seleq, S., Garg, P., Wang, M. T. M., Chieng, M., et al. (2021). A comparison of intravenous methylprednisolone and hydrocortisone for the treatment of acute inflammatory bowel disease. *J. Gastroenterol. Hepatol.* 36 (10), 2762–2768. doi:10.1111/jgh.15535

Schultz, M., and Butt, G. (2010). *E. coli* Nissle 1917 in the treatment of inflammatory bowel disease. *Pract. Gastroenterol.* 34 (3), 11–19. doi:10.1039/D2FO00226D

Scott, B. M., Gutiérrez-Vázquez, C., Sanmarco, L. M., Pereira, J., Quintana, F. J., Plasencia, A., et al. (2021). Self-tunable engineered yeast probiotics for the treatment of inflammatory bowel disease. *Nat. Med.* 27, 1212–1222. doi:10.1038/s41591-021-01390-x

Sen, S., and Mansell, T. J. (2020). Yeasts as probiotics: Mechanisms, outcomes, and future potential. *Fungal Genet. Biol.* 137, 103333. doi:10.1016/j.fgb.2020.103333

Seregin, S. S., Golovchenko, N., Schaf, B., Chen, J., Pudlo, N. A., Mitchell, J., et al. (2017). NLRP6 protects Il10/mice from colitis by limiting colonization of Akkermansia muciniphila. *Cell Rep.* 19 (10), 2174. doi:10.1016/j.celrep.2017.05.074

Shan, Y., Lee, M., and Chang, E. B. (2022). The gut microbiome and inflammatory bowel diseases. *Annu. Rev. Med.* 73 (1), 455–468. doi:10.1146/annurev-med-042320-021020)

Shi, Xie, Fu, Z., and Ma, J. (2017). The signaling axis of microRNA-31/interleukin-25 regulates Th1/Th17-mediated inflammation response in colitis. *Mucosal Immunol.* 10, 983. doi:10.1038/mi.2016.102

Singh, A. K., and Jiang, Y. (2003). Lipopolysaccharide (LPS) induced activation of the immune system in control rats and rats chronically exposed to a low level of the organothiophosphate insecticide, acephate. *Toxicol. Ind. Health* 19 (2-6), 93–108. doi:10.1191/0748233703th181oa

Song, L., Chang, R., Sun, X., Lu, L., Zhan, L., Lu, H., et al. (2021). Macrophage-derived EDA-A2 inhibits intestinal stem cells by targeting miR-494/EDA2R/β-catenin signaling in mice. *Commun. Biol.* 4 (1), 213. doi:10.1038/s42003-021-01730-0

Steidler, L., Hans, W., Schotte, L., Neirynck, S., Obermeier, F., Falk, W., et al. (2000). Treatment of murine colitis by Lactococcus lactis secreting Interleukin-10. *Science* 289, 1352–1355. doi:10.1126/science.289.5483.1352

Stojanov, S., Berlec, A., and Štrukelj, B. (2020). The influence of probiotics on the firmicutes/bacteroidetes ratio in the treatment of obesity and inflammatory bowel disease. *Microorganisms* 8 (11), 1715. doi:10.3390/microorganisms8111715

SvenjaMichael, L., Keubler, A., and Smoczek, F (2013). Quantitative phenotyping of inflammatory bowel disease in the IL-10-deficient mouse by use of noninvasive magnetic resonance imaging. *Inflamm. Bowel Dis.* 19, 185. doi:10.1002/ibd.23006

Talwadekar, M. D., Kale, V. P., and Limaye, L. S. (2015). Placenta-derived mesenchymal stem cells possess better immunoregulatory properties compared to their cord-derived counterparts–a paired sample study. *Sci. Rep.* 5 (1), 15784. doi:10.1038/srep15784

Tian, T., Zhou, Y., Feng, X., Ye, S., Wang, H., Wu, W., et al. (2016). MicroRNA-16 is putatively involved in the NF-κB pathway regulation in ulcerative colitis through adenosine A2a receptor (A2aAR) mRNA targeting. *Sci. Rep.* 6, 30824. doi:10.1038/srep30824

Tian, Y., Xu, J., Li, Y., Zhao, R., Du, S., Lv, C., et al. (2019). MicroRNA-31 reduces inflammatory signaling and promotes regeneration in colon epithelium, and delivery of mimics in microspheres reduces colitis in mice. *Gastroenterology* 156, 2281–2296.e6. doi:10.1053/j.gastro.2019.02.023

TingZhang, Q., Li, L., Faming, C., Buch, H., and Zhang, F. (2019). Akkermansia muciniphila is a promising probiotic. *Microb. Biotechnol.* 12, 1109–1125. doi:10.1111/1751-7915.13410

Ukena, S. N., Singh, A., Dringenberg, U., Engelhardt, R., Seidler, U., Hansen, W., et al. (2007). Probiotic *Escherichia coli* Nissle 1917 inhibits leaky gut by enhancing mucosal integrity. *Plos One* 2 (12), e1308. doi:10.1371/journal.pone.0001308

Vanhoutvin, S., Troost, F. J., Hamer, H. M., Lindsey, P. J., Koek, G. H., Jonkers, D., et al. (2009). Butyrate-induced transcriptional changes in human colonic mucosa. *Plos One* 4, e6759. doi:10.1371/journal.pone.0006759

Wang, H., Vilches-Moure, J. G., Cherkaoui, S., Tardy, I., Alleaume, C., Bettinger, T., et al. (2019). Chronic model of inflammatory bowel disease in IL-10-/- transgenic mice: Evaluation with ultrasound molecular imaging. *Theranostics* 9 (21), 6031–6046. doi:10.7150/thno.37397

Wang, J., Tian, M., Li, W., and Hao, F. (2019). Preventative delivery of IL-35 by Lactococcus lactis ameliorates DSS-induced colitis in mice. *Appl. Microbiol. Biotechnol.* 103 (19), 7931–7941. doi:10.1007/s00253-019-10094-9

Wang, L., Tang, L., Feng, Y., Zhao, S., Han, M., Zhang, C., et al. (2020). A purified membrane protein from Akkermansia muciniphila or the pasteurised bacterium blunts colitis associated tumourigenesis by modulation of CD8+ T cells in mice. *Gut* 69, 1988–1997. gutjnl-2019-320105. doi:10.1136/gutjnl-2019-320105

Wang, Y. D., Chen, W. D., Yu, D., Forman, B. M., and Huang, W. (2011). The G-Protein-coupled bile acid receptor, Gpbar1 (TGR5), negatively regulates hepatic inflammatory response through antagonizing nuclear factor kappa light-chain enhancer of activated B cells (NF-κB) in mice. *Hepatology* 54, 1421–1432. doi:10.1002/hep.24525

Watanabe, K., Ohira, H., Orikasa, H., Saito, K., Kanno, K., Shioya, Y., et al. (2006). Anti-calreticulin antibodies in patients with inflammatory bowel disease. *Fukushima J. Med. Sci.* 52 (1), 1–11. doi:10.5387/fms.52.1

Weber, C. (2015). Ibd: Lactococcus lactis alleviates oxidative stress and colitis in mice. *Nat. Rev. Gastroenterol. Hepatol.* 12 (8), 429. doi:10.1038/nrgastro.2015.109

Williams, C. N. (1994). Overview of 5-ASA in the therapy of IBD: Inflammatory bowel disease. *IBD – Ther.* 8. doi:10.1155/1994/756973

Wrzosek, L., Miquel, S., Noordine, M.-L., Bouet, S., Joncquel Chevalier-Curt, M., Robert, V., et al. (2013). Bacteroides thetaiotaomicron and Faecalibacterium prausnitzii influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent. *BMC Biol.* 11, 61. doi:10.1186/1741-7007-11-61

Wu, W., He, C., Liu, C., Cao, A. T., Xue, X., Evans-Marin, H. L., et al. (2015). miR-10a inhibits dendritic cell activation and Th1/Th17 cell immune responses in IBD. *Gut* 64 (11), 1755–1764. doi:10.1136/gutjnl-2014-307980

Xinyi, W., Wan, F., Volpetti, E., Petrova, J., Maerkl, S. J., and Wang, B. (2019). Cascaded amplifying circuits enable ultrasensitive cellular sensors for toxic metals. *Nat. Chem. Biol.* 15, 540–548. doi:10.1038/s41589-019-0244-3

Xue, Y., He, Q., Guo, Z., Fei, X., Yi, L., Yan, P., et al. (2018). MicroRNA-425 facilitates pathogenic Th17 cell differentiation by targeting forkhead box O1 (Foxo1) and is associated with inflammatory bowel disease. *Biochem. Biophys. Res. Commun.* 496 (2), 352–358. doi:10.1016/j.bbrc.2018.01.055

Yang, Y., Ma, Y., Shi, C., Chen, H., Zhang, H., Chen, N., et al. (2013). Overexpression of miR-21 in patients with ulcerative colitis impairs intestinal epithelial barrier function through targeting the Rho GTPase RhoB. *Biochem. Biophys. Res. Commun.* 434 (4), 746–752. doi:10.1016/j.bbrc.2013.03.122

Yin, J., Yang, L., Mou, L., Dong, K., Ye, H., Xue, S., et al. (2019). A green tea-triggered genetic control system for treating diabetes in mice and monkeys. *Sci. Transl. Med.* 11 (515), eaav8826. eaav8826-. doi:10.1126/scitranslmed.aav8826

Yosuke, S., Makoto, K., Kazuo, H., Masafumi, M., Kazunori, M., Hisako, K., et al. (2013). Commensal bacteria-dependent indole production enhances epithelial barrier function in the colon. *Plos One* 8 (11), e80604. doi:10.1371/journal.pone.0080604

Zenewicz, L. A., Yancopoulos, G. D., Valenzuela, D. M., Murphy, A. J., Stevens, S., Flavell, R. A., et al. (2008). Innate and adaptive interleukin-22 protects mice from inflammatory bowel disease. *Immunity* 29 (6), 947–957. doi:10.1016/j.immuni.2008.11.003

Zeng, B., Shi, S., Ashworth, G., Dong, C., Liu, J., and Xing, F. (2019). ILC3 function as a double-edged sword in inflammatory bowel diseases. *Cell Death & Disease* 10, 315. doi:10.1038/s41419-019-1540-2

Zhai, Q., Gong, X., Wang, C., Zhao, J., Zhang, H., Tian, F., et al. (2019a). Food-borne patulin toxicity is related to gut barrier disruption and can be prevented by docosahexaenoic acid and probiotic supplementation. *Food Funct.* 10, 1330–1339. doi:10.1039/c8fo02292e

Zhai, R., Xue, X., Zhang, L., Yang, X., and Zhang, C. (2019b). Strain-specific anti-inflammatory properties of two Akkermansia muciniphila strains on chronic colitis in mice. *Front. Cell. Infect. Microbiol.* 9, 239. doi:10.3389/fcimb.2019.00239

Zhang, L., Shen, J., Cheng, J., and Fan, X. (2015). MicroRNA-21 regulates intestinal epithelial tight junction permeability. *Cell biochem. Funct.* 33 (4), 235–240. doi:10.1002/cbf.3109

Zhang, T., Ji, X., Lu, G., and Zhang, F. (2021). The potential of Akkermansia muciniphila in inflammatory bowel disease. *Appl. Microbiol. Biotechnol.* 105 (4), 5785–5794. doi:10.1007/s00253-021-11453-1

Zwa, B., Kai, G. C., Yl, B., Cha, D., and Min, W. B. (2021). Dynamic impact of virome on colitis and colorectal cancer: Immunity, inflammation, prevention and treatment - ScienceDirect. *Seminars Cancer Biol.* doi:10.1016/j.semcancer.2021.10.004

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Review
PUBLISHED 23 August 2022
DOI 10.3389/fbioe.2022.968437

OPEN ACCESS

EDITED BY
Shota Atsumi,
University of California, Davis,
United States

REVIEWED BY
Jian Wang,
University of Georgia, United States
Yinjie Tang,
Washington University in St. Louis,
United States

*CORRESPONDENCE
Taek Soon Lee,
tslee@lbl.gov

SPECIALTY SECTION
This article was submitted to Synthetic Biology,
a section of the journal
Frontiers in Bioengineering and Biotechnology

RECEIVED 14 June 2022
ACCEPTED 21 July 2022
PUBLISHED 23 August 2022

CITATION
Carruthers DN and Lee TS (2022), Translating advances in microbial bioproduction to sustainable biotechnology.
Front. Bioeng. Biotechnol. 10:968437.
doi: 10.3389/fbioe.2022.968437

Translating advances in microbial bioproduction to sustainable biotechnology

David N. Carruthers[1,2] and Taek Soon Lee[1,2]*

[1]Joint BioEnergy Institute, Emeryville, CA, United States, [2]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States

Advances in synthetic biology have radically changed our ability to rewire microorganisms and significantly improved the scalable production of a vast array of drop-in biopolymers and biofuels. The success of a drop-in bioproduct is contingent on market competition with petrochemical analogues and weighted upon relative economic and environmental metrics. While the quantification of comparative trade-offs is critical for accurate process-level decision making, the translation of industrial ecology to synthetic biology is often ambiguous and assessment accuracy has proven challenging. In this review, we explore strategies for evaluating industrial biotechnology through life cycle and techno-economic assessment, then contextualize how recent developments in synthetic biology have improved process viability by expanding feedstock availability and the productivity of microbes. By juxtaposing biological and industrial constraints, we highlight major obstacles between the disparate disciplines that hinder accurate process evaluation. The convergence of these disciplines is crucial in shifting towards carbon neutrality and a circular bioeconomy.

KEYWORDS

synthetic biology, bioproduction, life cycle assessment (LCA), techno-economic assessment (TEA), bioeconomy

Introduction

Industrial decarbonization is crucial for combating anthropogenic climate change. Recent assessments published by the Intergovernmental Panel on Climate Change have inspired international pledges for global carbon emissions reduction that generally aim to mitigate global warming to 1.5°C as per the Paris Accords (Masson-Delmotte et al., 2018; Hoegh-Guldberg et al., 2019). Despite near universal recognition of the deleterious effects of climate change, pledges for emissions reduction, and ratification of the UN Sustainable Development Goals, actualizing climate commitments has proven extraordinarily challenging in part due to the global inertia of a fossil-based economy (Peters et al., 2019). Today, approximately 90% of chemicals are generated from fossil fuels as carbon and energy feedstocks (Blank et al., 2020).

Recent proposals have advocated shifting material flows towards a circular economy through upcycling of waste streams, increasing efficiency, and shifting raw material acquisition to biobased pathways. A major advantage of bioproducts is the ease with

which they may be substituted into traditional petrochemical systems as "drop-in" alternatives. Drop-in bioproducts have minimal impact on infrastructure, contemporary behavior, or sociopolitical attitudes and often maintain lower carbon emissions. The substitution of petrochemicals with biobased alternatives could reduce fuel associated greenhouse gas (GHG) emissions by 75–80% (Balakrishnan et al., 2015) and further substitution of bio-based plastics could reduce plastic production GHG emissions by 66% (Spierling et al., 2018; Zheng and Suh, 2019).

Metabolic engineering is a promising strategy for the delineation and decarbonization of plastics and fuels by supplanting high volume chemical production with biosynthetic routes. The last decade has witnessed a veritable revolution in synthetic biology, characterized by the relative ease with which microbes can be modified and optimized for high titer bioproduction via the canonical design-build-test-learn (DBTL) cycle, utterly transforming our ability to engineer complex biosynthetic pathways (Cameron et al., 2014; Robinson et al., 2020). Perhaps most remarkable is the horizontal expansion of metabolic pathways to generate a diverse portfolio of chemicals with the propensity to disrupt the traditional manufacture of food, fuels, materials, and medicines (Voigt, 2020). Recent synthetic biology advances like genome editing [CRISPR(Sander and Joung, 2014) (Kim J. et al., 2021); CRISPRi (Wu et al., 2017) (Banerjee et al., 2020)], adaptive laboratory evolution (Mohamed et al., 2020), the explosion of computationally informed protein folding prediction software (Jumper et al., 2021), and advanced library screening technologies (Saleski et al., 2021) have cumulatively improved sustainable bioproduction platforms. Specifically, metabolically engineered microbes have demonstrated generation of food supplements (Torres-Tiji et al., 2020), biofuels for blendstocks [e.g., limonene (Alonso-Gutierrez et al., 2013), isoprenol (Kim J. et al., 2021), isobutanol (Saleski et al., 2021)], biopolymers [e.g., PHA (Aurisano et al., 2021)], and bulk chemical precursors [e.g., adipic acid, lactic acid, L-lysine 3-hydroxypropionic acid, among others (Bozell and Petersen, 2010)]. This array of bioproducts is complemented by diverse engineering strategies. The most common approach is a consolidated rewiring of a microbial chassis for enhanced bioproduction, though researchers have also pursued the integration of variable bioproduction routes within the same chassis (Lee and Kim, 2015) as well as overhauling carbon assimilation to diversify substrate availability (Kim et al., 2020). At a community level, investigations of the rhizosphere have helped to identify constitutive microbes, nutrient exchanges, and microbial interactions that may improve community-level production, stability, and nutrient recycling thereby reducing material inputs (Ahkami et al., 2017; Lawson et al., 2019).

Despite significant improvements in pathway engineering, few platforms have proven commercially viable due in part to the lack of accessible economic assessments that address early bottlenecks and adequately scale laboratory cultures to pilot settings (Lynch, 2021). This deficiency is coupled with the difficult task of actually determining whether a bioproduction platform is more sustainable than current practices, which historically is not necessarily the case (Searchinger et al., 2008; DeCicco et al., 2016). Thorough accounting requires not only the quantification of cost, energy, and emissions throughout the process, but the inclusion of nontrivial yet ambiguous inputs like, for example, land-use-change, which considers how a production platform displaces native habitat. And, often most importantly, the purported sustainability of a process must align also with its commercial viability. Summarily, emerging bioproduction platforms are confronted with the formidable task of simultaneously approaching carbon neutrality while maintaining price parity with long established chemical syntheses (Lee et al., 2021).

Life cycle assessment (LCA) and techno-economic assessment (TEA) are the two dominant strategies for quantifying the relative environmental and economic costs of a production strategy. While TEAs have been methodologically applied throughout the industrial era, LCAs are relatively new and have matured from a Boolean choice between products to a multifaceted environmental report. In this review, we first describe how LCAs and TEAs are traditionally conducted to yield empirical metrics and their extrapolation to synthetic biology with emphasis on biofuels, biopolymers, and biochemical precursors of specific importance. We further elaborate on recent metabolic engineering of microorganisms for enhanced bioproduction, namely through the improvement of titer, rate, yield, and substrate utilization with specific attention to studies conducting LCAs/TEAs. Finally, we underline the necessity of assessment harmonization, accessibility, and transparency to better translate results between disciplines and, quintessentially, facilitate proper allocation of research investment for informed policy decision making.

Economic and environmental evaluation of bioproduction platforms

Both LCA and TEA aim to provide practitioners with useful indicators for improving the affordability or sustainability of a given process. Importantly, assessment accuracy is contingent on data reliability and availability, factors that can prove inhibitive for bioproduction platforms at low technology readiness levels (TRLs) (Guinée et al., 2011; Thomassen et al., 2019). Similarly, the integration of LCA and TEA has also been recognized as a major obstacle in assessing the combined economic and environmental burdens of specific processes (Mahmud et al., 2021). Here, we provide an overview of the traditional LCA, TEA,

and then explore how recent assessments have addressed the challenges of emerging bioproduction technologies.

Life cycle assessment

The nominal strategy for determining bioproduct sustainability is through an LCA, a technique that originally arose as a tool for product selection based on potential environmental impact (e.g., "paper or plastic?") (Guinée et al., 2011). Though the philosophy behind LCA has matured over the last several decades, the overall framework remains the same. A product life cycle consists of five key stages: raw material acquisition, manufacturing, packaging and transportation, use-phase, and end-of-life. The LCA framework is qualitatively outlined by ISO (International Standards Organization) 14040: 2006 and 14044:2006 with intentional ambiguity to enable its translation to various production pathways, inputs, impacts, and general assumptions (Standardization International Organization for, 2006). At its core, the framework is remarkably simple and consists of four principal components:

1. Scope and Goal: A system boundary and functional unit of production are allocated as a basis for process calculations. The system boundary dictates the overall processes to be considered and may extend from raw material acquisition to the final product ("cradle-to-gate" or "well-to-pump" for fuels) or to the use-phase and end-of-life ("cradle-to-grave" or "well-to-wheels" for fuels).
2. Life Cycle Inventory (LCI): The inventory accounts for the material and energy flows within the system boundary. Several LCI software packages have been developed to facilitate practitioner accounting and are especially valuable for multi-input multi-output mass and energy flows (Suh and Huppes, 2005). Many dedicated LCI packages have been developed, including the popular EcoInvent v3.1 (Wernet et al., 2016), which is used for commercial product generation. However, practitioners often choose more robust process engineering tools like Aspen Plus® and SuperPro Designer® that account for mass and energy flows with thermodynamic modeling.
3. Life Cycle Impact Assessment (LCIA): Next, specific impacts of interest are selected to measure the effects of the functional unit within the system by coupling environmental allocations to materials and energy sources. The impacts usually include specific, physical midpoint indicators like greenhouse gas emissions (CO_2eq), water intensity, and energy return, though may extend to environmental factors such as eutrophication potential, human toxicity, acidification, and direct or indirect land use change. LCIAs may also include more ambiguous endpoint indicators such as biodiversity loss, ecosystem damage, and human health effects through defined weighting of midpoint indicators (Souza et al., 2015). Common impact databases include Tool for Reduction and Assessment of Chemical and other environmental Impacts (TRACI) by the US EPA (Bare, 2011) and ReCiPe 2016 (Huijbregts et al., 2017) among others. Traditionally, bioproduction assessments utilize midpoint indicators to communicate the relative change in emissions and energy return between production systems. More recently, works have employed neural networks to better understand how different inventories characterize chemicals and thus predict impacts of novel chemical syntheses (Song et al., 2017).
4. Interpretation: Lastly, comparative impacts are weighed within the context of the system inventory and boundary to inform process selection, policy, and investment decisions (e.g., does the bioproduct have lower carbon intensity compared to its petrochemical analogue?).

The life cycle framework may also be subdivided into either an attributional LCA (aLCA) that allocates the environmental impacts of a production pathway to a given project or a consequential LCA (cLCA) that considers how environmental impacts change in response to product generation, the energy/material displaced by the product, and product demand (Earles and Halog, 2011). Each category has distinct importance for process interpretation and intercommunication between the four individual components is crucial for improving confidence in LCA results.

Several overarching software tools have been developed for LCA, including openLCA (https://openlca.org/) and SimaPro (https://simapro.com/). Such tools facilitate inventory construction and selection of impact indicators from a library of different methodologies. In the case of bioproducts, Greenhouse gasses, Regulated Emissions, and Energy in Transportation (Greet, 2022), a tool published by Argonne National Laboratory, is popularly applied to draft emerging bioproduction schemes including biofuels ranging from corn ethanol as a gasoline additive to algal biofuels for biodiesel (Argonne GREET Model). It has also been employed in conjunction with Aspen Plus® to evaluate bioproduction of 12 high performance platform chemicals (Dunn et al., 2015).

In general, the flexibility of ISO 14040 complements the diversity of bioproduction platforms, which generate not only disparate products, but utilize dramatically different feedstocks from climatically distinctive regions (e.g., Brazilian sugarcane vs. American corn). Practitioners therefore rely upon sensitivity analyses like Monte Carlo simulation and scenarios forecasting to elucidate inputs that disproportionately affect impacts (Patouillard et al., 2019). For LCA, scenarios may present different conversion rates, throughput, recycling fraction, or total yields (Krömer et al., 2020). Increasing emphasis has been placed on the inclusion of LCA in academic publications as an important tool for guiding policy as well as further research and development (Subramaniam et al., 2021). Nonetheless, accurate environmental accounting of

diverse and numerous bioproduction platforms represents a daunting challenge.

Techno-economic assessment

While LCAs are crucial for comparing the relative sustainability of biological and petrochemical products, TEAs actualize the commercial viability of biotechnology platforms. TEAs differ from other accounting strategies (e.g., cost-benefit analysis or cost-effectiveness analysis) by focusing on emerging technologies and, in general, there are few explicit methodological guidelines (Van Dael et al., 2015). TEAs tend to be highly specialized to less mature production pathways and are typically completed at lower TRL to assess early design decisions (Scown et al., 2021). Depending on TRL, TEA may be used for attracting stakeholders and encouraging investment by governmental entities, research and development departments, or early-stage investors to ensure efficient technology maturation (Sick et al., 2020).

At their most basic level, TEAs help to identify economic indicators in production pathways to reduce the minimum selling price (MSP) of a given product. The MSP is the price at which the net present value of a production system is zero and, when compared to market prices, the MSP succinctly describes the profitability of that system. While TEAs parallel the same overall design logic as LCAs in their accounting of material and energy flows, they maintain obvious emphasis on reducing overall process expenditures rather than environmental burden. Unlike the clear methodological demarcation of LCAs, TEA methodologies are far less uniform, often owing to divergent objectives depending on TRL (Thomassen et al., 2019). Scenarios forecasting, for example, is often contingent upon market uncertainty, which is especially relevant for bioproducts competing with volatile petrochemical prices prone to significant fluctuations in global supply and demand. Furthermore, price modeling must account for coproducts (Biddy et al., 2016).

Techno-economic assessments divide costs into two categories: capital expenditures (CAPEX) that include fixed assets and operational expenditures (OPEX) that encompass feedstock prices, utilities, operator salaries, and similar day-to-day expenses (Lynch, 2021). CAPEX may be further expanded by estimating asset interest rates and modified accelerated cost recovery system (MACRS) for system cost depreciation. Calculation of total expenditures within a discounted cash flow analysis based on an internal rate of return enables the elucidation of the MSP of a product at scale in an "nth" plant which is readily comparable to current price of the good in the market. Green technologies range from carbon capture to lithium-ion batteries to lignocellulosic biofuels, all of which are geared at displacing current technologies for emissions reduction. TEAs are therefore a pivotal tool in separating viable production by quantifying the current state of technology and highlighting design areas for the most efficient improvement. Recently, TEA methodologies have been tutorialized for practitioners to better assess obstacles and opportunities in low TRL bioproduction (Buchner et al., 2018; Thomassen et al., 2019).

Translation of life cycle assessments and techno-economic assessments to biological systems

Initial considerations of biochemical sustainability date to the development of the 12 principles of green chemistry (Anastas and Williamson, 1996), which were gradually parametrized in early LCA frameworks and ultimately applied to industrial ecology as a whole (Anastas and Lankey, 2000). Microbial metabolic engineering has the propensity to generate economically competitive products with vastly improved sustainability metrics compared to archetypical petrochemical production (Dunn, 2019). Assessments of biological systems traditionally focus on agricultural feedstocks and biofuels, but more recently have expanded to a broader range of bulk, value-added chemical precursors (Dunn et al., 2015).

The central challenge to integrating synthetic biology within the TEA/LCA framework is the mitigation of uncertainty. Longstanding challenges for synthetic biologists, for example, have been the unpredictability, incompatibility, volatility, and lack of knowledge in complex circuits that complicate the extrapolation of bench to pilot scale production (Kwok, 2010) as well as experimental reproducibility (Baker, 2016). Physical constraints also strongly influence bioproduction at scale. Large-scale bioreactors can exacerbate expression stochasticity due to gradients in dissolved gas and nutrient concentrations, which lend to suboptimal production and potentially the selection of deleterious mutations (Wehrs et al., 2019; Czajka et al., 2020). This phenomenon is especially well-studied in *E. coli*, where large-scale production often results in lower overall biomass and growth rates accompanied by significant organic acid accumulation (Neubauer et al., 1995).

From an analytical standpoint, the difficulty of intercomparing assessments cannot be understated. Even when utilizing the same production pathway and operating in accordance with ISO 14040:2006, the extrapolation of low confidence assumptions from low TRL, bench scale production strategies often yields assessments with starkly different conclusions and suboptimal design decisions (Hellweg and Milà i Canals, 2014).

Algal biofuels remain a useful case study for effective translation of LCA/TEA to biological systems. Algal biofuels continue to be popular candidates for LCA due to the abundance of extraction methods (e.g., hydrothermal liquefaction, lipid extraction, or combined algal processing), strains (e.g.,

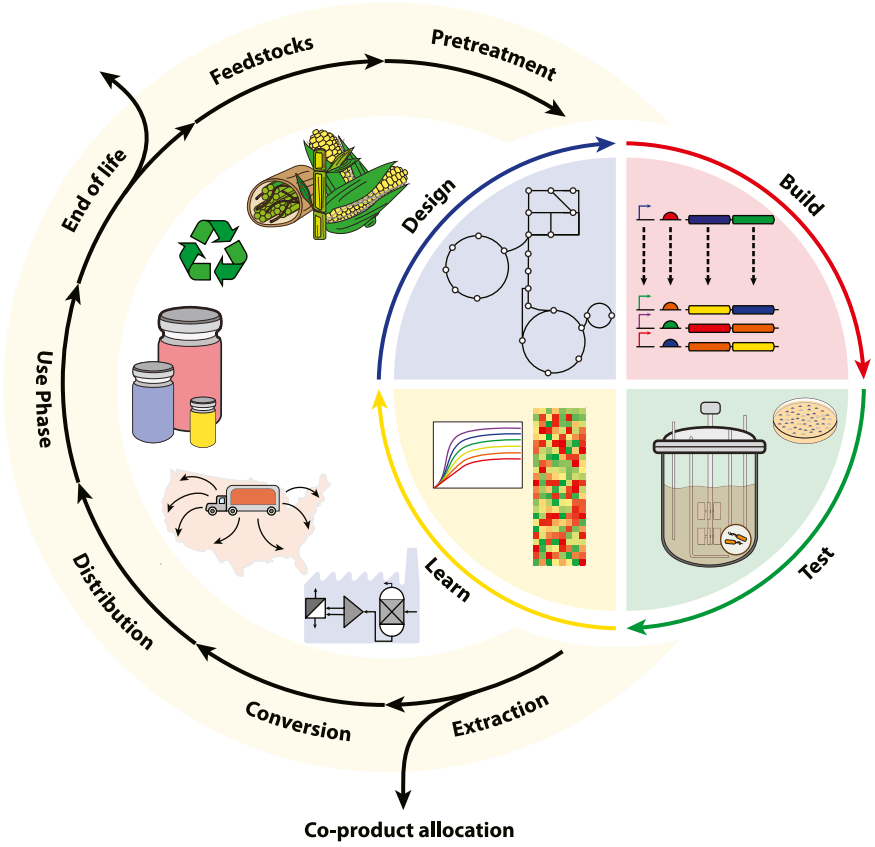


FIGURE 1
The integration of LCA and synthetic biology within a circular bioeconomy. Here, the LCA framework is subdivided into the cradle-to-grave stages of raw material acquisition (feedstocks and pretreatment), manufacturing (DBTL, extraction, and conversion), distribution, use-phase, and end-of-life. For bioproduction platforms, extraction may be further subdivided into bioseparation and media recycling with co-product allocation also accounting for biological waste treatment. Synthetic biology is represented by the DBTL cycle and integrated into the core manufacturing stage of the LCA. The framework enables iterative benchmarking for scaling bioproduction according toTRL.

halophilic *Dunelia* sp. or freshwater *Chlorella* sp.), and culture conditions (e.g., outdoor pond, outdoor raceway, photobioreactor) alongside the ease with which algal biodiesel can be integrated into contemporary fuel systems for valorization of photosynthetically sequestered CO_2. However, meta-analyses and harmonization reports have highlighted how the range of modeling assumptions, process data, and altogether divergent production schemes can heavily impact the translation of experimental data from flask to field (Quinn and Davis, 2015; Tu et al., 2017; Cruce et al., 2021). Quinn et al. note the variability in TEA/LCA metrics, citing reports of GHG emissions varying from -75–532 g CO_2eq per 1 MJ fuel as well as approximately $1.34 to $30 per gallon of fuel (Quinn and Davis, 2015).

More broadly, assessment variability is ubiquitous across bioproducts. A recent bioplastic critical review further reported high variation between production strategies with energy usage per kilogram of bio-PET varying by 400% and bio-PLA by over 1,000% (Walker and Rothman, 2020). Harmonization reports have attempted to synthesize these data to better assess the state of technology and note areas of greatest sensitivity. In the case of algal bioproduction, harmonization reports have emphasized pond productivity as and material inputs as major targets for optimization (Cruce et al., 2021). Perhaps most importantly, is that these studies stress the necessity for data reliability, transparency, availability, and model homogenization such that assumptions between studies can be controlled for and conclusions better validated. Specific studies elaborating on the complexities and deficiencies typical in current LCAs of bioproduction platforms have also been conducted (McKone et al., 2011; McManus et al., 2015; Ögmundarson et al., 2020).

Addressing the uncertainty of emergent, low TRL bioproduction platforms is a formidable challenge for TEAs and LCAs, yet early assessments are imperative for spurring investment in the bioeconomy (Hillson et al., 2019). Recent order-of-magnitude type assessments have been developed to

better characterize the economic and environmental fingerprint of low TRL biotechnologies (Shi and Guest, 2020; Lynch, 2021; Scown et al., 2021). These "agile TEA-LCAs" aid biologists in addressing process level bottlenecks and viability early in the DBTL pipeline. The integration of synthetic biology and the archetypical LCA are depicted in Figure 1.

Three novel assessment tools of particular importance include the Biorefinery Simulation and Techno-Economic Analysis Modules (BioSTEAM) (Cortes-Peña et al., 2020; Shi and Guest, 2020), the Bioprocess TEA calculator (Lynch, 2021), and Early State Technoeconomic Analysis ESTEA2 (Viswanathan et al., 2020). Their hallmarks include open-source access, good correlations with more robust tools like GREET and Ecoinvent, and tailored design for practitioners grounded in biological sciences. Low TRL bioproduction platforms are typically hampered by high OPEX due to strain performance, which may be characterized by titer, rate, and yield (TRY) per unit feed as well as other process-level functions like nutrient recycling and co-product allocation (Huang et al., 2021). The fact that improving TRY consequentially improves bioproduction environmental and economic metrics is fundamentally intuitive even at bench scale. However, agile assessments are useful in contextualizing how biological improvements translate to changes in process viability early in process design. Using these tools, synthetic biologists can harness the DBTL cycle to enhance TRY and benchmark improvements *in vivo*. This is exemplified by recent application of python-based BioSTEAM to lignocellulosic production of lactic acid and acrylic acid in which practitioners identified that separations efficiency and titer as well as titer and yield to be the greatest opportunities for improvement, respectively (Li et al., 2021). Another study by McClelland et al., consolidated experimental strain engineering and catalyst loading for optimized linear alpha-olefin production, concluding that attaining 40 g/L titer, 0.5 g/L/hr rate, and 80% theoretical yield would enable economic viability (McClelland et al., 2021). Importantly, these assessments provided comparable results to gold-standard tools like ecoinvent and GREET (Bhagwat et al., 2021; Li et al., 2021).

In the following sections we explore how the burden of feedstocks can be addressed through translation of production platforms to inexpensive and prevalent carbon sources, then consider how synthetic biology has improved the TRY of a specific set of bioproducts.

Advances in feedstock engineering

Costly feedstocks are a major barrier to efficient bioproduction due to agricultural energy accounting (Cherubini et al., 2009). Assessments of sugar substrates for microbial bioproduction are commonly limited to either Brazilian sugarcane or American corn/beet because of their prevalence as feedstocks for bioethanol. These two sources alone have extraordinarily different environmental impacts. Brazilian sugarcane, for example, tends to offset some nonrenewable energy usage (NREU) due to bagasse combustion and its high sugar content, yet quantifying carbon cost is challenging due to land use change (e.g., pasture vs. rainforest), potential biodiversity loss, and so forth (Renouf et al., 2008; Chaplin-Kramer et al., 2017). Corn, on the other hand, has comparatively higher GHGs, NREU, and eutrophication potential due to fertilizer demand but significantly lower water intensity. These disparate environmental fingerprints showcase how sugar feedstocks are complicated by regional differences in sourcing and agricultural practices and, more broadly, market demand and sociopolitical attitudes. Each factor impacts process uncertainty. In any case, pure sugars are an environmentally and economically expensive substrate. The production of raw sugar, namely glucose, routinely accounts for 40–60% of overall bioproduction OPEX as well as considerable NREU and GHG emissions (Tsiropoulos et al., 2013; Gunukula and Anex, 2017; Rios et al., 2021).

Unlocking recalcitrant substrates has tremendous potential for reducing input costs and GHG emissions, valorizing waste streams, and decoupling food from fuel. Yet gaining access to these substrates through hydrolysis or chemical extraction has proven challenging (Jansen and van Gulik, 2014; Lim et al., 2021). In recent years, significant breakthroughs have been made in engineering microbial uptake of atypical carbon substrates in high production chassis as well as integrating production pathways into microbes with atypical carbon pathways. Regardless of approach, the most promising strategies for feedstock cost reduction utilize either lignocellulosic biomass (i.e., second generation biofuel production) or C1 substrates (i.e., CO_2, CO, CH_4, methanol, or formate). Furthermore, the engineering of microbial consumption of C1 feedstocks enables conversion of common waste streams to value-added chemicals (Clomburg et al., 2017; Zang et al., 2021).

In this section, we first describe current advances in lignocellulosic bioproduction, then discuss how C1 metabolism could be harnessed to decrease the economic and environmental burden of carbon feedstocks in bioproduction. An overview of the respective feedstocks, metabolic pathways, as well as selected products and precursors is depicted in Figure 2.

Lignocellulosic biomass

Lignocellulosic biomass is principally composed of cellulose (30–50%), hemicellulose (15–30%), and lignin (15–30%) (Bugg et al., 2011). Cellulose consists of polymeric glucose with 1,4-glycosidic bonds whereas hemicellulose is a heterogeneous, branched polysaccharide comprised of a mix of hexose (e.g., glucose and galactose) and pentose (e.g., xylose and arabinose) sugars. Both lignocellulosic components may be biologically or

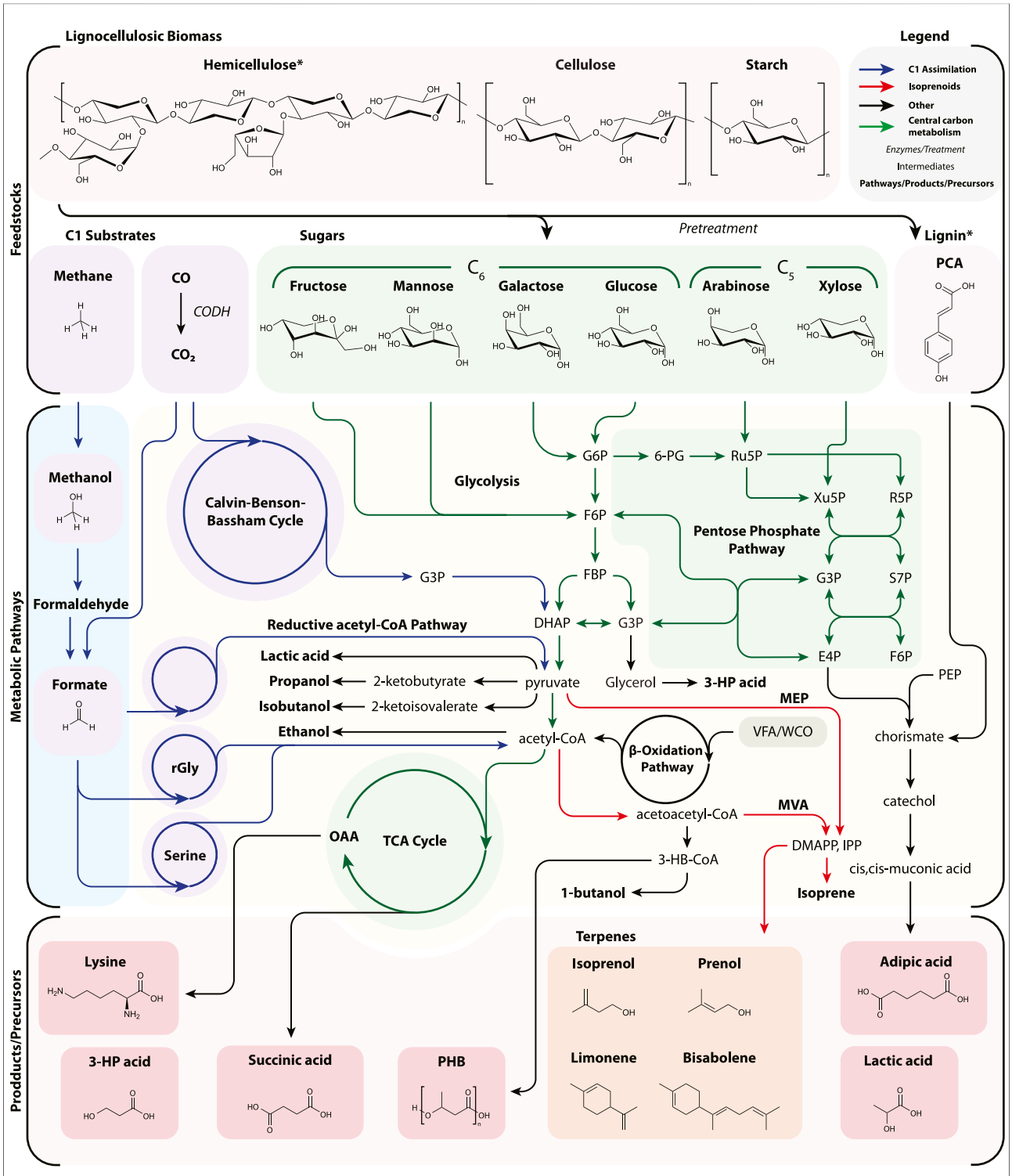


FIGURE 2
A simplified diagram of feedstock assimilation into bioproduction pathways with specific attention to C1 (CO_2, CO, CH_4, methanol, and formate) as well as lignocellulosic biomass. Pathways are not exhaustive nor necessarily the most efficient, but representative to the selection of bioproducts reviewed. Many other routes have been successfully demonstrated. For example, adipic acid production is depicted from *p*-coumaric acid via the shikimate pathway and from lignin derivatives, though may be generated from glucaric acid, TCA intermediates, and the β-oxidation pathway. Lignin is represented by a single aromatic, *p*-coumaric acid, and *cis*,*cis*-muconic acid is usually chemically hydrogenated to adipic acid. Likewise, certain pathways have been simplified for clarity (e.g., lysine and CO assimilation). 3-HP acid, 3-hydroxypropionic acid; 6PG, 6-phosphogluconate; CoA, coenzyme A; CODH, carbon monoxide dehydrogenase; DHAP, dihydroxyacetone phosphate; DMAPP, dimethylallyl diphosphate; E4P, erythrose-4-phosphate; F6P, fructose 6-phosphate; FBP, fructose 1,6-bisphosphatase; G3P, glyceraldehyde-3-phosphate; G6P, glucose-6-
(*Continued*)

FIGURE 2
phosphate; IPP, isopentenyl diphosphate; MEP, methylerythritol 4-phosphate; MVA, mevalonate; OAA, oxaloacetate; PCA, p-coumaric acid; PEP, phosphoenolpyruvate; PHB, poly-3-hydroxybutyrate; R5P, ribose 5-phosphate; rGly, reductive glycine pathway; Ru5P, ribulose 5-phosphate; S7P, sedoheptulose-7-phosphate; Serine, serine cycle; TCA, tricarboxylic acid cycle; VFA, volatile fatty acids; WCO, waste cooking oil Xu5P, xylulose 5-phosphate.

chemically degraded into utilizable monomeric substrates. Conversely, lignin is an extremely heterogeneous substrate characterized by heavily crosslinked phenolic compounds that contribute to rigidity of woody biomass and have proven especially recalcitrant for microbial production. Collectively, the cost of pretreatment and tolerance remain the principal limitations to the utilization of lignocellulosic biomass for microbial bioproduction.

Deconstruction strategies to catalytically "unlock" lignocellulosic sugar and aromatic monomers may be subdivided into chemical, mechanical, physicochemical, or biological pretreatments with ionic liquids, milling, ammonia fiber explosion (AFEX), or microbial degradation, respectively, as representative examples (Davis et al., 2013; Kumar and Sharma, 2017). Each technology has certain trade-offs, generally pertaining to high cost, efficient conversion strategies as in the use of ionic liquids (ILs), or low cost, slow conversion in the case of microbial degradation (Beckham et al., 2016). Broad LCAs that compare first and second-generation sugar production describe significant variation between processing costs due to choice of deconstruction technology (Bello et al., 2021). Contingent on sourcing, biomass processing results in higher overall GHG emissions and NREU, but comparatively less eutrophication and acidification potential (Bello et al., 2021). Deconstruction reviews have also highlighted the high cost of processing reagents, which are partly due to technology nascency but have spurred optimization studies (Baral et al., 2019). One such study demonstrated efficient deconstruction of sorghum with 50% less IL in a third of the time, then paired the resultant hydrolysate with a naturally tolerant, genetically modified strain of *Rhodosporidium toruloides* (Magurudeniya et al., 2021). The strategy improved bisabolene production on hydrolysate with an estimated 10% lower MSP while further achieving emissions and cost reduction through optimized IL recovery (Magurudeniya et al., 2021). Importantly, integrating deconstruction with bioproduction significantly improved the viability of lignocellulosic production.

While cellulose and hemicellulose are degraded into broadly utilizable hexose and pentose sugars, lignin is degraded predominantly into aromatic compounds like *p*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohols (Singhvi and Gokhale, 2019). Ideally a strain can simultaneously metabolize hexose and pentose sugars while tolerating concentrations of lignin-derived aromatics that are not only difficult to metabolize but often inhibit microbial fermentation altogether, as is the case for *Saccharomyces cerevisiae* and *Escherichia coli*. Likewise, *E. coli* demonstrates diauxic growth on mixed sugar substrates due to carbon catabolite repression (Deutscher, 2008). Some oleaginous yeasts like *Yarrowia lipolytica* (Sun et al., 2021) and *R. toruloides* (Kirby et al., 2021) have shown high tolerance to lignocellulosic components as well as genetic tractability, with the latter organism having demonstrated terpene production. On the other hand bacterial lignin depolymerization is limited to a subset of α-proteobacteria, γ-proteobacteria, and actinobacteria (Bugg et al., 2011), with *Pseudomonas* sp. As the most popular host. In particular, *Pseudomonas putida* is naturally resistant to ILs due in part to the presence of cholinium catabolizing pathways (Park et al., 2020). Capitalizing on the native predispositions of *P. putida* and *R. toruloides*, adaptive laboratory evolution (ALE) and tolerance ALE (TALE) further improved resistance to ILs (Sundstrom et al., 2018; Lim et al., 2020) as well as enabled growth on the lignin aromatics *p*-coumaric acid and ferulic acid (Mohamed et al., 2020; Liu Z. et al., 2021). In *P. putida*, integration of *xylAB* genes for xylose catabolism enabled simultaneous uptake of hexose, pentose, and aromatic compounds from hydrolysate (Elmore et al., 2020). Simultaneously, other works have expanded the repertoire of *P. putida* and *R. toruloides* bioproducts with value-added muconic acid and terpenoids (Sundstrom et al., 2018; Johnson et al., 2019; Kirby et al., 2021).

Although TALE has increased IL tolerance in *E. coli*, uptake of recalcitrant aromatics remains a major challenge (Mohamed et al., 2017). A recent study explored production of catechol in *E. coli* from vanillin, a common lignin-derived chemical. The aromatic transporter *CouP* was co-expressed with *LigV* and *LigM* for vanillin degradation and the protocatechuate decarboxylase *AroY* in *E. coli* all under a ADH7 vanillin inducible promoter to regulate inhibition and ultimately reduce toxicity (Wu et al., 2018). While successfully increasing catechol titer by 40%, final titers remained comparatively lower than the titer from *P. putida*. Nonetheless, these studies collectively demonstrate how adaptive evolution paired with product-oriented metabolic engineering can valorize lignocellulosic biomass into useful bioproducts.

To summarize, most approaches to valorize lignocellulosic biomass either integrate production pathways into organisms such as *P. putida* and *R. toruloides* with natural tolerance to aromatic compounds or engineer tolerance toward aromatic compounds in common production chassis like *E. coli*. While both have made significant strides over the last decade, most lignocellulosic bioproduction platforms are not yet commercially viable.

C1 substrates

Many C1 substrates are generated through anthropogenic waste streams in the form of flue gas (CO_2, CH_4, and CO), anaerobic digestion (CH_4, CO_2), and as byproducts of the petroleum industry (CH_4). CO_2 may then be electrochemically converted into methanol (CH_3OH) and formate (HCOOH). Collectively, this library of chemicals has the potential to revolutionize value-added production by methylotrophic, formatotrophic, or phototrophic organisms. Methane and methanol are especially appealing due to their high energy-to-carbon ratios compared to glucose such that pairing methylotrophy with glucose consumption could balance production, a framework that has already demonstrated improvements in titers and yields while reducing costs (Liu et al., 2020).

Although at the Frontier of microbial bioproduction, substrate utilization is challenged by the unique physical constraints of C1 chemicals, which demand specialized fermentation reactors, media, and culture conditions. While CO_2 may be bubbled directly into a reactor, for example, photosynthetic organisms require sufficient illumination that increase production costs dramatically. Likewise, carboxidotrophic bioproduction is hampered by the low solubility of CO in aqueous media requiring specialized reactors with high gas-liquid volumetric transfer coefficients to maximize solubility (Köpke et al., 2011a). Furthermore, the use of CH_4 or CO in any fermentation strategy raises significant safety concerns at scale and, conversely, formic acid and methanol maintain a high energetic cost of synthesis and microbial toxicity, respectively. Nonetheless, these pathways enable the valorization of industrial waste otherwise emitted into the atmosphere and represent a tremendous opportunity from a TEA/LCA perspective.

As with lignocellulosic biomass, the two predominant strategies of C1 bioproduction are through engineering metabolically tractable chassis like *E. coli* and *S. cerevisiae* for C1 assimilation or through heterologous expression of production pathways in native assimilators. Here, we subdivide microbial C1 metabolism primarily into photoautotrophic, methylotrophic, and formatotrophic carbon assimilation with an emphasis on recent advances and improvements.

Carbon dioxide

Photoautotrophic microbes photosynthetically fix carbon dioxide to generate carbohydrates via the Calvin-Bassham-Benson (CBB) cycle. Photosynthetic bioproduction is especially appealing due to the conversion of CO_2 and sunlight into valuable products. Select species of algae (e.g., *Chlorella* sp., *Duniella* sp., *Nannochloropsis* sp., etc) are attractive due to their ability to quickly accrue biomass that, depending on the strain and culture conditions, is often rich in triglycerides and carbohydrates ideal for low value, high volume catalytic conversion. Their natural accumulation of high energy metabolites is further complemented by their uncanny ability to thrive on wastewater, saline, or freshwater environments. From a synthetic biology perspective, the model alga *Chlamydomonas reinhardtii* has a well characterized metabolic toolkit for genetic modification (Tran and Kaldenhoff, 2020), but the translation of these tools and strategies to other algal biofuel candidates has proven extraordinarily challenging (Mosey et al., 2021). Counterintuitively, the greatest improvement in algal biomass accumulation stemmed from a reduction in the light harvesting antenna size. Large light harvesting complexes, for example, enable growth in low light conditions but may lend to high photon absorption and photochemical quenching to avoid photodamage under high light conditions (Kirst et al., 2014). By limiting the potential burden of photon overabsorption in high light conditions, synthetic biologists circumvented the necessary dissipation of excess photons and improved culture biomass accumulation (Melis, 1999; Polle et al., 2003).

As prokaryotes, cyanobacteria (especially *Synechococcus elongatus*, *Synechocystis* sp.) are comparatively easy to genetically modify and maintain a high theoretical photosynthetic energy conversion efficiency of ~10% (Lewis and Nocera, 2006). As with algae, engineering photosynthetic efficiency is complicated by evolutionary adaptations to natural light cycling, where organisms must propagate under both indirect and direct sunlight. As a result, cyanobacterial photosynthetic efficiency is typically 1–2%. Following promising improvements in algae, *Synechocystis* was also modified with truncated light harvesting antennae, yielding a 57% increase in productivity in limited to high light conditions (Kirst et al., 2014). More recently, sink engineering has also arisen as a novel strategy for rerouting carbon or electron flux towards specific chemicals. Analogous to antenna size reduction, sink engineering reroutes utilizable photosynthetic energy towards bioproducts. The concept has been demonstrated with sucrose (Ducat et al., 2012), 2,3-butanediol (Oliver et al., 2013), and isoprene (Lindberg et al., 2010) through a hypothesized reduction of accumulating inhibitory intermediates. One work also combined sucrose export and P450 expression, noting an additive benefit that further increased photosynthetic efficiency by improving utilization of the electron transport chain (ETC) (Santos-Merino et al., 2021). As a whole, these works establish that heterologous bioproduction or fine tuning of metabolic pathways in canonical cyanobacteria enable access to excess energy from photosynthesis otherwise dissipated to avoid overreduction of the ETC, accumulation of reactive oxygen species, photodamage, and photoinhibition (Santos-Merino et al., 2021). Emerging insight of cyanobacterial photosynthesis is coupled with a slew of metabolic engineering tools like promoter libraries, CRISPR-cas systems, and genome scale models (Santos-Merino et al., 2019) and, collectively, these works provide encouraging strategies of how

conditional rewiring of photosynthetic machinery can spur advancements in cyanobacterial bioproduction.

Many thorough reviews, LCAs, TEAs, and subsequent harmonization reports have investigated varied photoautotrophic biofuel production methodologies in an attempt to better understand limitations barring commerciality (Quinn and Davis, 2015; Tu et al., 2017; Cruce et al., 2021). In general, the scaling of algal platforms has proven economically prohibitive due to high CAPEX of the cultivation methodology (e.g., outdoor raceways, photobioreactors, or cattle tanks) combined with culture instability and relative extraction efficiencies. Overall, the most prohibitive obstacles are slow growth, tolerance to environmental fluctuations, and resistance to invasion by grazers, disease, or other phototrophs. While clever strategies have been developed to address culture instability, including polycultivation of multiple strains to improve culture stability (Narwani et al., 2016) or unnatural phosphite feeding (González-Morales et al., 2020), few have proven immediately viable at scale (Carruthers et al., 2019). Thus while rough approximations of, for example, *S. elongatus* sucrose yield outstrip sugarcane on a per hectare basis (Ducat et al., 2012), the realization of these yields in the context of LCA and TEA demands a more holistic engineering approach that integrates laboratory strain improvements and practical cultivation strategies within the biorefinery framework.

Methane and methanol

Methylotrophs are enticing candidates for capturing and valorizing industrial methane waste streams. Common candidates include the bacteria *Methylorubrum extorquens* AM1 (formally *Methylobacterium extorquens* AM1) and *Methylobacillus glycogenes* as well as the prominent yeast *Pichia pastoris* (now *Komagataella phaffi*), which incorporate methane or methanol through a variety of highly regulated processes. All methylotrophic organisms first oxidize methanol to formaldehyde through a cofactor dependent oxidoreductase. In bacteria, formaldehyde may be assimilated into carbon metabolism through either the ribulose monophosphate (RuMP) cycle (Type I) or the serine cycle (Type II). Alternatively, yeasts compartmentalize methanol-derived metabolites and ultimately assimilate formaldehyde in the xylulose monophosphate (XuMP) cycle. *K. phaffi*, for example, natively expresses methanol inducible alcohol oxidase genes *AOX1* and *AOX2* to enable methanol assimilation through peroxisome biogenesis, thereby sequestering toxic formaldehyde and hydrogen peroxide (Peña et al., 2018). Leveraging the tightly regulated inducible *AOX1* promoter has enabled biphasic culturing strategies and heterologous production, usually entailing an initial glycerol growth phase followed by a methanol production phase. Recent developments have led to rapid expansion of the methylotrophic production portfolio, including terpenoids ranging from C5 isoprene to C30 squalene, though with lower comparative titers to traditional cultivation (Carruthers and Lee, 2021).

Recently, interest in synthetic methylotrophs has led to metabolic engineering of *E. coli*, *S. cerevisiae,* and *Cornyebacterium glutamicum*. In bacteria, *de novo* C1 assimilation has been achieved through heterologous expression of a methanol dehydrogenase, which enabled growth upon C13 labeled methanol with limited sugar supplementation and subsequent metabolomic profiling of central carbon metabolism (Tuyishime et al., 2018; Kim et al., 2020). From there, various studies have investigated formaldehyde assimilation via native aldolases into the homoserine cycle, or further oxidation to formate, discussed later (He et al., 2020). On the other hand, a recent study identified that *S. cerevisiae* has native tolerance to methanol as well as low levels of assimilation. Methanol assimilation was enhanced by 44% using ALE (Espinosa et al., 2020). Heterologous expression of C1 assimilation pathways has tremendous potential for reducing the economic and environmental burden that have hitherto limited commercial viability, however many obstacles remain to actualize C1 bioproduction including, as an example, safely scaling methane fermentation systems.

Formate

Many organisms are capable of oxidizing formate to CO_2 via a formate dehydrogenase either as a tolerance mechanism or to supply reducing power. Native formate assimilation is usually accomplished by either the reductive pentose phosphate cycle (CBB cycle), the serine cycle, or the reductive acetyl-CoA pathway with varied efficiencies and bottlenecks. Yet few organisms are naturally capable of surviving solely on formate and formate assimilation is phenotypically isolated to a subset of fastidious microbes (Yishai et al., 2016).

Fortunately, formaldehyde assimilation pathways are necessary for methylotrophy. Elucidation of formate and formaldehyde assimilation has coincided with a revolution in our understanding of metabolic pathways. Combinedly, these advances have led to hypothetical synthetic pathways designed to tune thermodynamic efficiencies and redox balances (Bar-Even, 2016; Yishai et al., 2016). Of those synthetic pathways, the reductive glycine pathway (rGly) has proven a popular *in silico* target due to its relatively few enzymatic steps and favorable ATP/NAD(P)H consumption. Remarkably, a novel natural pathway reducing CO_2 to formate via the rGly was recently elucidated in an isolate of dissimilatory phosphite-oxidizing microorganisms (Figueroa et al., 2018), specifically *Desulfovibrio desulfuricans* (strain G11) (Sánchez-Andrea et al., 2020), demonstrating the propensity of the rGly to naturally drive central carbon metabolism. Indeed, *de novo* formate assimilation pathways in *E. coli* and *S. cerevisiae* have demonstrated *in vivo* serine and glycine production (Yishai et al., 2018; Gonzalez de la Cruz et al., 2019). One seminal study combined overexpression of the rGly cycle with ALE to yield

a strain of *E. coli* capable of growth solely on formate and establishing a novel strategy for reduced feedstock cost in a common metabolic chassis (Kim et al., 2020). The confluence of biological and computational approaches in the elucidation and application of the rGly also has tremendous implications for further pathway optimization. It also presents an opportunity for rewiring or heterologous expression of individual genes from the remaining four CO_2 fixation pathways (the reductive TCA cycle, the 3-hydroxypropionate cycle, dicarboxylate/4-hydroxybutyrate cycle, and the 3-hydroxypropionate/4-hydroxybutyrate cycle).

Complementary to biological advances, LCAs have shown that electrochemical derivation of formic acid from CO_2 is far more dependent on grid specific electricity composition compared to other C1 chemicals. As electricity and H_2 are increasingly generated from a renewable grid, formic acid becomes increasingly viable as a substrate (Sternberg et al., 2017). Integration and tuning of *de novo* formate assimilation pathways represents the Frontier of feedstock engineering and, in the context of advances in carbon sequestration, could enable upcycling of CO_2 into valuable chemicals.

Carbon monoxide

Carbon monoxide is a major constituent of syngas (CO_2, CH_4, H_2, and CO) commonly generated through gasification reactions of biological or fuel carbon or through steam reforming in ammonia synthesis. Historically, the value of syngas lies in its canonical catalytic conversion into valuable hydrocarbons through the Fischer-Tropsch process or into hydrogen via the water-gas shift (Teixeira et al., 2018). Yet, from a microbial perspective, syngas represents a unique feedstock for niche anaerobic. CO metabolism is accomplished anaerobically or aerobically through the expression of a specialized heterometaloenzyme, carbon monoxide dehydrogenase (CODH). Microbial CODHs are classified as either monofunctional or bifunctional. Monofunctional CODHs reversibly oxidize CO to CO_2 whereas bifunctional CODHs, coupled with an acetyl-CoA synthetase, also catalyze the condensation of CO, a CO_2 derived methyl group, and a CoA-SH to generate acetyl-CoA (Oelgeschläger and Rother, 2008; Schuchmann and Müller, 2014).

Carboxidotrophs represent a specific class of bacteria capable of sole growth on CO through its oxidation to CO_2 and subsequent incorporation into central carbon metabolism via the reductive acetyl-CoA pathway. Reducing equivalents from CO oxidation then drive respiratory functions with further capacity to generate hydrogen, acetate, methane, and so forth (Oelgeschläger and Rother, 2008; Schuchmann and Müller, 2014). As a result, this unique metabolic class is promising in its capacity to further valorize syngas. Microbial growth on CO_2, CO, and H_2 principally using *Clostridium* sp. Has demonstrated production of isoprene, 2,3-butanediol, ethanol, succinic acid among other products (Köpke et al., 2011b; Liew et al., 2016). Commercial interest by LanzaTech (Köpke et al., 2011a; Heijstra et al., 2017) has spurred strain development and pilot coupling of microbial systems with industrial waste. Although much of the information is proprietary, a recent LCA of the LanzaTech process asserted that their ethanol production platform could generate ethanol from syngas at 60% reduced GHG emissions compared to conventional processes, a figure that matches the EPA standards set for cellulosic biomass (Handler et al., 2016). Importantly, the diversity of microbial C1 metabolism presents a fascinating opportunity for valorizing syngas beyond conventional conversion approaches as we describe in the following section.

Consortial approaches

Engineered microbial communities divide a metabolic pathway between constituent organisms to increase overall pathway efficiency. Consortial approaches may integrate microbial and chemical syntheses into a single, unified platform that parallels agricultural or industrial waste stream effluxes (Roell et al., 2019). Emblematic of natural ecological processes like lignolytic detritus degradation, synthetic biologists can co-culture microbes to access recalcitrant substrates. Pairing *Trichoderma reesei*, a filamentous yeast renowned for secreting cellulolytic enzymes, with *E. coli,* for example, enabled access to sugar monomers and ultimately to 1.88 g/L isobutanol production (Minty et al., 2013).

More recently, researchers have engineered consortia with C2 metabolite cross feeding (Kim Y. et al., 2021; Ma et al., 2021). Ethanol and acetate are typically generated via overflow metabolism, which is characterized by incomplete oxidation of sugars in high growth rate production organisms, or by acetogenic microbes under anaerobic conditions. However, assimilation of C2 substrates is challenging due to imbalances in ATP, NADPH, and acetyl-CoA generation. Acetate, for example, may be directly converted into acetyl-CoA, but ATP and NADPH generation typically require the addition of sugars (Park et al., 2019). A specific study expressed acetaldehyde and alcohol dehydrogenase enzymes *ada* and *adha*1 in *E. coli* for direct conversion of ethanol to acetyl-CoA, demonstrating low titer production of PHB and prenol in rich medium (Liang et al., 2021). Co-feeding of acetate and gluconate in *Y. lipolytica* has also been reported for lipid production in which the low NADPH barrier is overcome by limited sugar addition, presenting its possible candidacy in a microbial consortia (Park et al., 2019).

Pairing specialized microbes has emerged as a strategy to reduce nutrient input costs as in the case of photoautotrophic and diazotrophic bacteria, which natively fix carbon dioxide and nitrogen gas, respectively. Fixation and excretion of sucrose by *S. elongatus* demonstrated co-culture production of low titer PHBs from an engineered *P. putida* (Löwe et al., 2017) and native *Halomonas boliviensis* (Weiss et al., 2017). Similarly, biologically fixed and subsequently secreted nitrogen from *Azotobacter*

vinelandii promoted syntrophic growth with green algae, thereby reducing Haber-Bosch derived nitrogen costs (Villa et al., 2014). Following these examples, microbial consortia can extend to more complex systems like engineered mycorrhizal communities that improve agricultural productivity while reducing input costs (Garrido-Oter et al., 2018; Wurtzel et al., 2019).

Valorization of waste streams is an obvious yet challenging opportunity for reducing production cost and environmental burden. An important caveat is that the success of these microbial platforms appears contingent on coupling bioproduction with traditional energy waste streams like flue gas. As energy systems continue to decarbonize, these platforms may adapt through cross-platform consortia that pair electrochemical processes like carbon capture and storage (CSS) and CO_2 hydrogenation with microbial fermentation (Bui et al., 2018). Of special interest are combinatorial approaches using CSS technologies to convert sequestered carbon into bioproducts (Faber et al., 2021) in "bioenergy CSS" platforms (Hanssen et al., 2020). Many proof of concept studies have already been established, including electrochemical conversion of CO_2 to formic acid for bioproduction in *C. necator* (Li et al., 2012).

Assessments of multifaceted systems are inherently challenging due to necessary approximations. For example, fluctuations in temporal microbial composition of mycorrhizal consortia must be reduced to generalized terms (e.g., productivity, yield, elemental composition). The challenge is therefore in summarizing complex processes without compromising on data resolution to elucidate important bottlenecks with some certainty, a challenge that may require more specialized tools within the LCA-TEA community.

Biofuels, biopolymers, and precursors

The substitution of petrochemicals with biobased sources is a major opportunity for reducing global GHG emissions (Zheng and Suh, 2019). In 2004, the United States Department of Energy published a report of fifteen chemical targets for biorefineries (Werpy and Petersen, 2004), which was further updated in 2010 based on the current state of technology. The chemicals are categorized by nine technological criteria that range from possible co-products in a scalable biorefinery to conversion and TRY (Bozell and Petersen, 2010). More recently, efforts have shifted to the concept of the BioFoundry that can rapidly generate an array of products from beachhead molecules, which include metabolic precursors like pyruvate, acetyl-CoA, malonyl-CoA, and the like (Hillson et al., 2019; Benavides et al., 2021). Importantly, these projects embody the use of LCA/TEA to streamline and productionize synthetic biology, while explicitly considering non-model organisms and atypical carbon substrates (Benavides et al., 2021).

It is notoriously difficult to compare bioproduction LCA/TEAs due to critical differences in assumptions, parameters, and process-level design decisions. More exhaustive meta-analysis and critical reviews of assessment strategies tend to highlight limited LCI data or sparsity of available assessments in general. Parsing variability between studies to determine process viability or sustainability is extremely challenging and often requires product specific reviews with a case-by-case analysis of LCA/TEA (Ögmundarson et al., 2020).

While more general reviews of bio-feedstocks are available (Cywar et al., 2021; Keasling et al., 2021), here we highlight a limited subset of biochemical precursors leveraged for bulk microbial production (lactic acid, succinic acid, adipic acid, 3-hydroxypropionic acid, and L-lysine), biopolymers (polyhydroxyalkanoates), and isoprenoid biofuels (bisabolene, limonene, and isoprenol) with specific attention to recent improvements in TRY, growth on non-glucose substrates, metabolically proximal co-products, and studies that include LCA/TEA. Rather than focusing on actual assessments, we consider how the metabolic engineering of strains for improved characteristics could translate into increased economic and environmental performance at scale.

Biochemical precursors

Lactic acid

Many industrial schemes have explored lactic acid production due in part to the natural abundance of lactic acid accumulating bacteria (e.g., *Lactobacillus* sp., *Lactococcus* sp., *Lacticaseibacillus* sp.) that generally outcompete pathway expression in common bioproduction chassis. Lactic acid is typically produced by a reduction of pyruvate under anaerobic conditions and can be readily condensed to polylactic acid (PLA) either through direct condensation or a ring-opening reaction involving the lactide intermediate (Vink et al., 2003). PLA is an attractive biopolymer not only due to its comparable thermal and mechanical properties to polystyrene and polyethylene terephthalate, but its high biodegradability (Zheng and Suh, 2019).

An LCA predating ISO 14044 was published to measure the relative GHG and energetics of the process outlined by Cargill Dow's NatureWorks™ PLA in 2003 (Vink et al., 2003) with estimated GHG emissions of 1.6 kg CO_2eq/kg PLA and with a required 54 MJ/kg, astonishing figures that outstripped their displaced plastic counterparts. Notably, these metrics stem from the cradle-to-gate system boundary, few process details, and lack of sensitivity data. Despite these clear deficiencies, the primordial LCA established feedstocks as a major opportunity for improving sustainability metrics and advocated that transitioning from pure sugars to corn stover could provide a 10-fold reduction in process energy demand. Many publications have since varied production strategies on glycerol and lignocellulosic biomass with a veritable portfolio of production

organisms (Morales et al., 2015; Adom and Dunn, 2017; Li et al., 2021). In particular, a recent work demonstrated production of 0.6 g/L of lactate from methane by *Methylomicrobium buryatense* with a TEA yielding optimistic values of 5.83–2.17 $/kg MSP, approaching those of lignocellulosic-based production (Garg et al., 2018; Fei et al., 2020).

Succinic acid

As with lactic acid, succinic acid (SA) is shortlisted as a top bulk biochemical precursor for use in generation of polybutylene succinate, polyester polyols, polyurethanes, 1,4-butanediol, and adipic acid among other chemicals (Jansen and van Gulik, 2014). Bio-succinic acid production has steadily increased over the last decade and culminated in a number of burgeoning commercial platforms via heterologous production in *E. coli*, *Actinobacillus succinogenes,* and *C. glutamicum* (Nghiem et al., 2017; Dickson et al., 2021). Optimization strategies typically involve channeling metabolic flux towards succinate by elimination of alternative anaerobic byproduct pathways, often down-regulating or completely removing native *ldh*, *acka*, *pta*, and *pfl*, which encode for a lactate dehydrogenase, acetate kinase, phosphate acetyltransferase, and formate acetyltransferase, respectively. Candidate strains have historically achieved titers close to or greater than 100 g/L and recent SA LCAs have tended to deem competitive or comparable to petrochemical pathways (Ögmundarson et al., 2020).

One study introduced the *M. extorquens* gene *fhd*2 for formic acid assimilation into a strain of *Mannheimia succiniciproducens* with significant modifications to mixed acid fermentation pathways (Ahn et al., 2017). Although supplementing formic acid and mixed sugars at a 1:5 ratio, the authors ultimately demonstrated 76.1 g/L SA production (4.08 g/L/h and 1.28 M yield) with C13 analysis (Ahn et al., 2017), approaching conventional production on glucose (Ögmundarson et al., 2020). Furthermore, TEAs of lignocellulosic biomass derived biofuel production have also highlighted SA as an exemplary value-added co-product for improved process valorization, which may prove pivotal for achieving favorable economics in other production pathways (Biddy et al., 2016).

Adipic acid

Approximately three million tons of adipic acid are generated annually, mainly to produce nylon (Kruyer and Peralta-Yahya, 2017). Adipic acid is generated through chemical synthesis using nitric acid and cyclohexane, generating nitric oxide as an extremely potent GHG byproduct in unabated systems (Kruyer and Peralta-Yahya, 2017). A plethora of biosynthetic pathways have been designed for adipic acid production and are generally divisible into production of muconic acid and glutaric acid precursors or direct production of adipic acid itself (Kruyer and Peralta-Yahya, 2017; Skoog et al., 2018). Engineering approaches for muconic acid have successfully utilized an extended shikimate pathway in which catechol is generated from chorismate either via a variety of intermediates including salicylic acid 4-hydroxybenzoic acid, 2,3-dihydroxybenzoic acid, or protocatechuic acid (Kruyer and Peralta-Yahya, 2017). Catechol is then converted to muconic acid via a heterologous catechol 1,2-dioxygenase with titers in *E. coli* reaching 36.8 g/L *cis,cis*-muconic acid (Niu et al., 2002). Other works have investigated direct adipic acid production via the reverse β-oxidation pathway, ω-oxidation pathway, and reverse adipate degradation pathway. By heterologously expressing a combination of five key genes from *Thermobifida fusca* and a CRISPR-mediated deletion of *ldhA*, *sucB*, and *atoB*, the latter approach ultimately generated 68.1 g/L adipic acid (0.381 g-adipic acid/g-glucose) in super rich medium (Zhao et al., 2018).

A recent TEA noted that, assuming theoretical yields of adipic acid and excellent catalyst properties, a fully biologically derived adipic acid route could achieve 1.36 $/kg, well below the MSP of 1.60 $/kg (Gunukula and Anex, 2017), but obtaining such biological yields is extraordinarily challenging. Alternatively, an estimated 41.79% of total adipic acid expenditures stem from growth on sugar feedstocks (Johnson et al., 2019). Lignin-based production, albeit at very high efficiencies, could achieve prices as low as 0.88 $/kg adipic acid (Rios et al., 2021). Furthermore, generation of adipic acid from lignin could reduce an estimated 62%–78% emissions compared to chemical synthesis (Corona et al., 2018). *P. putida* has arisen as a promising candidate in addressing feedstock burden due to its genetic tractability and native resistance to lignin aromatic toxicity. Following this predisposition, an engineered strain of *P. putida* demonstrated 13.5 g/L titer from M9 minimal medium fed with a 2:1 glucose to *p*-coumaric acid ratio (Vardon et al., 2015). More recently, an investigation demonstrated low titer muconic acid production in a modified *P. putida* grown solely on variably sourced lignocellulosic hydrolysate without exogenous sugar supplementation (Sonoki et al., 2018). Crucially, these advances move biologically derived adipic acid towards higher TRL and potential viability at scale.

3-Hydroxypropionic acid

Generation of 3-hydroxypropionic acid (3-HP) is achieved firstly through the dehydration of glycerol via a B12-dependent dehydratase and reduction by an alcohol dehydrogenase (Nakamura and Whited, 2003). The major derivatives of 3-HP include other bulk biochemicals like acrylic acid, 1,3-propanediol (1,3-PDO) and 3-hydroproprionaldehyde. For many anaerobic organisms, production of 3-HP from glycerol serves as an electron sink to enable NAD + regeneration (Nakamura and

Whited, 2003). 1,3-propanediol (1,3-PDO) is of special interest as a product of 3-HP reactions due to its potential as a biopolymer precursor, the proprietary generation of which has been completed biologically in *E. coli* by DuPont for several decades on glucose (Nakamura and Whited, 2003; Bozell and Petersen, 2010). More recently, *C. glutamicum* derived 1,3-PDO from glucose has demonstrated a final titer of 110.4 g/L, a yield of 0.42 (g-1,3-PDO/g-glucose), and a productivity of 2.30 g/L/h in fed-batch fermentation (Li et al., 2022). By further integrating xylose metabolism for simultaneous uptake of glucose and xylose, 1,3-PDO production from mixed sugar approached those metrics from pure glucose (98.2 g/L vs. 110.4 g/L titer, 0.38 (g-3-HP/g-mixed substrate) vs. 0.42 (g-3-HP/g-glucose yield)) (Li et al., 2022).

In an utterly different strategy, 3-HP production has been demonstrated in the Type II methanotroph *Methylosinus trichosporium* OB3b. Type II methanotrophs maintain high acetyl-CoA flux, which is a particularly useful trait in deriving valuable chemicals from C1 feedstocks (Nguyen et al., 2021). By engineering the malonyl-CoA pathway, methane fed cultures in nitrate mineral salt medium achieved 60.59 mg/L 3-HP (Nguyen et al., 2020). An optimized titer of 69.8 mg/L 3-HP has also been demonstrated in the methylotroph *M. extorquens* on a supplemented minimal medium (Yang et al., 2017). While low, such works demonstrate feedstock ingenuity and the potential for less common organisms to serve as metabolic chassis.

Finally, an assessment of 3-HP derived acrylic acid in BioSTEAM for lignocellulosic substrates noted a baseline MSP of 1.83 $/kg, assuming a titer of 54.8 g/L, productivity of 0.76 g/L/hr, and theoretical yield of 49% in *C. glutamicum* on lignocellulosic glucose and xylose (Bhagwat et al., 2021). Although marginally higher than current market prices, Bhagwat et al. establish critical advances necessary to achieve market competitiveness. As a result, the recent improvements in mixed substrate *C. glutamicum* yields are especially encouraging.

L-lysine

L-Lysine is an essential amino acid and critical precursor to several industrially relevant chemicals including glutaric acid, diamines, and 5-aminolevulinic acid (5-AVA), which are of special interest in their polymerization to polyamides, namely Nylon-6 and Nylon-510, which are almost ubiquitously derived from petrochemicals. *E. coli* has commonly been employed as a chassis for lysine derivatives with significant lysine supplementation. The heterologous expression of *P. putida* genes *davB* and *davA* for a lysine monooxygenase and delta-aminovaleramidase, respectively, in lysine supplemented medium led to 3.6 g/L 5-AVA production (Park et al., 2013). Further addition of the *P. putida* 5-AVA aminotransferase, glutarate semialdehyde dehydrogenase, and supplementation with alpha-ketoglutarate led to 1.7 g/L glutarate production (Park et al., 2013). The diamine cadaverine has been generated in *E. coli* through overexpression of the lysine production pathway, diamine degradation knockouts, and heterologous expression of the lysine decarboxylase *cadA* for conversion of lysine to cadaverine (Qian et al., 2011). Ultimately, Qian et al. demonstrated 9.62 g/L cadaverine production at a rate of 0.32 g/L/h and 0.12 (g-lysine/g-glucose) yield on minimal medium without lysine supplementation (Qian et al., 2011).

More recent works have explored polyamide precursor production in *C. glutamicum*, which naturally accumulates L-lysine due to a lack of native degradation enzymes (Kalinowski et al., 2003). Iterative optimization of *C. glutamicum* has included tuning the pentose phosphate pathway for improved NADPH cofactor production and a systematic overexpression of lysine biosynthesis genes, namely *lysA*, *dapB*, *lysC*, and *ddh* coupled with reduced expression of threonine dehydrogenase and specific TCA modification (Becker et al., 2011). Collectively, these modifications resulted in fed-batch production of 0.55 g-lysine/g-glucose and a final titer of 120 g/L lysine at a production rate of 4 g/L. While *C. glutamicum* appears an obvious chassis for lysine derivatives, initial generation has been hampered by co-production of non-target derivatives, like N-acetylcadaverine in the case of cadaverine production (Kind et al., 2010). Nonetheless, *C. glutamicum* maintains significant resistance to 5-AVA and glutarate toxicity and, unlike the diamines, demonstrated tunable product selectivity (Rohles et al., 2018).

Lysine derivatization poses a unique dilemma between organisms, specifically high precursor production, production on inexpensive substrates, and optimization of derivative pathways. More recently, *C. glutamicum* showed high titer conversion of lysine to glutaric acid (105.3 g/L), which is recognized as an important chemical precursor to polyamides and polyurethanes (Bozell and Petersen, 2010; Han et al., 2020). Nylon precursor production has also been demonstrated on nonsugar substrates like methanol and CO_2 though at markedly lower concentrations of 6.5 g/L cadaverine in *Bacillus methanolicus* (Naerdal et al., 2015) and 1.74 mM lysine in *Synechococcus* sp. (Korosh et al., 2017), respectively.

Isoprenoid biofuels

Biofuels comprise a broad category of drop-in chemical compounds that may serve as fuel additives to improve fuel characteristics (e.g., octane and cetane numbers, oxygen sensitivity, engine performance) or supplement conventional diesel/gasoline entirely. Advanced biofuels are produced from inedible carbon substrates and are especially attractive due to their propensity to displace conventional fossil fuels while valorizing the waste streams described previously. Advanced biofuels derived from metabolic routes for isoprenoids, fatty

acids, branched amino acids, and ketones, have arisen as important candidates in the energy market (Keasling et al., 2021).

Isoprenoids are naturally derived from C5 precursors (isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP)), which are in turn generated through either the mevalonate (MVA) or the methylerythritol 4-phosphate (MEP) pathways. While these pathways maintain different efficiencies, cofactors, and initial precursors, both have been candidates of extensive optimization to produce pharmaceuticals, fragrances, solvents, and biofuels. While hundreds of thousands of terpenes exist, biofuel candidates are generally limited to C5-C15 chemicals, especially those leveraged for bulk microbial production. Of special interest are the expansion of isoprenoid production platforms to non-model microbial chassis including *R. toruloides*, *Y. lipolytica*, as well as production on C1 carbon substrates (Carruthers and Lee, 2021). Major isoprenoid-derived biofuel candidates may be classified by chain length and include the hemiterpenes (C5) isoprenol, prenol, and isoprene, the monoterpenes (C10) limonene and 1,8-cineole, and the sesquiterpenes (C15) bisabolene and farnesene, among others.

From an energetic perspective, microbial biofuel production is a biotransformation of a feedstock—a critical reason why photoautotrophic cyanobacteria and algae have historically attracted significant research attention. On the other hand, methylotrophs or lignolytic organisms must also compete thermodynamically with simply combusting lignocellulose or methane, respectively. As with any energy transformation, the efficiency or energy return on energy invested (EROI) of these bioconversions may dictate their viability (Fei et al., 2014). While most microbial biofuel platforms continue to use pure C5 and C6 sugar feedstocks, they remain instrumental to actualizing advanced biofuel production. Here, we focus on isoprenoid biofuels that have demonstrated high titer and are approaching commercial viability.

Bisabolene

While isoprenoid derived biofuels have attracted significant attention as fuel enhancing additives, their titers and scalability of current technologies remain variable. Bisabolene, for example, demonstrated yields of approximately 1 g/L in *S. cerevisiae* and *E. coli*, by careful MVA pathway balancing and quorum sensing mediated pathway expression, respectively (Peralta-Yahya et al., 2011; Kim et al., 2017). Production has also expanded into *R. toruloides*, an oleaginous yeast with attractive natural characteristics for production on unfiltered sorghum hydrolysate (Sundstrom et al., 2018). Optimization of this microbe with multiple genomic copies of the MVA pathway resulted in 2.6 g/L bisabolene in a 2-L fermentation reaction, accounting for approximately 10% theoretical yield (Kirby et al., 2021).

Limonene

Production of limonene in *E. coli* was demonstrated by Alonso-Gutierrez et al. via fine tuning of MVA pathway genes on a plasmid also harboring a limonene synthase from *Mentha spicata* and a truncated geranyl diphosphate synthase from *Abies grandis* (JBEI-6410). Pathway expression in *E. coli* DH1 ultimately achieved a titer of approximately 435 mg/L limonene on 1% glucose (compared to 0.32 g-limonene/g-glucose theoretical maximum) (Alonso-Gutierrez et al., 2013) with a similar titer recently achieved in *R. toruloides* (Liu S. et al., 2021). Titer was further improved to 3.6 g/L by tuning culture conditions with mixed glucose and glycerol feeding in *E. coli* BL21 (DE3) (Rolf et al., 2020). The dramatic improvement was hypothesized to derive from combined flux through both the endogenous MEP and heterologous MVA pathways (Rolf et al., 2020). Lastly, a recent work explored expression in *S. cerevisiae* by exploiting metabolite sequestration to peroxisomes, thereby partitioning product toxicity while maintaining proximal generation of acetyl-CoA (Dusséaux et al., 2020). The approach ultimately achieved a titer of 2.6 g/L under fed-batch conditions in synthetic medium (Dusséaux et al., 2020).

These recent improvements have made remarkable strides towards economic viability, though TEAs still estimate the MSP of limonene to be between 20 $/kg and ~7 $/kg if yields are improved to 45% (0.144 g-limonene/g-glucose) (Sun et al., 2020) or 30% (0.096 g-limonene/g-glucose) if coupled with significant feed and culture optimizations (Wu and Maravelias, 2018). Both figures are far higher than current methodologies, though highlight the burden of pure sugar substrate on overall cost.

Isoprenol

The C5 alcohols isoprenol and prenol have enormous potential as biofuel additives and precursors. Prenol has demonstrated a unique blendstock characteristic called hyperboosting in which its addition increases the research octane number (RON) of the blendstock above the RON of the individual components (Monroe et al., 2019). On the other hand, isoprenol can also serve as a precursor to 1,4-dimethylcyclooctane (DMCO), a drop-in jet fuel (Baral et al., 2021). From a metabolic perspective, isoprenol is generated by a simple sequential dephosphorylation of IPP. However, IPP accumulation is inhibitory such that high titer isoprenol bioproduction has proven challenging. Recently, heterologous expression and subsequent mutagenesis of the *S. cerevisiae*

mevalonate diphosphate decarboxylase enzyme was demonstrated to avoiding intracellular accumulation of IPP, a known toxic intermediate hypothesized to limit overall isoprenol efficiency (Kang et al., 2017; George et al., 2018). This "IPP-bypass" enhanced isoprenol titer by 2.4-fold to 1.1 g/L (Kang et al., 2017). Further incorporation of an optimized upstream MVA pathway into *E. coli* DH1 also harboring acetate pathway knockouts (Δ*ackA*, Δ*pta*, and Δ*poxB*) resulted in 10.38 g/L titer in 2-L fermenters on minimal medium (0.105 g-isoprenol/g-glucose and a productivity of 0.157 g/L/hr) (Kang et al., 2019). This titer is especially encouraging for production of DMCO, with baseline unoptimized metrics of 9.0 $/L-Jet-A-eq and 61.4 g CO_2eq/MJ (Baral et al., 2021). Another work demonstrated isoprenol production in yeast by knocking out an endogenous kinase and overexpressing a heterologous phosphatase to yield 380 mg/L isoprenol (Kim J. et al., 2021).

Polyhydroxyalkanoates

Polyhydroxyalkanoates (PHAs) are a large class of natural polyesters including short chain length (scl) monomers poly (3-hydroxybutyrate) (PHB), poly (3-hydroxyvalerate) (PHV), and their co-polymer PHBV with poor thermal and mechanical properties (Li et al., 2016). PHAs are an amalgam of fatty acids, with C5 and C4 acids generating the scl-PHAs of PHV and PHB, respectively. The fatty acids themselves may be generated through varied metabolic pathways including fatty acid synthesis and β-oxidation, which may also generate medium and long chain length PHAs (Mezzina et al., 2021). A nominal production strategy for PHB is a three-step process initiating with a Claisen condensation of two acetyl-CoA molecules to form acetoacetyl-CoA, reduction to 3-hydroxybutyryl-CoA, and polymerization to PHB via phaA, phaB and phaC, respectively. PHV may proceed via a similar pathway, though with a propionyl-CoA precursor. PHAs are commonly produced by bacteria as a carbon storage mechanism under nutrient stress, though exhibit varying degrees of biodegradability and elasticity due to differences in monomeric composition and chain length. While these characteristics limit the use of PHAs for thermomechanical applications, PHAs have attracted attention as single-use bioplastics for decades.

Accordant with most of the bioproducts described in this review, the commercialization of PHAs has been hampered by feedstock costs, low conversion, and bioprocessing demands amounting to approximately 4–6 $/kg, values far higher than comparable petrochemical products like polyethylene (Chen et al., 2020; Tan et al., 2021). The vast majority of PHA production strategies have been conducted on glucose and are typically limited to *E. coli* and *C. necator*, both of which have achieved titers well over 100 g/L (Zheng et al., 2020; Tan et al., 2021). Metabolic engineering of atypical production organisms, namely the halotolerant *Halomonas* sp., *P. putida*, and *M. extorquens*, has arisen as a promising opportunity for reducing culturing costs and enabling consumption of inexpensive feedstocks. As previously noted, *P. putida* has gained traction as a metabolic chassis due to its metabolism of mixed aromatics and potential for valorizing lignin (Mezzina et al., 2021). Feeding alkaline pretreated liquor, a medium rich in lignin monomers, to *P. putida* demonstrated 34% and 39% conversion of *p*-coumaric acid and ferulic acid to mcl-PHAs, respectively (Linger et al., 2014). Overall titer was then improved to ~1 g/L mcl-PHAs on pretreated corn stover with 77.6% yield from lignin in the liquid stream (Liu et al., 2017). Using a feedstock of waste cooking oil and a carboxylic acid transport gene pathway knockout (*actA*) resulted in a titer of 1.91 g/L mcl-PHAs (Borrero-de Acuña et al., 2018). Likewise, *M. extorquens* AM1 demonstrated PHA accumulation on methanol, ultimately achieving 43.6% PHAs by weight with 96.6 mol% PHV through overexpression of *phaAB* though at low overall titers (Orita et al., 2014). Although titers in both strains are well below economic viability, they provide low TRL opportunities for improvement and, ideally, significant overall cost reduction.

Perspectives and conclusion

In this review we have elaborated on the challenges associated with *ex ante* life cycle and technoeconomic analysis of emergent, low TRL biotechnologies while highlighting significant synthetic biology advances that have elevated the sustainable and economic viability of certain bioproduction platforms. Assessments of microbial bioproduction are often surmised by a core set of parameters - feedstock burden as well as fermentation titer, rate, and yield. Advances in metabolic flux analyses, -omics studies, and genome scale modeling have also facilitated construction of computationally informed theoretical or stoichiometric approaches to guide and inform laboratory research. While scaling of the current state of technology still often lends to platform infeasibility, the associated sensitivity analyses are vital for guiding future metabolic engineering and have realized astonishing improvements in many bioproduction pathways.

Nonetheless, the plurality of bioproduction assessments of emergent technologies poses a challenge to comparative analyses due to the sheer range of inputs, strategies, and overall frameworks, which are especially true for LCAs (Cruce et al., 2021). Assessments usually require simplification of the system. The paucity of collaborative studies by experts in the fields of biology and industrial ecology can lend to complications in design accuracy, interpretation, and poor policy decision making (DeCicco et al., 2016). Extrapolating lessons learned from algal biofuels, namely encouraging multidisciplinary collaboration and early LCA-TEA integration to better understand trade-offs, input sensitivity, and viability at scale is necessary for spurring research investment into promising platforms (Mahmud et al., 2021).

Some critics assert that a simple translation of the LCA framework to synthetic biology is fundamentally flawed due to

the incompatible extrapolation of materials accounting to biological systems, blurring of ecology and industry, and the intersection of engineered organisms with the natural environment (Seager et al., 2017). And indeed, the recapitulation of the LCA framework for biological systems demands integration not only of industrial ecology and biology, but of physical and social sciences to better inform the implications and applications of synthetic biology, which will hopefully guide governmental policy (Trump et al., 2019). Fortunately, these factors may readily be addressed through careful determination of impact factors. Many LCAs already include esoteric factors like biodiversity loss and human health but have only recently ventured into other critical socio-political factors like environmental justice and a social cost of carbon. Incorporation of a social cost of carbon in assessments will help valorize emissions reduction, facilitating translation between LCA and TEA metrics.

Assessments of bioproduction platforms may also be ameliorated by a combination of transparency and sensitivity to temper theoretical and demonstrated production in the context of environmental risk (e.g., biosecurity, horizontal gene transfer, escape mutants, etc.). The increased uncertainty of bioproduction is a characteristic of biological systems themselves. For example, scaled *E. coli* bioproduction is commonly avoided due to concerns with strain stability, product toxicity, limited feedstock pool, poor production of proteins, and metabolite feedback inhibition (Neubauer et al., 1995; Calero and Nikel, 2019). While the uncertainty is not to be understated, the quantification of economic and environmental tradeoffs is nonetheless critical for benchmarking processes and guiding policies. Effective integration of synthetic biology with environmental and economic assessments is critical for actualizing industrial biotechnology and decarbonizing bulk chemical production.

Author contributions

DC and TL contributed to conception and design of the study. DC wrote the first draft of the manuscript and TL supervised the manuscript writing. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This work was supported by the DOE Joint BioEnergy Institute (http://www.jbei.org), supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC0205CH11231 with Lawrence Berkeley National Laboratory. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.

References

Adom, F. K., and Dunn, J. B. (2017). Life cycle analysis of corn-stover-derived polymer-grade l-lactic acid and ethyl lactate: Greenhouse gas emissions and fossil energy consumption. *Biofuels, Bioprod. Bioref.* 11, 258–268. doi:10.1002/bbb.1734

Ahkami, A. H., Allen White, R., Handakumbura, P. P., and Jansson, C. (2017). Rhizosphere engineering: Enhancing sustainable plant ecosystem productivity. *Rhizosphere* 3, 233–243. doi:10.1016/j.rhisph.2017.04.012

Ahn, J. H., Bang, J., Kim, W. J., and Lee, S. Y. (2017). Formic acid as a secondary substrate for succinic acid production by metabolically engineered Mannheimia succiniciproducens. *Biotechnol. Bioeng.* 114, 2837–2847. doi:10.1002/bit.26435

Alonso-Gutierrez, J., Chan, R., Batth, T. S., Adams, P. D., Keasling, J. D., Petzold, C. J., et al. (2013). Metabolic engineering of *Escherichia coli* for limonene and perillyl alcohol production. *Metab. Eng.* 19, 33–41. doi:10.1016/j.ymben.2013.05.004

Anastas, P. T., and Lankey, R. L. (2000). Life cycle assessment and green chemistry: The yin and yang of industrial ecology. *Green Chem.* 2, 289–295. doi:10.1039/b005650m

Aurisano, N., Weber, R., and Fantke, P. (2021). Enabling a circular economy for chemicals in plastics. *Curr. Opin. Green Sustain. Chem.* 31, 100513. doi:10.1016/j.cogsc.2021.100513

Baker, M. (2016). 1, 500 scientists lift the lid on reproducibility. *Nature* 533, 452–454. doi:10.1038/533452a

Balakrishnan, M., Sacia, E. R., Sreekumar, S., Gunbas, G., Gokhale, A. A., Scown, C. D., et al. (2015). Novel pathways for fuels and lubricants from biomass optimized using life-cycle greenhouse gas assessment. *Proc. Natl. Acad. Sci. U. S. A.* 112, 7645–7649. doi:10.1073/pnas.1508274112

Banerjee, D., Eng, T., Lau, A. K., Sasaki, Y., Wang, B., Chen, Y., et al. (2020). Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale. *Nat. Commun.* 11, 5385. doi:10.1038/s41467-020-19171-4

Bar-Even, A. (2016). Formate assimilation: The metabolic architecture of natural and synthetic pathways. *Biochemistry* 55, 3851–3863. doi:10.1021/acs.biochem.6b00495

Baral, N. R., Sundstrom, E. R., Das, L., Gladden, J. M., Eudes, A., Mortimer, J., et al. (2019). Approaches for more efficient biological conversion of lignocellulosic feedstocks to biofuels and bioproducts. *ACS Sustain. Chem. Eng.* 7, 9062–9079. doi:10.1021/acssuschemeng.9b01229

Baral, N. R., Yang, M., Harvey, B. G., Simmons, B. A., Mukhopadhyay, A., Lee, T. S., et al. (2021). Production cost and carbon footprint of biomass-derived dimethylcyclooctane as a high performance jet fuel blendstock. *ACS Sustain. Chem. Eng.* 9, 11872–11882. doi:10.26434/chemrxiv.14761002.v1

Bare, J. (2011). Traci 2.0: The tool for the reduction and assessment of chemical and other environmental impacts 2.0. *Clean. Techn. Environ. Policy* 13, 687–696. doi:10.1007/s10098-010-0338-9

Becker, J., Zelder, O., Häfner, S., Schröder, H., and Wittmann, C. (2011). From zero to hero--design-based systems metabolic engineering of Corynebacterium glutamicum for L-lysine production. *Metab. Eng.* 13, 159–168. doi:10.1016/j.ymben.2011.01.003

Beckham, G. T., Johnson, C. W., Karp, E. M., Salvachúa, D., and Vardon, D. R. (2016). Opportunities and challenges in biological lignin valorization. *Curr. Opin. Biotechnol.* 42, 40–53. doi:10.1016/j.copbio.2016.02.030

Bello, S., Salim, I., Feijoo, G., and Moreira, M. T. (2021). Inventory review and environmental evaluation of first- and second-generation sugars through life cycle assessment. *Environ. Sci. Pollut. Res.* 28, 27345–27361. doi:10.1007/s11356-021-12405-y

Benavides, P. T., Klein, B., and Davis, R. (2021). Economic and environmental assessment of biological conversions of agile BioFoundry (ABF) bio-derived chemicals. Available at: https://www.nrel.gov/docs/fy21osti/78106.pdf.

Bhagwat, S. S., Li, Y., Cortés-Peña, Y. R., Brace, E. C., Martin, T. A., Zhao, H., et al. (2021). Sustainable production of acrylic acid via 3-hydroxypropionic acid from lignocellulosic biomass. *ACS Sustain. Chem. Eng.* 9, 16659–16669. doi:10.1021/acssuschemeng.1c05441

Biddy, M. J., Davis, R., Humbird, D., Tao, L., Dowe, N., Guarnieri, M. T., et al. (2016). The techno-economic basis for coproduct manufacturing to enable hydrocarbon fuel production from lignocellulosic biomass. *ACS Sustain. Chem. Eng.* 4, 3196–3211. doi:10.1021/acssuschemeng.6b00243

Blank, L. M., Narancic, T., Mampel, J., Tiso, T., and O'Connor, K. (2020). Biotechnological upcycling of plastic waste and other non-conventional feedstocks in a circular economy. *Curr. Opin. Biotechnol.* 62, 212–219. doi:10.1016/j.copbio.2019.11.011

Borrero-de Acuña, J. M., Aravena-Carrasco, C., Gutierrez-Urrutia, I., Duchens, D., and Poblete-Castro, I. (2018). Enhanced synthesis of medium-chain-length poly(3-hydroxyalkanoates) by inactivating the tricarboxylate transport system of Pseudomonas putida KT2440 and process development using waste vegetable oil. *Process Biochem.* 77, 23–30. doi:10.1016/j.procbio.2018.10.012

Bozell, J. J., and Petersen, G. R. (2010). Technology development for the production of biobased products from biorefinery carbohydrates—The US department of energy's "top 10" revisited. *Green Chem.* 12, 539. doi:10.1039/b922014c

Buchner, G. A., Zimmermann, A. W., Hohgräve, A. E., and Schomäcker, R. (2018). Techno-economic assessment framework for the chemical industry—based on technology readiness levels. *Ind. Eng. Chem. Res.* 57, 8502–8517. doi:10.1021/acs.iecr.8b01248

Bugg, T. D. H., Ahmad, M., Hardiman, E. M., and Rahmanpour, R. (2011). Pathways for degradation of lignin in bacteria and fungi. *Nat. Prod. Rep.* 28, 1883–1896. doi:10.1039/c1np00042j

Bui, M., Adjiman, C. S., Bardow, A., Anthony, E. J., Boston, A., Brown, S., et al. (2018). Carbon capture and storage (CCS): The way forward. *Energy Environ. Sci.* 11, 1062–1176. doi:10.1039/C7EE02342A

Calero, P., and Nikel, P. I. (2019). Chasing bacterial chassis for metabolic engineering: A perspective review from classical to non-traditional microorganisms. *Microb. Biotechnol.* 12, 98–124. doi:10.1111/1751-7915.13292

Cameron, D. E., Bashor, C. J., and Collins, J. J. (2014). A brief history of synthetic biology. *Nat. Rev. Microbiol.* 12, 381–390. doi:10.1038/nrmicro3239

Carruthers, D. N., Godwin, C. M., Hietala, D. C., Cardinale, B. J., Lin, X. N., and Savage, P. E. (2019). Biodiversity improves life cycle sustainability metrics in algal biofuel production. *Environ. Sci. Technol.* 53, 9279–9288. doi:10.1021/acs.est.9b00909

Carruthers, D. N., and Lee, T. S. (2021). Diversifying isoprenoid platforms via atypical carbon substrates and non-model microorganisms. *Front. Microbiol.* 12, 791089. doi:10.3389/fmicb.2021.791089

Chaplin-Kramer, R., Sim, S., Hamel, P., Bryant, B., Noe, R., Mueller, C., et al. (2017). Life cycle assessment needs predictive spatial modelling for biodiversity and ecosystem services. *Nat. Commun.* 8, 15065. doi:10.1038/ncomms15065

Chen, G.-Q., Chen, X.-Y., Wu, F.-Q., and Chen, J.-C. (2020). Polyhydroxyalkanoates (PHA) toward cost competitiveness and functionality. *Adv. Industrial Eng. Polym. Res.* 3, 1–7. doi:10.1016/j.aiepr.2019.11.001

Cherubini, F., Bird, N. D., Cowie, A., Jungmeier, G., Schlamadinger, B., and Woess-Gallasch, S. (2009). Energy- and greenhouse gas-based LCA of biofuel and bioenergy systems: Key issues, ranges and recommendations. *Resour. Conservation Recycl.* 53, 434–447. doi:10.1016/j.resconrec.2009.03.013

Clomburg, J. M., Crumbley, A. M., and Gonzalez, R. (2017). Industrial biomanufacturing: The future of chemical production. *Science* 355, aag0804. doi:10.1126/science.aag0804

Corona, A., Biddy, M. J., Vardon, D. R., Birkved, M., Hauschild, M. Z., and Beckham, G. T. (2018). Life cycle assessment of adipic acid production from lignin. *Green Chem.* 20, 3857–3866. doi:10.1039/C8GC00868J

Cortes-Peña, Y., Kumar, D., Singh, V., and Guest, J. S. (2020). BioSTEAM: A fast and flexible platform for the design, simulation, and techno-economic analysis of biorefineries under uncertainty. *ACS Sustain. Chem. Eng.* 8, 3302–3310. doi:10.1021/acssuschemeng.9b07040

Cruce, J. R., Beattie, A., Chen, P., Quiroz, D., Somers, M., Compton, S., et al. (2021). Driving toward sustainable algal fuels: A harmonization of techno-economic and life cycle assessments. *Algal Res.* 54, 102169. doi:10.1016/j.algal.2020.102169

Cywar, R. M., Rorrer, N. A., Hoyt, C. B., Beckham, G. T., and Chen, E. Y.-X. (2021). Bio-based polymers with performance-advantaged properties. *Nat. Rev. Mat.* 7, 83–103. doi:10.1038/s41578-021-00363-3

Czajka, J. J., Okumuş, B., Koffas, M. A. G., Blenner, M., and Tang, Y. J. (2020). Mitigation of host cell mutations and regime shift during microbial fermentation: a perspective from flux memory. *Curr. Opin. Biotechnol.* 66, 227–235.

Davis, R., Tao, L., Tan, E. C. D., Biddy, M. J., Beckham, G. T., and Scarlata, C. (2013). *Process design and economics for the conversion of lignocellulosic biomass to hydrocarbons: Dilute -acid and enzymatic deconstruction of biomass to sugars and biological conversion of sugars to hydrocarbons*. Golden, CO: National Renewable Energy Laboratory.

DeCicco, J. M., Liu, D. Y., Heo, J., Krishnan, R., Kurthen, A., and Wang, L. (2016). Carbon balance effects of U.S. biofuel production and use. *Clim. Change* 138, 667–680. doi:10.1007/s10584-016-1764-4

Deutscher, J. (2008). The mechanisms of carbon catabolite repression in bacteria. *Curr. Opin. Microbiol.* 11, 87–93. doi:10.1016/j.mib.2008.02.007

Dickson, R., Mancini, E., Garg, N., Woodley, J. M., Gernaey, K. V., Pinelo, M., et al. (2021). Sustainable bio-succinic acid production: Superstructure optimization, techno-economic, and lifecycle assessment. *Energy Environ. Sci.* 14, 3542–3558. doi:10.1039/D0EE03545A

Ducat, D. C., Avelar-Rivas, J. A., Way, J. C., and Silver, P. A. (2012). Rerouting carbon flux to enhance photosynthetic productivity. *Appl. Environ. Microbiol.* 78, 2660–2668. doi:10.1128/AEM.07901-11

Dunn, J. B., Adom, F., Sather, N., Han, J., Snyder, S., He, C., et al. (2015). *Life-cycle analysis of bioproducts and their conventional counterparts in GREET*. Argonne, IL (United States): Argonne National Laboratory. doi:10.2172/1250468

Dunn, J. B. (2019). Biofuel and bioproduct environmental sustainability analysis. *Curr. Opin. Biotechnol.* 57, 88–93. doi:10.1016/j.copbio.2019.02.008

Dusséaux, S., Wajn, W. T., Liu, Y., Ignea, C., and Kampranis, S. C. (2020). Transforming yeast peroxisomes into microfactories for the efficient production of high-value isoprenoids. *Proc. Natl. Acad. Sci. U. S. A.* 117, 31789–31799. doi:10.1073/pnas.2013968117

Earles, J. M., and Halog, A. (2011). Consequential life cycle assessment: A review. *Int. J. Life Cycle Assess.* 16, 445–453. doi:10.1007/s11367-011-0275-9

Elmore, J. R., Dexter, G. N., Salvachúa, D., O'Brien, M., Klingeman, D. M., Gorday, K., et al. (2020). Engineered Pseudomonas putida simultaneously catabolizes five major components of corn stover lignocellulose: Glucose, xylose, arabinose, p-coumaric acid, and acetic acid. *Metab. Eng.* 62, 62–71. doi:10.1016/j.ymben.2020.08.001

Espinosa, M. I., Gonzalez-Garcia, R. A., Valgepea, K., Plan, M. R., Scott, C., Pretorius, I. S., et al. (2020). Adaptive laboratory evolution of native methanol assimilation in *Saccharomyces cerevisiae*. *Nat. Commun.* 11, 5564. doi:10.1038/s41467-020-19390-9

Faber, G., Mangin, C., and Sick, V. (2021). Life cycle and techno-economic assessment templates for emerging carbon management technologies. *Front. Sustain.* 2. doi:10.3389/frsus.2021.764057

Fei, Q., Guarnieri, M. T., Tao, L., Laurens, L. M. L., Dowe, N., and Pienkos, P. T. (2014). Bioconversion of natural gas to liquid fuel: Opportunities and challenges. *Biotechnol. Adv.* 32, 596–614. doi:10.1016/j.biotechadv.2014.03.011

Fei, Q., Liang, B., Tao, L., Tan, E. C. D., Gonzalez, R., Henard, C., et al. (2020). Biological valorization of natural gas for the production of lactic acid: Techno-economic analysis and life cycle assessment. *Biochem. Eng. J.* 158. doi:10.1016/j.bej.2020.107500

Figueroa, I. A., Barnum, T. P., Somasekhar, P. Y., Carlström, C. I., Engelbrektson, A. L., and Coates, J. D. (2018). Metagenomics-guided analysis of microbial chemolithoautotrophic phosphite oxidation yields evidence of a seventh natural CO2 fixation pathway. *Proc. Natl. Acad. Sci. U. S. A.* 115, E92–E101. doi:10.1073/pnas.1715549114

Garg, S., Clomburg, J. M., and Gonzalez, R. (2018). A modular approach for high-flux lactic acid production from methane in an industrial medium using engineered Methylomicrobium buryatense 5GB1. *J. Ind. Microbiol. Biotechnol.* 45, 379–391. doi:10.1007/s10295-018-2035-3

Garrido-Oter, R., Nakano, R. T., Dombrowski, N., Ma, K.-W., McHardy, A. C., and Schulze-Lefert, P. (2018). Modular traits of the rhizobiales root microbiota and their evolutionary relationship with symbiotic rhizobia. *Cell Host Microbe* 24, 155–167. e5. doi:10.1016/j.chom.2018.06.006

George, K. W., Thompson, M. G., Kim, J., Baidoo, E. E. K., Wang, G., Benites, V. T., et al. (2018). Integrated analysis of isopentenyl pyrophosphate (IPP) toxicity in isoprenoid-producing *Escherichia coli*. *Metab. Eng.* 47, 60–72. doi:10.1016/j.ymben.2018.03.004

Gonzalez de la Cruz, J., Machens, F., Messerschmidt, K., and Bar-Even, A. (2019). Core catalysis of the reductive glycine pathway demonstrated in yeast. *ACS Synth. Biol.* 8, 911–917. doi:10.1021/acssynbio.8b00464

González-Morales, S. I., Pacheco-Gutiérrez, N. B., Ramírez-Rodríguez, C. A., Brito-Bello, A. A., Estrella-Hernández, P., Herrera-Estrella, L., et al. (2020). Metabolic engineering of phosphite metabolism in Synechococcus elongatus PCC 7942 as an effective measure to control biological contaminants in outdoor raceway ponds. *Biotechnol. Biofuels* 13, 119. doi:10.1186/s13068-020-01759-z

Greet (2022). Argonne GREET model. Available at: https://greet.es.anl.gov/ (Accessed January 31, 2022).

Guinée, J. B., Heijungs, R., Huppes, G., Zamagni, A., Masoni, P., Buonamici, R., et al. (2011). Life cycle assessment: Past, present, and future. *Environ. Sci. Technol.* 45, 90–96. doi:10.1021/es101316v

Gunukula, S., and Anex, R. P. (2017). Techno-economic analysis of multiple bio-based routes to adipic acid. *Biofuels, Bioprod. Bioref.* 11, 897–907. doi:10.1002/bbb.1797

Han, T., Kim, G. B., and Lee, S. Y. (2020). Glutaric acid production by systems metabolic engineering of an l-lysine-overproducing Corynebacterium glutamicum. *Proc. Natl. Acad. Sci. U. S. A.* 117, 30328–30334. doi:10.1073/pnas.2017483117

Handler, R. M., Shonnard, D. R., Griffing, E. M., Lai, A., and Palou-Rivera, I. (2016). Life cycle assessments of ethanol production via gas fermentation: Anticipated greenhouse gas emissions for cellulosic and waste gas feedstocks. *Ind. Eng. Chem. Res.* 55, 3253–3261. doi:10.1021/acs.iecr.5b03215

Hanssen, S. V., Daioglou, V., Steinmann, Z. J. N., Doelman, J. C., Van Vuuren, D. P., and Huijbregts, M. A. J. (2020). The climate change mitigation potential of bioenergy with carbon capture and storage. *Nat. Clim. Chang.* 10, 1023–1029. doi:10.1038/s41558-020-0885-y

He, H., Höper, R., Dodenhöft, M., Marlière, P., and Bar-Even, A. (2020). An optimized methanol assimilation pathway relying on promiscuous formaldehyde-condensing aldolases in *E. coli*. *Metab. Eng.* 60, 1–13. doi:10.1016/j.ymben.2020.03.002

Heijstra, B. D., Leang, C., and Juminaga, A. (2017). Gas fermentation: Cellular engineering possibilities and scale up. *Microb. Cell Fact.* 16, 60. doi:10.1186/s12934-017-0676-y

Hellweg, S., and Milà i Canals, L. (2014). Emerging approaches, challenges and opportunities in life cycle assessment. *Science* 344, 1109–1113. doi:10.1126/science.1248361

Hillson, N., Caddick, M., Cai, Y., Carrasco, J. A., Chang, M. W., Curach, N. C., et al. (2019). Building a global alliance of biofoundries. *Nat. Commun.* 10, 2040. doi:10.1038/s41467-019-10079-2

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Guillén Bolaños, T., Bindi, M., Brown, S., et al. (2019). The human imperative of stabilizing global climate change at 1.5°C. *Science* 365, eaaw6974. doi:10.1126/science.aaw6974

Huang, K., Peng, X., Kong, L., Wu, W., Chen, Y., and Maravelias, C. T. (2021). Greenhouse gas emission mitigation potential of chemicals produced from biomass. *ACS Sustain. Chem. Eng.* 9, 14480–14487. doi:10.1021/acssuschemeng.1c04836

Huijbregts, M. A. J., Steinmann, Z. J. N., Elshout, P. M. F., Stam, G., Verones, F., Vieira, M., et al. (2017). ReCiPe2016: A harmonised life cycle impact assessment method at midpoint and endpoint level. *Int. J. Life Cycle Assess.* 22, 138–147. doi:10.1007/s11367-016-1246-y

Jansen, M. L. A., and van Gulik, W. M. (2014). Towards large scale fermentative production of succinic acid. *Curr. Opin. Biotechnol.* 30, 190–197. doi:10.1016/j.copbio.2014.07.003

Johnson, C. W., Salvachúa, D., Rorrer, N. A., Black, B. A., Vardon, D. R., St. John, P. C., et al. (2019). Innovative chemicals and materials from bacterial aromatic catabolic pathways. *Joule* 3, 1523–1537. doi:10.1016/j.joule.2019.05.011

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2

Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., et al. (2003). The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J. Biotechnol.* 104, 5–25. doi:10.1016/S0168-1656(03)00154-8

Kang, A., Meadows, C. W., Canu, N., Keasling, J. D., and Lee, T. S. (2017). High-throughput enzyme screening platform for the IPP-bypass mevalonate pathway for isopentenol production. *Metab. Eng.* 41, 125–134. doi:10.1016/j.ymben.2017.03.010

Kang, A., Mendez-Perez, D., Goh, E.-B., Baidoo, E. E. K., Benites, V. T., Beller, H. R., et al. (2019). Optimization of the IPP-bypass mevalonate pathway and fed-batch fermentation for the production of isoprenol in *Escherichia coli*. *Metab. Eng.* 56, 85–96. doi:10.1016/j.ymben.2019.09.003

Keasling, J., Garcia Martin, H., Lee, T. S., Mukhopadhyay, A., Singer, S. W., and Sundstrom, E. (2021). Microbial production of advanced biofuels. *Nat. Rev. Microbiol.* 19, 701–715. doi:10.1038/s41579-021-00577-w

Kim, E.-M., Woo, H. M., Tian, T., Yilmaz, S., Javidpour, P., Keasling, J. D., et al. (2017). Autonomous control of metabolic state by a quorum sensing (QS)-mediated regulator for bisabolene production in engineered *E. coli*. *Metab. Eng.* 44, 325–336. doi:10.1016/j.ymben.2017.11.004

Kim, J., Baidoo, E. E. K., Amer, B., Mukhopadhyay, A., Adams, P. D., Simmons, B. A., et al. (2021a). Engineering *Saccharomyces cerevisiae* for isoprenol production. *Metab. Eng.* 64, 154–166. doi:10.1016/j.ymben.2021.02.002

Kim, S., Lindner, S. N., Aslan, S., Yishai, O., Wenk, S., Schann, K., et al. (2020). Growth of *E. coli* on formate and methanol via the reductive glycine pathway. *Nat. Chem. Biol.* 16, 538–545. doi:10.1038/s41589-020-0473-5

Kim, Y., Lama, S., Agrawal, D., Kumar, V., and Park, S. (2021b). Acetate as a potential feedstock for the production of value-added chemicals: Metabolism and applications. *Biotechnol. Adv.* 49, 107736. doi:10.1016/j.biotechadv.2021.107736

Kind, S., Jeong, W. K., Schröder, H., Zelder, O., and Wittmann, C. (2010). Identification and elimination of the competing N-acetyldiaminopentane pathway for improved production of diaminopentane by Corynebacterium glutamicum. *Appl. Environ. Microbiol.* 76, 5175–5180. doi:10.1128/AEM.00834-10

Kirby, J., Geiselman, G. M., Yaegashi, J., Kim, J., Zhuang, X., Tran-Gyamfi, M. B., et al. (2021). Further engineering of R. toruloides for the production of terpenes from lignocellulosic biomass. *Biotechnol. Biofuels* 14, 101. doi:10.1186/s13068-021-01950-w

Kirst, H., Formighieri, C., and Melis, A. (2014). Maximizing photosynthetic efficiency and culture productivity in cyanobacteria upon minimizing the phycobilisome light-harvesting antenna size. *Biochimica Biophysica Acta - Bioenergetics* 1837, 1653–1664. doi:10.1016/j.bbabio.2014.07.009

Köpke, M., Mihalcea, C., Bromley, J. C., and Simpson, S. D. (2011a). Fermentative production of ethanol from carbon monoxide. *Curr. Opin. Biotechnol.* 22, 320–325. doi:10.1016/j.copbio.2011.01.005

Köpke, M., Mihalcea, C., Liew, F., Tizard, J. H., Ali, M. S., Conolly, J. J., et al. (2011b). 2, 3-butanediol production by acetogenic bacteria, an alternative route to chemical synthesis, using industrial waste gas. *Appl. Environ. Microbiol.* 77, 5467–5475. doi:10.1128/AEM.00355-11

Korosh, T. C., Markley, A. L., Clark, R. L., McGinley, L. L., McMahon, K. D., and Pfleger, B. F. (2017). Engineering photosynthetic production of L-lysine. *Metab. Eng.* 44, 273–283. doi:10.1016/j.ymben.2017.10.010

Krömer, J. O., Ferreira, R. G., Petrides, D., and Kohlheb, N. (2020). Economic process evaluation and environmental life-cycle assessment of bio-aromatics production. *Front. Bioeng. Biotechnol.* 8, 403. doi:10.3389/fbioe.2020.00403

Kruyer, N. S., and Peralta-Yahya, P. (2017). Metabolic engineering strategies to bio-adipic acid production. *Curr. Opin. Biotechnol.* 45, 136–143. doi:10.1016/j.copbio.2017.03.006

Kumar, A. K., and Sharma, S. (2017). Recent updates on different methods of pretreatment of lignocellulosic feedstocks: A review. *Bioresour. Bioprocess.* 4, 7. doi:10.1186/s40643-017-0137-9

Kwok, R. (2010). Five hard truths for synthetic biology. *Nature* 463, 288–290. doi:10.1038/463288a

Lawson, C. E., Harcombe, W. R., Hatzenpichler, R., Lindemann, S. R., Löffler, F. E., O'Malley, M. A., et al. (2019). Common principles and best practices for engineering microbiomes. *Nat. Rev. Microbiol.* 17, 725–741. doi:10.1038/s41579-019-0255-9

Lee, J.-Y., Lee, S.-E., and Lee, D.-W. (2021). Current status and future prospects of biological routes to bio-based products using raw materials, wastes, and residues as renewable resources. *Crit. Rev. Environ. Sci. Technol.* 52, 2453–2509. doi:10.1080/10643389.2021.1880259

Lee, S. Y., and Kim, H. U. (2015). Systems strategies for developing industrial microbial strains. *Nat. Biotechnol.* 33, 1061–1072. doi:10.1038/nbt.3365

Lewis, N. S., and Nocera, D. G. (2006). Powering the planet: Chemical challenges in solar energy utilization. *Proc. Natl. Acad. Sci. U. S. A.* 103, 15729–15735. doi:10.1073/pnas.0603395103

Li, H., Opgenorth, P. H., Wernick, D. G., Rogers, S., Wu, T.-Y., Higashide, W., et al. (2012). Integrated electromicrobial conversion of CO2 to higher alcohols. *Science* 335, 1596. doi:10.1126/science.1217643

Li, Y., Bhagwat, S. S., Cortés-Peña, Y. R., Ki, D., Rao, C. V., Jin, Y.-S., et al. (2021). Sustainable lactic acid production from lignocellulosic biomass. *ACS Sustain. Chem. Eng.* 9, 1341–1351. doi:10.1021/acssuschemeng.0c08055

Li, Z., Dong, Y., Liu, Y., Cen, X., Liu, D., and Chen, Z. (2022). Systems metabolic engineering of Corynebacterium glutamicum for high-level production of 1, 3-propanediol from glucose and xylose. *Metab. Eng.* 70, 79–88. doi:10.1016/j.ymben.2022.01.006

Li, Z., Yang, J., and Loh, X. J. (2016). Polyhydroxyalkanoates: Opening doors for a sustainable future. *NPG Asia Mat.* 8, e265. doi:10.1038/am.2016.48

Liang, H., Ma, X., Ning, W., Liu, Y., Sinskey, A. J., Stephanopoulos, G., et al. (2021). Constructing an ethanol utilization pathway in *Escherichia coli* to produce acetyl-CoA derived compounds. *Metab. Eng.* 65, 223–231. doi:10.1016/j.ymben.2020.11.010

Liew, F., Martin, M. E., Tappel, R. C., Heijstra, B. D., Mihalcea, C., and Köpke, M. (2016). Gas fermentation-A flexible platform for commercial scale production of low-carbon-fuels and chemicals from waste and renewable feedstocks. *Front. Microbiol.* 7, 694. doi:10.3389/fmicb.2016.00694

Lim, H. G., Eng, T., Banerjee, D., Alarcon, G., Lau, A. K., Park, M.-R., et al. (2021). Generation of *Pseudomonas putida* KT2440 strains with efficient utilization of xylose and galactose via adaptive laboratory evolution. *ACS Sustain. Chem. Eng.* 9, 11512–11523. doi:10.1021/acssuschemeng.1c03765

Lim, H. G., Fong, B., Alarcon, G., Magurudeniya, H. D., Eng, T., Szubin, R., et al. (2020). Generation of ionic liquid tolerant *Pseudomonas putida* KT2440 strains *via* adaptive laboratory evolution. *Green Chem.* 22, 5677–5690. doi:10.1039/D0GC01663B

Lindberg, P., Park, S., and Melis, A. (2010). Engineering a platform for photosynthetic isoprene production in cyanobacteria, using Synechocystis as the model organism. *Metab. Eng.* 12, 70–79. doi:10.1016/j.ymben.2009.10.001

Linger, J. G., Vardon, D. R., Guarnieri, M. T., Karp, E. M., Hunsinger, G. B., Franden, M. A., et al. (2014). Lignin valorization through integrated biological funneling and chemical catalysis. *Proc. Natl. Acad. Sci. U. S. A.* 111, 12013–12018. doi:10.1073/pnas.1410657111

Liu, N., Santala, S., and Stephanopoulos, G. (2020). Mixed carbon substrates: A necessary nuisance or a missed opportunity? *Curr. Opin. Biotechnol.* 62, 15–21. doi:10.1016/j.copbio.2019.07.003

Liu, S., Zhang, M., Ren, Y., Jin, G., Tao, Y., Lyu, L., et al. (2021a). Engineering Rhodosporidium toruloides for limonene production. *Biotechnol. Biofuels* 14, 243. doi:10.1186/s13068-021-02094-7

Liu, Z.-H., Olson, M. L., Shinde, S., Wang, X., Hao, N., Yoo, C. G., et al. (2017). Synergistic maximization of the carbohydrate output and lignin processability by combinatorial pretreatment. *Green Chem.* 19, 4939–4955. doi:10.1039/C7GC02057K

Liu, Z., Radi, M., Mohamed, E. T. T., Feist, A. M., Dragone, G., and Mussatto, S. I. (2021b). Adaptive laboratory evolution of Rhodosporidium toruloides to inhibitors derived from lignocellulosic biomass and genetic variations behind evolution. *Bioresour. Technol.* 333, 125171. doi:10.1016/j.biortech.2021.125171

Löwe, H., Hobmeier, K., Moos, M., Kremling, A., and Pflüger-Grau, K. (2017). Photoautotrophic production of polyhydroxyalkanoates in a synthetic mixed culture of Synechococcus elongatus cscB and Pseudomonas putida cscAB. *Biotechnol. Biofuels* 10, 190. doi:10.1186/s13068-017-0875-0

Lynch, M. D. (2021). The bioprocess TEA calculator: An online technoeconomic analysis tool to evaluate the commercial competitiveness of potential bioprocesses. *Metab. Eng.* 65, 42–51. doi:10.1016/j.ymben.2021.03.004

Ma, X., Liang, H., Panda, S., Fung, V. K. Y., Zhou, J. F. J., and Zhou, K. (2021). C2 feedstock-based biomanufacturing of value-added chemicals. *Curr. Opin. Biotechnol.* 73, 240–245. doi:10.1016/j.copbio.2021.08.017

Magurudeniya, H. D., Baral, N. R., Rodriguez, A., Scown, C. D., Dahlberg, J., Putnam, D., et al. (2021). Use of ensiled biomass sorghum increases ionic liquid pretreatment efficiency and reduces biofuel production cost and carbon footprint. *Green Chem.* 23, 3127–3140. doi:10.1039/D0GC03260C

Mahmud, R., Moni, S. M., High, K., and Carbajales-Dale, M. (2021). Integration of techno-economic analysis and life cycle assessment for sustainable process design – a review. *J. Clean. Prod.* 317, 128247. doi:10.1016/j.jclepro.2021.128247

Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D. C., Skea, J., Shukla, P. R., et al. (2018). *Global warming of 1.5 C: Summary for policy makers* (Incheon: IPCC-The Intergovernmental Panel on Climate Change).

McClelland, D. J., Wang, B.-X., Cordell, W. T., Cortes-Peña, Y. R., Gilcher, E. B., Zhang, L., et al. (2021). Renewable linear alpha-olefins by base-catalyzed dehydration of biologically-derived fatty alcohols. *Green Chem.* 23, 4338–4354. doi:10.1039/D1GC00243K

McKone, T. E., Nazaroff, W. W., Berck, P., Auffhammer, M., Lipman, T., Torn, M. S., et al. (2011). Grand challenges for life-cycle assessment of biofuels. *Environ. Sci. Technol.* 45, 1751–1756. doi:10.1021/es103579c

McManus, M. C., Taylor, C. M., Mohr, A., Whittaker, C., Scown, C. D., Borrion, A. L., et al. (2015). Challenge clusters facing LCA in environmental decision-making-what we can learn from biofuels. *Int. J. Life Cycle Assess.* 20, 1399–1414. doi:10.1007/s11367-015-0930-7

Melis, A. (1999). Photosystem-II damage and repair cycle in chloroplasts: What modulates the rate of photodamage *in vivo*? *Trends Plant Sci.* 4, 130–135. doi:10.1016/S1360-1385(99)01387-4

Mezzina, M. P., Manoli, M. T., Prieto, M. A., and Nikel, P. I. (2021). Engineering native and synthetic pathways in Pseudomonas putida for the production of tailored polyhydroxyalkanoates. *Biotechnol. J.* 16, e2000165. doi:10.1002/biot.202000165

Minty, J. J., Singer, M. E., Scholz, S. A., Bae, C.-H., Ahn, J.-H., Foster, C. E., et al. (2013). Design and characterization of synthetic fungal-bacterial consortia for direct production of isobutanol from cellulosic biomass. *Proc. Natl. Acad. Sci. U. S. A.* 110, 14592–14597. doi:10.1073/pnas.1218447110

Mohamed, E. T., Wang, S., Lennen, R. M., Herrgård, M. J., Simmons, B. A., Singer, S. W., et al. (2017). Generation of a platform strain for ionic liquid tolerance using adaptive laboratory evolution. *Microb. Cell Fact.* 16, 204. doi:10.1186/s12934-017-0819-1

Mohamed, E. T., Werner, A. Z., Salvachúa, D., Singer, C. A., Szostkiewicz, K., Rafael Jiménez-Díaz, M., et al. (2020). Adaptive laboratory evolution of Pseudomonas putida KT2440 improves p-coumaric and ferulic acid catabolism and tolerance. *Metab. Eng. Commun.* 11, e00143. doi:10.1016/j.mec.2020.e00143

Monroe, E., Gladden, J., Albrecht, K. O., Bays, J. T., McCormick, R., Davis, R. W., et al. (2019). Discovery of novel octane hyperboosting phenomenon in prenol biofuel/gasoline blends. *Fuel* 239, 1143–1148. doi:10.1016/j.fuel.2018.11.046

Morales, M., Dapsens, P. Y., Giovinazzo, I., Witte, J., Mondelli, C., Papadokonstantakis, S., et al. (2015). Environmental and economic assessment of lactic acid production from glycerol using cascade bio- and chemocatalysis. *Energy Environ. Sci.* 8, 558–567. doi:10.1039/C4EE03352C

Mosey, M., Douchi, D., Knoshaug, E. P., and Laurens, L. M. L. (2021). Methodological review of genetic engineering approaches for non-model algae. *Algal Res.* 54, 102221. doi:10.1016/j.algal.2021.102221

Naerdal, I., Pfeifenschneider, J., Brautaset, T., and Wendisch, V. F. (2015). Methanol-based cadaverine production by genetically engineered Bacillus methanolicus strains. *Microb. Biotechnol.* 8, 342–350. doi:10.1111/1751-7915.12257

Nakamura, C. E., and Whited, G. M. (2003). Metabolic engineering for the microbial production of 1, 3-propanediol. *Curr. Opin. Biotechnol.* 14, 454–459. doi:10.1016/j.copbio.2003.08.005

Narwani, A., Lashaway, A. R., Hietala, D. C., Savage, P. E., and Cardinale, B. J. (2016). Power of plankton: Effects of algal biodiversity on biocrude production

and stability. *Environ. Sci. Technol.* 50, 13142–13150. doi:10.1021/acs.est.6b03256

Neubauer, P., Häggström, L., and Enfors, S. O. (1995). Influence of substrate oscillations on acetate formation and growth yield in *Escherichia coli* glucose limited fed-batch cultivations. *Biotechnol. Bioeng.* 47, 139–146. doi:10.1002/bit.260470204

Nghiem, N., Kleff, S., and Schwegmann, S. (2017). Succinic acid: Technology development and commercialization. *Fermentation* 3, 26. doi:10.3390/fermentation3020026

Nguyen, D. T. N., Lee, O. K., Lim, C., Lee, J., Na, J.-G., and Lee, E. Y. (2020). Metabolic engineering of type II methanotroph, Methylosinus trichosporium OB3b, for production of 3-hydroxypropionic acid from methane via a malonyl-CoA reductase-dependent pathway. *Metab. Eng.* 59, 142–150. doi:10.1016/j.ymben.2020.02.002

Nguyen, D. T. N., Lee, O. K., Nguyen, T. T., and Lee, E. Y. (2021). Type II methanotrophs: A promising microbial cell-factory platform for bioconversion of methane to chemicals. *Biotechnol. Adv.* 47, 107700. doi:10.1016/j.biotechadv.2021.107700

Niu, W., Draths, K. M., and Frost, J. W. (2002). Benzene-free synthesis of adipic acid. *Biotechnol. Prog.* 18, 201–211. doi:10.1021/bp010179x

Oelgeschläger, E., and Rother, M. (2008). Carbon monoxide-dependent energy metabolism in anaerobic bacteria and archaea. *Arch. Microbiol.* 190, 257–269. doi:10.1007/s00203-008-0382-6

Ögmundarson, Ó., Herrgård, M. J., Forster, J., Hauschild, M. Z., and Fantke, P. (2020). Addressing environmental sustainability of biochemicals. *Nat. Sustain.* 3, 167–174. doi:10.1038/s41893-019-0442-8

Oliver, J. W. K., Machado, I. M. P., Yoneda, H., and Atsumi, S. (2013). Cyanobacterial conversion of carbon dioxide to 2, 3-butanediol. *Proc. Natl. Acad. Sci. U. S. A.* 110, 1249–1254. doi:10.1073/pnas.1213024110

Orita, I., Nishikawa, K., Nakamura, S., and Fukui, T. (2014). Biosynthesis of polyhydroxyalkanoate copolymers from methanol by Methylobacterium extorquens AM1 and the engineered strains under cobalt-deficient conditions. *Appl. Microbiol. Biotechnol.* 98, 3715–3725. doi:10.1007/s00253-013-5490-9

Park, J. O., Liu, N., Holinski, K. M., Emerson, D. F., Qiao, K., Woolston, B. M., et al. (2019). Synergistic substrate cofeeding stimulates reductive metabolism. *Nat. Metab.* 1, 643–651. doi:10.1038/s42255-019-0077-0

Park, M., Chen, Y., Thompson, M., Benites, V. T., Fong, B., Petzold, C. J., et al. (2020). Response of *Pseudomonas putida* to complex, aromatic-rich fractions from biomass. *ChemSusChem* 13, 4455–4467. doi:10.1002/cssc.202000268

Park, S. J., Kim, E. Y., Noh, W., Park, H. M., Oh, Y. H., Lee, S. H., et al. (2013). Metabolic engineering of *Escherichia coli* for the production of 5-aminovalerate and glutarate as C5 platform chemicals. *Metab. Eng.* 16, 42–47. doi:10.1016/j.ymben.2012.11.011

Patouillard, L., Collet, P., Lesage, P., Tirado Seco, P., Bulle, C., and Margni, M. (2019). Prioritizing regionalization efforts in life cycle assessment through global sensitivity analysis: A sector meta-analysis based on ecoinvent v3. *Int. J. Life Cycle Assess.* 24, 2238–2254. doi:10.1007/s11367-019-01635-5

Peña, D. A., Gasser, B., Zanghellini, J., Steiger, M. G., and Mattanovich, D. (2018). Metabolic engineering of Pichia pastoris. *Metab. Eng.* 50, 2–15. doi:10.1016/j.ymben.2018.04.017

Peralta-Yahya, P. P., Ouellet, M., Chan, R., Mukhopadhyay, A., Keasling, J. D., and Lee, T. S. (2011). Identification and microbial production of a terpene-based advanced biofuel. *Nat. Commun.* 2, 483. doi:10.1038/ncomms1494

Peters, G. P., Andrew, R. M., Canadell, J. G., Friedlingstein, P., Jackson, R. B., Korsbakken, J. I., et al. (2019). Carbon dioxide emissions continue to grow amidst slowly emerging climate policies. *Nat. Clim. Chang.* 10, 3–6. doi:10.1038/s41558-019-0659-6

Polle, J. E. W., Kanakagiri, S.-D., and Melis, A. (2003). tla1, a DNA insertional transformant of the green alga Chlamydomonas reinhardtii with a truncated light-harvesting chlorophyll antenna size. *Planta* 217, 49–59. doi:10.1007/s00425-002-0968-1

P. T. Anastas and T. C. Williamson (Editors) (1996). *Green chemistry: Designing chemistry for the environment* (Washington, DC: American Chemical Society). doi:10.1021/bk-1996-0626

Qian, Z.-G., Xia, X.-X., and Lee, S. Y. (2011). Metabolic engineering of *Escherichia coli* for the production of cadaverine: A five carbon diamine. *Biotechnol. Bioeng.* 108, 93–103. doi:10.1002/bit.22918

Quinn, J. C., and Davis, R. (2015). The potentials and challenges of algae based biofuels: A review of the techno-economic, life cycle, and resource assessment modeling. *Bioresour. Technol.* 184, 444–452. doi:10.1016/j.biortech.2014.10.075

Renouf, M. A., Wegener, M. K., and Nielsen, L. K. (2008). An environmental life cycle assessment comparing Australian sugarcane with US corn and UK sugar beet as producers of sugars for fermentation. *Biomass Bioenergy* 32, 1144–1155. doi:10.1016/j.biombioe.2008.02.012

Rios, J., Lebeau, J., Yang, T., Li, S., and Lynch, M. D. (2021). A critical review on the progress and challenges to a more sustainable, cost competitive synthesis of adipic acid. *Green Chem.* 23, 3172–3190. doi:10.1039/D1GC00638J

Robinson, C. J., Carbonell, P., Jervis, A. J., Yan, C., Hollywood, K. A., Dunstan, M. S., et al. (2020). Rapid prototyping of microbial production strains for the biomanufacture of potential materials monomers. *Metab. Eng.* 60, 168–182. doi:10.1016/j.ymben.2020.04.008

Roell, G. W., Zha, J., Carr, R. R., Koffas, M. A., Fong, S. S., and Tang, Y. J. (2019). Engineering microbial consortia by division of labor. *Microb. Cell Fact.* 18, 35. doi:10.1186/s12934-019-1083-3

Rohles, C. M., Gläser, L., Kohlstedt, M., Gießelmann, G., Pearson, S., del Campo, A., et al. (2018). A bio-based route to the carbon-5 chemical glutaric acid and to bionylon-6, 5 using metabolically engineered*Corynebacterium glutamicum*. *Green Chem.* 20, 4662–4674. doi:10.1039/C8GC01901K

Rolf, J., Julsing, M. K., Rosenthal, K., and Lütz, S. (2020). A gram-scale limonene production process with engineered *Escherichia coli*. *Molecules* 25, 1881. doi:10.3390/molecules25081881

Saleski, T. E., Chung, M. T., Carruthers, D. N., Khasbaatar, A., Kurabayashi, K., and Lin, X. N. (2021). Optimized gene expression from bacterial chromosome by high-throughput integration and screening. *Sci. Adv.* 7, eabe1767. doi:10.1126/sciadv.abe1767

Sánchez-Andrea, I., Guedes, I. A., Hornung, B., Boeren, S., Lawson, C. E., Sousa, D. Z., et al. (2020). The reductive glycine pathway allows autotrophic growth of Desulfovibrio desulfuricans. *Nat. Commun.* 11, 5090. doi:10.1038/s41467-020-18906-7

Sander, J. D., and Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32, 347–355. doi:10.1038/nbt.2842

Santos-Merino, M., Singh, A. K., and Ducat, D. C. (2019). New applications of synthetic biology tools for cyanobacterial metabolic engineering. *Front. Bioeng. Biotechnol.* 7, 33. doi:10.3389/fbioe.2019.00033

Santos-Merino, M., Torrado, A., Davis, G. A., Röttig, A., Bibby, T. S., Kramer, D. M., et al. (2021). Improved photosynthetic capacity and photosystem I oxidation via heterologous metabolism engineering in cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2021523118. doi:10.1073/pnas.2021523118

Schuchmann, K., and Müller, V. (2014). Autotrophy at the thermodynamic limit of life: A model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* 12, 809–821. doi:10.1038/nrmicro3365

Scown, C. D., Baral, N. R., Yang, M., Vora, N., and Huntington, T. (2021). Technoeconomic analysis for biofuels and bioproducts. *Curr. Opin. Biotechnol.* 67, 58–64. doi:10.1016/j.copbio.2021.01.002

Seager, T. P., Trump, B. D., Poinsatte-Jones, K., and Linkov, I. (2017). Why life cycle assessment does not work for synthetic biology. *Environ. Sci. Technol.* 51, 5861–5862. doi:10.1021/acs.est.7b01604

Searchinger, T., Heimlich, R., Houghton, R. A., Dong, F., Elobeid, A., Fabiosa, J., et al. (2008). Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science* 319, 1238–1240. doi:10.1126/science.1151861

Shi, R., and Guest, J. S. (2020). BioSTEAM-LCA: An integrated modeling framework for agile life cycle assessment of biorefineries under uncertainty. *ACS Sustain. Chem. Eng.* 8, 18903–18914. doi:10.1021/acssuschemeng.0c05998

Sick, V., Armstrong, K., Cooney, G., Cremonese, L., Eggleston, A., Faber, G., et al. (2020). The need for and path to harmonized life cycle assessment and techno-economic assessment for carbon dioxide capture and utilization. *Energy Technol.* 8, 1901034. doi:10.1002/ente.201901034

Singhvi, M. S., and Gokhale, D. V. (2019). Lignocellulosic biomass: Hurdles and challenges in its valorization. *Appl. Microbiol. Biotechnol.* 103, 9305–9320. doi:10.1007/s00253-019-10212-7

Skoog, E., Shin, J. H., Saez-Jimenez, V., Mapelli, V., and Olsson, L. (2018). Biobased adipic acid - the challenge of developing the production host. *Biotechnol. Adv.* 36, 2248–2263. doi:10.1016/j.biotechadv.2018.10.012

Song, R., Keller, A. A., and Suh, S. (2017). Rapid life-cycle impact screening using artificial neural networks. *Environ. Sci. Technol.* 51, 10777–10785. doi:10.1021/acs.est.7b02862

Sonoki, T., Takahashi, K., Sugita, H., Hatamura, M., Azuma, Y., Sato, T., et al. (2018). Glucose-free*cis* , *cis* -muconic acid production via new metabolic designs corresponding to the heterogeneity of lignin. *ACS Sustain. Chem. Eng.* 6, 1256–1264. doi:10.1021/acssuschemeng.7b03597

Souza, D. M., Teixeira, R. F. M., and Ostermann, O. P. (2015). Assessing biodiversity loss due to land use with life cycle assessment: Are we there yet? *Glob. Chang. Biol.* 21, 32–47. doi:10.1111/gcb.12709

Spierling, S., Knüpffer, E., Behnsen, H., Mudersbach, M., Krieg, H., Springer, S., et al. (2018). Bio-based plastics - a review of environmental, social and economic impact assessments. *J. Clean. Prod.* 185, 476–491. doi:10.1016/j.jclepro.2018.03.014

Standardization International Organization for (2006). Environmental management: Life cycle assessment; principles and framework. Available at: https://www.iso.org/standard/37456.html.

Sternberg, A., Jens, C. M., and Bardow, A. (2017). Life cycle assessment of CO_2 -based C1-chemicals. *Green Chem.* 19, 2244–2259. doi:10.1039/C6GC02852G

Subramaniam, B., Licence, P., Moores, A., and Allen, D. T. (2021). Shaping effective practices for incorporating sustainability assessment in manuscripts submitted to *acs sustainable chemistry and engineering*: An initiative by the editors. *ACS Sustain. Chem. Eng.* 9, 3977–3978. doi:10.1021/acssuschemeng.1c01511

Suh, S., and Huppes, G. (2005). Methods for life cycle inventory of a product. *J. Clean. Prod.* 13, 687–697. doi:10.1016/j.jclepro.2003.04.001

Sun, C., Theodoropoulos, C., and Scrutton, N. S. (2020). Techno-economic assessment of microbial limonene production. *Bioresour. Technol.* 300, 122666. doi:10.1016/j.biortech.2019.122666

Sun, T., Yu, Y., Wang, K., Ledesma-Amaro, R., and Ji, X.-J. (2021). Engineering Yarrowia lipolytica to produce fuels and chemicals from xylose: A review. *Bioresour. Technol.* 337, 125484. doi:10.1016/j.biortech.2021.125484

Sundstrom, E., Yaegashi, J., Yan, J., Masson, F., Papa, G., Rodriguez, A., et al. (2018). Demonstrating a separation-free process coupling ionic liquid pretreatment, saccharification, and fermentation with *Rhodosporidium toruloides* to produce advanced biofuels. *Green Chem.* 20, 2870–2879. doi:10.1039/C8GC00518D

Tan, D., Wang, Y., Tong, Y., and Chen, G.-Q. (2021). Grand challenges for industrializing polyhydroxyalkanoates (PHAs). *Trends Biotechnol.* 39, 953–963. doi:10.1016/j.tibtech.2020.11.010

Teixeira, L. V., Moutinho, L. F., and Romão-Dumaresq, A. S. (2018). Gas fermentation of C1 feedstocks: Commercialization status and future prospects. *Biofuels, Bioprod. Bioref.* 12, 1103–1117. doi:10.1002/bbb.1912

Thomassen, G., Van Dael, M., Van Passel, S., and You, F. (2019). How to assess the potential of emerging green technologies? Towards a prospective environmental and techno-economic assessment framework. *Green Chem.* 21, 4868–4886. doi:10.1039/C9GC02223F

Torres-Tiji, Y., Fields, F. J., and Mayfield, S. P. (2020). Microalgae as a future food source. *Biotechnol. Adv.* 41, 107536. doi:10.1016/j.biotechadv.2020.107536

Tran, N. T., and Kaldenhoff, R. (2020). Achievements and challenges of genetic engineering of the model green alga Chlamydomonas reinhardtii. *Algal Res.* 50, 101986. doi:10.1016/j.algal.2020.101986

Trump, B. D., Cegan, J., Wells, E., Poinsatte-Jones, K., Rycroft, T., Warner, C., et al. (2019). Co-evolution of physical and social sciences in synthetic biology. *Crit. Rev. Biotechnol.* 39, 351–365. doi:10.1080/07388551.2019.1566203

Tsiropoulos, I., Cok, B., and Patel, M. K. (2013). Energy and greenhouse gas assessment of European glucose production from corn – A multiple allocation approach for a key ingredient of the bio-based economy. *J. Clean. Prod.* 43, 182–190. doi:10.1016/j.jclepro.2012.12.035

Tu, Q., Eckelman, M., and Zimmerman, J. (2017). Meta-analysis and harmonization of life cycle assessment studies for algae biofuels. *Environ. Sci. Technol.* 51, 9419–9432. doi:10.1021/acs.est.7b01049

Tuyishime, P., Wang, Y., Fan, L., Zhang, Q., Li, Q., Zheng, P., et al. (2018). Engineering Corynebacterium glutamicum for methanol-dependent growth and glutamate production. *Metab. Eng.* 49, 220–231. doi:10.1016/j.ymben.2018.07.011

Van Dael, M., Kuppens, T., Lizin, S., and Van Passel, S. (2015). "Techno-economic assessment methodology for ultrasonic production of biofuels," in *Production of biofuels and chemicals with ultrasound biofuels and biorefineries*. Editors Z. Fang, R. L. Smith, and X. Qi (Dordrecht: Springer Netherlands), 317–345. doi:10.1007/978-94-017-9624-8_12

Vardon, D. R., Franden, M. A., Johnson, C. W., Karp, E. M., Guarnieri, M. T., Linger, J. G., et al. (2015). Adipic acid production from lignin. *Energy Environ. Sci.* 8, 617–628. doi:10.1039/C4EE03230F

Villa, J. A., Ray, E. E., and Barney, B. M. (2014). *Azotobacter vinelandii* siderophore can provide nitrogen to support the culture of the green algae *Neochloris oleoabundans* and *Scenedesmus* sp. BA032. *FEMS Microbiol. Lett.* 351, 70–77. doi:10.1111/1574-6968.12347

Vink, E. T. H., Rábago, K. R., Glassner, D. A., and Gruber, P. R. (2003). Applications of life cycle assessment to NatureWorks™ polylactide (PLA) production. *Polym. Degrad. Stab.* 80, 403–419. doi:10.1016/S0141-3910(02)00372-5

Viswanathan, M. B., Raman, D. R., Rosentrater, K. A., and Shanks, B. H. (2020). A technoeconomic platform for early-stage process design and cost estimation of Joint fermentative–catalytic bioprocessing. *Processes* 8, 229. doi:10.3390/pr8020229

Voigt, C. A. (2020). Synthetic biology 2020-2030: Six commercially-available products that are changing our world. *Nat. Commun.* 11, 6379. doi:10.1038/s41467-020-20122-2

Walker, S., and Rothman, R. (2020). Life cycle assessment of bio-based and fossil-based plastic: A review. *J. Clean. Prod.* 261, 121158. doi:10.1016/j.jclepro.2020.121158

Wehrs, M., Tanjore, D., Eng, T., Lievense, J., Pray, T. R., and Mukhopadhyay, A. (2019). Engineering robust production microbes for large-scale cultivation. *Trends Microbiol.* 27 (6), 524–537.

Weiss, T. L., Young, E. J., and Ducat, D. C. (2017). A synthetic, light-driven consortium of cyanobacteria and heterotrophic bacteria enables stable polyhydroxybutyrate production. *Metab. Eng.* 44, 236–245. doi:10.1016/j.ymben.2017.10.009

Wernet, G., Bauer, C., Steubing, B., Reinhard, J., Moreno-Ruiz, E., and Weidema, B. (2016). The ecoinvent database version 3 (part I): Overview and methodology. *Int. J. Life Cycle Assess.* 21, 1218–1230. doi:10.1007/s11367-016-1087-8

Werpy, T., and Petersen, G. (2004). *Top value added chemicals from biomass: Volume I – results of screening for potential candidates from sugars and synthesis gas*. US Department of Energy. Golden, CO: National Renewable Energy Lab. doi:10.2172/15008859

Wu, M.-Y., Sung, L.-Y., Li, H., Huang, C.-H., and Hu, Y.-C. (2017). Combining CRISPR and CRISPRi systems for metabolic engineering of *E. coli* and 1, 4-BDO biosynthesis. *ACS Synth. Biol.* 6, 2350–2361. doi:10.1021/acssynbio.7b00251

Wu, W., Liu, F., and Singh, S. (2018). Toward engineering *E. coli* with an autoregulatory system for lignin valorization. *Proc. Natl. Acad. Sci. U. S. A.* 115, 2970–2975. doi:10.1073/pnas.1720129115

Wu, W., and Maravelias, C. T. (2018). Synthesis and techno-economic assessment of microbial-based processes for terpenes production. *Biotechnol. Biofuels* 11, 294. doi:10.1186/s13068-018-1285-7

Wurtzel, E. T., Vickers, C. E., Hanson, A. D., Millar, A. H., Cooper, M., Voss-Fels, K. P., et al. (2019). Revolutionizing agriculture with synthetic biology. *Nat. Plants* 5, 1207–1210. doi:10.1038/s41477-019-0539-0

Yang, Y.-M., Chen, W.-J., Yang, J., Zhou, Y.-M., Hu, B., Zhang, M., et al. (2017). Production of 3-hydroxypropionic acid in engineered Methylobacterium extorquens AM1 and its reassimilation through a reductive route. *Microb. Cell Fact.* 16, 179. doi:10.1186/s12934-017-0798-2

Yishai, O., Bouzon, M., Döring, V., and Bar-Even, A. (2018). *In vivo* assimilation of one-carbon via a synthetic reductive Glycine pathway in *Escherichia coli*. *ACS Synth. Biol.* 7, 2023–2028. doi:10.1021/acssynbio.8b00131

Yishai, O., Lindner, S. N., Gonzalez de la Cruz, J., Tenenboim, H., and Bar-Even, A. (2016). The formate bio-economy. *Curr. Opin. Chem. Biol.* 35, 1–9. doi:10.1016/j.cbpa.2016.07.005

Zang, G., Sun, P., Elgowainy, A., and Wang, M. (2021). Technoeconomic and life cycle analysis of synthetic methanol production from hydrogen and industrial byproduct CO2. *Environ. Sci. Technol.* 55, 5248–5257. doi:10.1021/acs.est.0c08237

Zhao, M., Huang, D., Zhang, X., Koffas, M. A. G., Zhou, J., and Deng, Y. (2018). Metabolic engineering of *Escherichia coli* for producing adipic acid through the reverse adipate-degradation pathway. *Metab. Eng.* 47, 254–262. doi:10.1016/j.ymben.2018.04.002

Zheng, J., and Suh, S. (2019). Strategies to reduce the global carbon footprint of plastics. *Nat. Clim. Chang.* 9, 374–378. doi:10.1038/s41558-019-0459-z

Zheng, Y., Chen, J.-C., Ma, Y.-M., and Chen, G.-Q. (2020). Engineering biosynthesis of polyhydroxyalkanoates (PHA) for diversity and cost reduction. *Metab. Eng.* 58, 82–93. doi:10.1016/j.ymben.2019.07.004

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Review
PUBLISHED 25 August 2022
DOI 10.3389/fbioe.2022.890155

OPEN ACCESS

EDITED BY
Jean Marie François,
Institut Biotechnologique de Toulouse (INSA), France

REVIEWED BY
Mario Andrea Marchisio,
Tianjin University, China
Mingfeng Cao,
Xiamen University, China

*CORRESPONDENCE
Jingyu Chen,
chenjy@cau.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 05 March 2022
ACCEPTED 28 July 2022
PUBLISHED 25 August 2022

CITATION
Li W, Huang C and Chen J (2022), The application of CRISPR /Cas mediated gene editing in synthetic biology: Challenges and optimizations. *Front. Bioeng. Biotechnol.* 10:890155. doi: 10.3389/fbioe.2022.890155

The application of CRISPR /Cas mediated gene editing in synthetic biology: Challenges and optimizations

Wenqian Li[1,2†], Can Huang[1,2†] and Jingyu Chen[1,2*]

[1]MOE Key Laboratory of Precision Nutrition and Food Quality, College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, China, [2]Key Laboratory of Food Bioengineering (China National Light Industry), College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, China

Clustered regularly interspaced short palindromic repeats (CRISPR) and its associated enzymes (Cas) is a simple and convenient genome editing tool that has been used in various cell factories and emerging synthetic biology in the recent past. However, several problems, including off-target effects, cytotoxicity, and low efficiency of multi-gene editing, are associated with the CRISPR/Cas system, which have limited its application in new species. In this review, we briefly describe the mechanisms of CRISPR/Cas engineering and propose strategies to optimize the system based on its defects, including, but not limited to, enhancing targeted specificity, reducing toxicity related to Cas protein, and improving multi-point editing efficiency. In addition, some examples of improvements in synthetic biology are also highlighted. Finally, future perspectives of system optimization are discussed, providing a reference for developing safe genome-editing tools for new species.

KEYWORDS

CRISPR/Cas, off-target effects, toxicity, gene-editing efficiency, synthetic biology

Introduction

The features of clustered regularly interspaced short palindromic repeats (CRISPR) were discovered serendipitously in the genomes of various bacteria and archaea during molecular biology studies (Ishino et al., 1987; Mojica et al., 1993). The system was originally used as an adaptive immune defense system by bacteria against the invasion of foreign nucleic acids, such as viruses and plasmids (Ishino et al., 1987; Sovova et al., 2017; Albitar et al., 2018). The immune process is divided into three stages: 1) adaptation: the invading nucleotide fragments are first captured by the host organism, and subsequently, a Cas integrase-derived nucleic acid sequence is inserted into the CRISPR array; 2) biogenesis: the CRISPR array is transcribed into a pre-crRNA (pre-crRNA) containing a spacer and a portion of the repeat, which is subsequently cleaved in the repeat to induce the generation of mature guide crRNA (gRNA); and 3) interference: mature crRNA-Cas complexes recognize specific sites on target nucleic acids through complementary base pairing, triggering Cas enzymes to catalyze the cleavage of effector complexes of invading

nucleotides (Debin Zhang et al., 2020; Makarova et al., 2020). CRISPR provides acquired immunity to prokaryotes, in addition to other unrelated functions, such as gene regulation (Mojica and Rodriguez-Valera, 2016). The discovery and application of the CRISPR-Cas enzyme system in prokaryotes and eukaryotes have revolutionized genome engineering by providing directly accessible or editable tools (Barrangou, 2014).

CRISPR/Cas systems usually consist of Cas proteins and CRISPR arrays, in which the Cas proteins are involved in the acquisition and protection of invading nucleotides. Currently, CRISPR/Cas systems are divided into classes 1 and 2 according to the composition of their effectors. Class 1 systems include many Cas protein subunits, including types I, III, and IV (Makarova et al., 2011). In contrast, only one Cas protein involving II, V, and VI subdomains is included in Class 2 (Ishino et al., 2018). Class 2 CRISPR/Cas systems have emerged as an attractive option for developing next-generation genome-editing technologies because of their simple structural design and easy implementation in gene editing and manipulation of cell-free nucleic acids (Tang and Fu, 2018; Debin Zhang et al., 2020). Therefore, we consider a typical Class 2 system as an example to introduce the basic components of the system.

Generally, the CRISPR/Cas9 system is considered the smallest CRISPR/Cas system because its pre-CrRNA processing occurs solely with Cas9, a member of the Cas family. Apart from Cas9, single-guide RNA (sgRNA) is the other main component of this simple and easy Class 2 system. sgRNA structure comprises two parts, one of which plays a role in guiding cleavage by binding to Cas9 (Mojica et al., 2005; Pourcel et al., 2005), whereas the other is capable of binding to the target sequence through complementary base pairing to guide positioning. A conserved sequence called the proto-spacer adjacent motif (PAM) exists downstream of the target sequence (Jinek et al., 2012). The presence or absence of the PAM sequence determines whether Cas9 protein cleaves the target sequence. In this system, Cas9 cleaves the target DNA sequence under the guidance of gRNA to create double-strand breaks (DSB). CRISPR systems in prokaryotes often require the addition of donor DNA fragments because they only have homology-directed repair (HDR) (Liang et al., 1998). In addition, the deletion or insertion of target DNA sequences can also be achieved by the addition of donor DNA fragments. Eukaryotes have access to another mode of non-homologous end joining (NHEJ) repair, allowing bacteria to repair without DSBs in the presence of donor DNA (Lieber, 2010). Another vital element, the Cas9 protein, is regarded as a nuclease with two cleavage domains: RuvC and HNH (Deltcheva et al., 2011). While both domains are activated, the HNH nuclease domain cleaves the DNA strand complementary to the sgRNA, and RuvC cleaves the DNA strand that is not complementary to the sgRNA, forming a nicked double-strand break near the PAM. Based on these properties, Cas9 can be directed to the target gene to generate DSBs by designing the sgRNA spacer sequence. In some studies, the two nuclease domains of RuvC and HNH were mutated to introduce novel Cas9 or completely inactivate Cas9, which has excellent effects in improving target efficiency (Makarova et al., 2011).

In recent years, the emergence of CRISPR/Cas and its derivative editing technology has become a pillar of experimental biology, as well as a tool for synthetic biology, because of convenient implementation, high mutation efficiency, and great potential for therapeutic applications (Zeballos and Gaj, 2021). However, the system itself has the risk of being off-target, with low multiple editing efficiency. The toxicity of Cas9 affects the broad application of this gene-editing technology in cell factories and synthetic biology (Figure 1). Ensuring the high editing efficiency of the CRISPR/Cas system and reducing off-target effects and toxicity to cells is urgently required to study CRISPR/Cas in new species. In this study, the mechanism of the CRISPR/Cas system and its composition are briefly reviewed. This review focuses on three current problems that have limited the development of this system and summarizes optimization strategies, as well as the latest practical application cases, to provide a reference for more researchers in system optimization and new species development.

Minimizing off-target editing of CRISPR/Cas system

Causes of off-target effects

In the CRISPR/Cas system, sgRNA can guide the Cas protein to be cut or modified by recognizing the PAM sequence and binding to the same sequence of the genome. Nucleotide-binding does not require an identical sequence, and the mismatching of a single base pair is acceptable. However, this leads to mismatching between the designed sgRNA and non-target DNA sequence, resulting in unexpected genetic mutation, known as the off-target effect (Haeussler, 2020).

This effect of random cleavage on the genome may considerably limit the development of CRISPR/Cas systems for practical applications. In early studies, Pattanayak et al. found that the specificity of the sgRNA-Cas9 complex depended on the complementary pairing of the 7–12 base sequence of sgRNA adjacent to PAM with the target gene, with distal sequences facilitating mutations, causing unexpected off-target events (Pattanayak et al., 2013). However, many groups using high-throughput sequencing have observed thousands of off-target effects on chromosomes. In some cases, even nucleotide insertions or deletions exist, increasing the off-target range. To address this, the development of genetic sequencing technology and the search for off-target locations *in vitro* using biochemical methods have allowed researchers to determine that most of the off-target sequences are similar to the target sequences, which could be

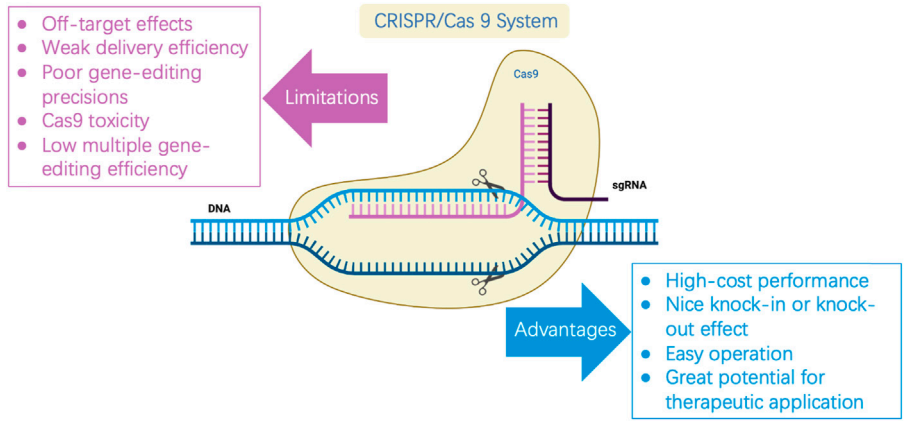


FIGURE 1
Limitations and advantages of CRISPR/Cas9.

used to forecast possible off-target effects. In prokaryotes, the genome is small with few similar sgRNA sequences, ensuring that the off-target effect has a weak impact. In addition, most prokaryotes only have a homology-directed repair (HDR) repair mechanism, such that DNA double-strand breaks (DSBs) caused by Cas9 are repaired with the assistance of donor DNA. Otherwise, the cells die, further reducing the possibility of off-target effects. Notably, this may also cause mutations in the PAM sequence or its adjacent guide sequence to escape the cleavage by Cas9.

Strategies to improve on-target effects

In most cases, off-target mutations can be improved by selecting sgRNA with higher specificities. A positive correlation exists between the GC content of sgRNA "seed region" and gene editing efficiency. Moreover, effective cleavage cannot occur with more than three mismatches in sgRNA beyond the "seed region" (Ren et al., 2014). Therefore, when designing sgRNA sequences, sgRNAs with a high GC content and low homology with genomic DNA other than target gene sequences can be selected to improve specificity. The length of sgRNA is also closely related to the targeting efficiency. sgRNA with less than 20 nucleotide sequences can effectively reduce off-target effects without affecting gene editing results. However, the mechanism underlying this strategy is currently unclear (Fu et al., 2014). Various software tools for designing high-precision sgRNAs have been widely developed, mainly focusing on identifying attractive off-target regions in the genome (Martin et al., 2016), allowing for a certain number of mismatches (Haeussler and Concordet, 2016). Many online networking tools, such as CRISPR-2.0 E-Crisp breaking-CAS system (Heigwer et al., 2014; Oliveros et al., 2016; Liu et al., 2017), have been applied to detect off-target mutations, including but not limited to providing services for sgRNA sequence design and reducing the possibility of off-target editing.

In recent years, the modification of wild-type Cas9 has been a crucial strategy to improve CRISPR/Cas9 specificity. This includes the inactivation of a single enzyme digestion domain in Cas9 to obtain NickCas9, which can only cut a single strand, thus requiring the simultaneous participation of two sgRNAs to induce DNA fracture. Therefore, this reduces the off-target rate since the simultaneous misalignment of both sgRNAs can lead to missing the target. Ran et al. utilized this strategy and found that cell lines can significantly reduce the off-target effect by 50–1,500 times and have broad applicability in different cell types (Ran et al., 2013b). Additionally, a series of high-fidelity Cas9 variants, such as xCas9 and SpCas9, can extend the PAM recognition range and DNA specificity, making it another revolutionary tool in gene editing (Zhaohui et al., 2020). Four evoCas9 mutants with improved editing efficiency have also been reported in yeast, showing good targeting and shearing activity. The fidelity of the SpCas9 mutant was found to be 79 times higher than that of the wild-type (Casini et al., 2018). Furthermore, a series of high-fidelity Cas9 variants have also been developed in recent years (Table 1), including Cas9n (Ran et al., 2013a), dCas9 (Qi et al., 2013), spCas9 (Kleinstiver et al., 2016), espCas9 (Cong et al., 2013), Cpf1 (Crosetto et al., 2013), HypaCas9 (Chen et al., 2017), and Sniper-Cas (Lee et al., 2019) (Table 1).

In addition to optimizing sgRNA and Cas9, reducing the concentration and exposure time of Cas9 are critical strategies to improve specificity. The discovery of anti-CRISPR (ACR) proteins has provided technical possibilities for realizing this strategy. The *ACR* gene expresses the anti-CRISPR-associated protein (ACA), which binds to upstream promoters and regulates *ACR* gene expression during transcription. ACR

TABLE 1 Improved CRISPR/Cas system involving Cas protein, gRNA and donor DNA.

Strategies	Descriptions	Characteristics	Ref
Cas proteins optimization strategy			
xCas9 or Cas9-NG	Engineered versions of *Streptococcus pyogenes* Cas9	Improve target specificity and expand target range	Hu et al. (2018)
Cas9n	Inactivating the HNH or RuvC nuclease domain of Cas9	Edit specific sites	Ran et al. (2013a)
dCas9	Inactivating both HNH and RuvC nuclease domain of Cas9	Base editing without generation of DSBs	Qi et al. (2013)
spCas9-HF1	Mutation of the key amino acid residues of SpCas9 responsible for contact with the target sequence	Variants that reduced nonspecific DNA interactions	Kleinstiver et al. (2016)
espCas9	*Streptococcus pyogenes* K848A, K1003A, R1060A	Requirement of a high specificity	Cong et al. (2013)
Cpf1	Type Ⅱ-Ⅴ CRISPR system	Cpf1 recognizes T-rich PAM, and is degraded by the endogenous protease system after editing	Crosetto et al. (2013)
HypaCas9	Balance nuclease activation and target recognition	Higher genome-wide fidelity	Chen et al. (2017)
Sniper-Cas	An *Escherichia coli* based directed evolution method, Sniper-screen to obtain a Cas9 variant	Higher specificity	Lee et al. (2019)
sgRNA optimization strategy			
Individual promoters/ terminators	Each sgRNA has independent promoter and terminator control	High efficiency and wide application, but unstable because of large structure and repeat sequences	Yu et al. (2020)
Type Ⅱ CRISPR crRNA array	Multiple sgRNA expression using crRNA array via one promoter and one terminator	The structure is simple, but the complementary pairing of pre-crRNA and trans-activating crRNA and the intracellular RNase III nuclease need to be considered	Li et al. (2020)
ABEs or CBEs	Transcriptome-wide gRNA-independent editing of RNA bases	Low DNA off-target and indels formation activity	Grunewald et al. (2019)
tRNA processing	Multiple sgRNA are controlled by one promoter and one terminator based on endogenous ribozymes, tRNA processing, and exogenous Csy4 protein	No need to introduce heterologous Cas protein, and showed more stable in multi-site editing	Zhang et al. (2019)
Prime editing (PE)	No DSB or donor DNA is required	89% of known genetic variations associated with human disease can be corrected	Tong et al. (2021)
DONOR DNA optimization strategy			
ssDNA	Single-stranded DNA	Simple to prepare, but limited in length	Lin et al. (2015)
dsDNA	Double-stranded DNA	More stable than ssDNA	Lin et al. (2015)
plasmids	Donor DNA templates are provided as plasmids	The plasmid transformation efficiency is high and stable, but the operation is complicated, and the number of plasmids available in some engineered strains is limited	Zhao et al. (2016)

protein has been shown to inhibit CRISPR/Cas immune function at different stages and prevent Cas9 from binding to DSBs, as well as cleavage, CrRNA loading, or effector complex formation, thus avoiding the continuous expression of Cas9 protein in cells, ultimately reducing miss efficiency (Stanley and Maxwell, 2018). Li et al. demonstrated that the *in vitro* injection of *ACR* genes reduced Cas9-related cytotoxicity and improved transplantation outcomes (Chang Li et al., 2018). Moreover, multiple protein families that naturally inhibited Cas9 have been identified. For example, Pawluk et al. identified three naturally occurring protein families that bind directly to *Neisseria meningitidis* Cas9 (NmeCas9) and act as effective inhibitors of this system in human cells (Pawluk et al., 2016).

Notably, these strategies are mainly based on eukaryotic studies. In contrast, owing to the low off-target rate in prokaryotes, the corresponding optimization strategies have not yet been fully developed in these organisms. However, using these strategies remains an excellent choice as a reference for improving the on-target effects in prokaryotes.

Minimizing the toxicity of CRISPR/Cas9

The versatility of the CRISPR system has given rise to derivative techniques based on nuclear editing, which may contribute to cytotoxicity owing to the unique nature of prokaryotic gene profiling, especially in microorganisms. CRISPR/Cas can lead to fatal chromosomal breaks, resulting in inefficient transformation and gene editing failure (Jing Zhao

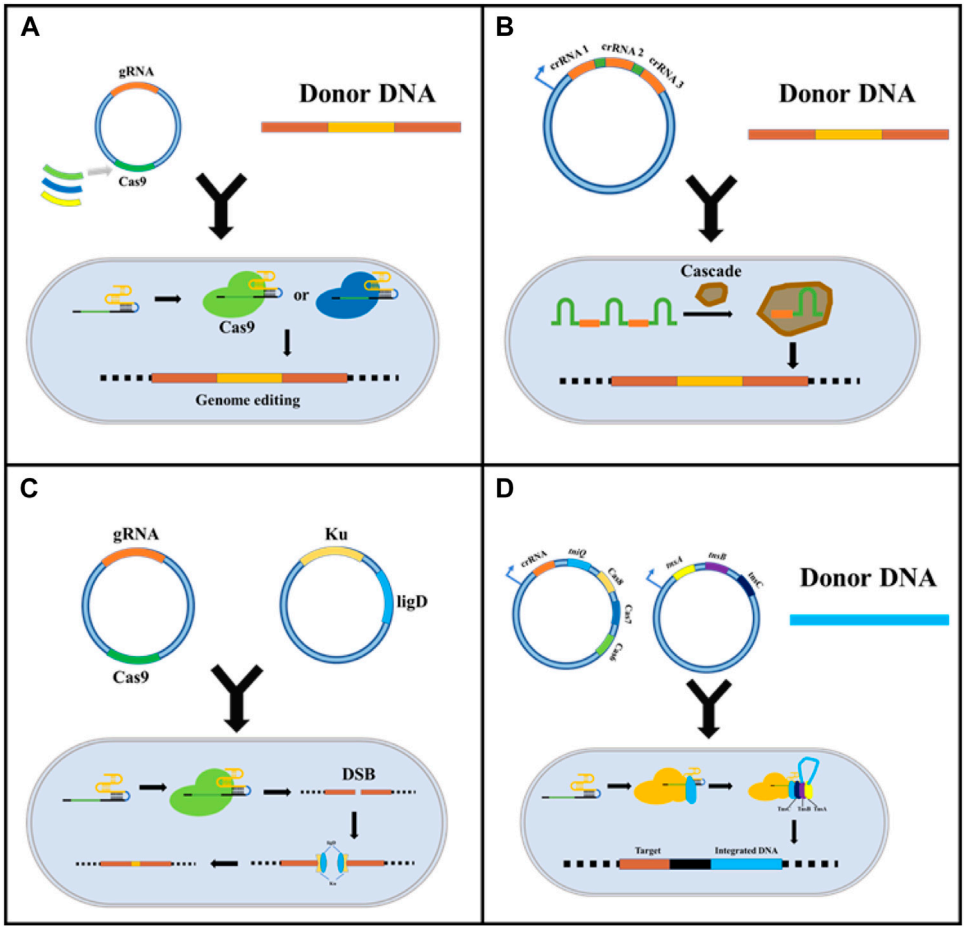


FIGURE 2
Strategies for reducing CRISPR/Cas9 system cytotoxicity. **(A)** Replacing Cas9 with other less toxic Cas proteins, such as dCas9 and Cpf1; **(B)** Using the intracellular endogenous CRISPR/Cas immune system to express crRNA and realized target sites-editing; **(C)** Introducing exogenous NHEJ recombination system to reduce DSB induced cytotoxicity; **(D)** Emerging CRISPR gene editing based on transposons.

et al., 2020). In this review, four methods to mitigate Cas9 toxicity are discussed (Figure 2): substituting for Cas proteins with lower toxicity (McNeely et al., 2010); using endogenous immune systems in prokaryotes to relieve the toxicity, and low transformation efficiency of Cas9 (Luo et al., 2015; Wang et al., 2016; Qin et al., 2021); reducing the cytotoxicity caused by DSBs by introducing the non-homologous end joining (NHEJ) repair mechanism (Huang et al., 2019; Su et al., 2019); and emerging novel CRISPR gene editing based on transposons (Strecker et al., 2019).

Although the CRISPR-Cas9 system from *Streptococcus pyogenes* is widely used in bacterial genome editing, its toxicity has limited research on *Streptomycetes* (Yeo et al., 2019) and *Corynebacterium glutamicum* (Jiang et al., 2017). Thus, establishing CRISPR systems in industrial microorganisms by substituting Cas proteins with an alternative with reduced toxicity is essential. *Cyanobacterium* has been successfully designed and transformed into a strain capable of producing biomass fuels and bioactive substances, such as alkanes and terpenoids (McNeely et al., 2010). However, Cas9 cytotoxicity has hindered the application of the CRISPR/Cas system in cyanobacteria. Unlike Cas9, Cas12a (Cpf1), which is modulated by the lac promoter, can thrive both in crRNA processing and target gene cleavage, resulting in sticky ends (Pattharaprachayakul et al., 2020). Cas12a provides an effective solution for cyanobacteria. However, the reasons for the lower toxicity of Cas12a protein are unknown (McNeely et al., 2010). The same strategy has been applied to *Streptomyces*, in which the FnCas12a protein from *Francisella novicida* was found to be effective (Lei Li et al., 2018). Meanwhile, the author found that the CRISPR-FNCas12A3 system of genetically engineered FnCas12a mutant EP16 can recognize the PAM sequence for precise site mutation and insertion. The CRISPR-FNCAS12A3 system solves the limitations of TTN PAM recognition by

Streptomyces with a high GC content (Jun Zhang et al., 2020). Additionally, the effector Cas9 (TdCas9_m) from *Treponema denticola*, Cas9 (NmCas9) from *Neisseria meningitides*, and *Corynebacterium glutamicum* codon-optimized Cpf1 (FnCpf1_cg) from *Francisella tularensis* do not affect cell growth (Sun et al., 2018).

In addition to replacing the exogenous Cas protein, using the endogenous CRISPR/Cas system is also an excellent choice to reduce cytotoxicity and the cellular stress and compatibility caused by the external immune system. Luo et al. first attempted to knockout the *Cas3* gene in *E. coli* K-12 by using the endogenous CRISPR/Cas system to enable target gene expression and multiple regulations. Accidental self-targeting is assumed to be forced in the case of *Cas* gene deletion-mediated deactivation of the endogenous CRISPR/Cas system. Specifically, this study used Class 1 CRISPR/Cas systems containing Cascade and Cas3 for DNA targeting. The researchers inactivated Cas3 protein and added a constitutive promoter upstream of the cascade operon, which drove transcription to become individual crRNAs, which could participate in binding DNA and insert into the PAM sequence to inhibit transcription (Figure 1B). This study verified the possibility of endogenous CRISPR system development (Luo et al., 2015). Subsequently, based on this principle, Qin et al. constructed a system for the industrial strain *Gluconobacter oxydans*. They speculated that Cas3 was naturally inactivated by the nucleic acid sequence, and the effectiveness of the system was verified by metabolic engineering. In their study, the endogenous CRISPR interference system (CRISPRi, dCas9 mediated system) was used to study the central carbon metabolic pathway PPP and EDP of *G. oxydans*, achieving flexible and reliable genome editing of *G. oxydans* (Qin et al., 2021).

Unlike eukaryotes, most prokaryotes only have an HDR repair mechanism; thus, DSBs caused by the CRISPR/Cas9 system cannot be repaired in cells that have not successfully transformed donor DNA, giving rise to non-viable cells. Therefore, introducing the NHEJ repair mechanism is vital in reducing DSB-induced toxicity. Huang et al. introduced the NHEJ repair pathway by overexpressing the Ku protein and ligase D derived from *Mycobacterium* in *E. coli*. They constructed an efficient gene knockout system without donor DNA based on CRISPR/Cas9 and demonstrated for the first time the potential of this system in genetic engineering. The researchers reported that the system was able to delete a chromosome fragment of up to 83 kb with an efficiency of over 85%. Notably, this system does not need to rely on efficient, competent cells, possibly because of its lower cytotoxicity compared to other CRISPR/Cas9 systems (Huang et al., 2019). Similarly, Yan et al. innovatively established genome editing tools that reduced cytotoxicity based on CRISPR cutting and NHEJ repair pathways. Ultimately, this resulted in the formation of deletion mutants in *Mycobacterium tuberculosis*, wherein the system had access to double mutations simultaneously, demonstrating its potential for screening a large-scale targeted point in drugs (Yan et al., 2020). In addition to Ku protein and ligase D, Su et al. (2019) found that T4 phage ligase could efficiently repair the NHEJ system alone, enabling *E. coli* to achieve a higher survival rate after producing DSB (Figure 2C).

The transposon is a class of DNA sequences that can be transcribed or reverse transcribed in the genome. The present study has found that a transposase from *Scytonema Hofmanni* can associate with the CRISPR effector (Cas12k) to form CRISPR-associated transposase (CAST). CAST system utilized Cas12k to identify and bind to specific sites in the genome, and directly inserts exogenous gene fragments into target sites by transposase. Cas12k in this system does not revolve in the cleavage of DNA and homologous recombination due to the absence of endonuclease activity. Experimental results showed that the system could successfully insert 2.5 kb DNA fragments at the target site and the success rate is up to 80%, which was far better than the traditional CRISPR system based on HDR (Strecker et al., 2019). Similarly, Klompe et al. also found the Tn7-like transposon in *Vibrio cholerae*. They found that after Cascade recognized the target DNA, Cascade directly combined with the protein (TniQ) to guide the transposon into the target site in the genome. After many experiments, transposons were accurately and completely delivered to the target site of the bacterial genome (Klompe et al., 2019).

Increasing the target editing efficiency

Strategies for better gene-editing efficiency

Based on the CRISPR/Cas system components, we distinguished the optimization schemes from Cas protein, gRNA, and donor DNA (Figure 3). Many optimizations may integrate two, all of them, or in combination with other homologous recombination systems, all of which aim to improve the editing efficiency of the CRISPR/Cas system.

Cas9 mutants may be selected to improve the editing efficiency of the CRISPR/Cas system (Table 1). The Cas9 variants mentioned earlier in this review showed a higher specificity than the wild type. Several other examples have been introduced to improve editing efficiency. A variant of the V-type CRISPR system, Cpf1, was promoted, which showed that Cpf1 produced staggered DSBs, followed by the formation of adhesive endpoints. This is more conducive to HDR pathway repair, thus improving the accuracy of gene editing (Zetsche et al., 2015). Additionally, Cpf1 is more sensitive to sgRNA/DNA mismatches, its PAM sequence is more stringent, and Cpf1 is more suitable for targeting T-rich fragments of the genome (Ungerer and Pakrasi, 2016). The CRISPR/Cpf1-RecT system

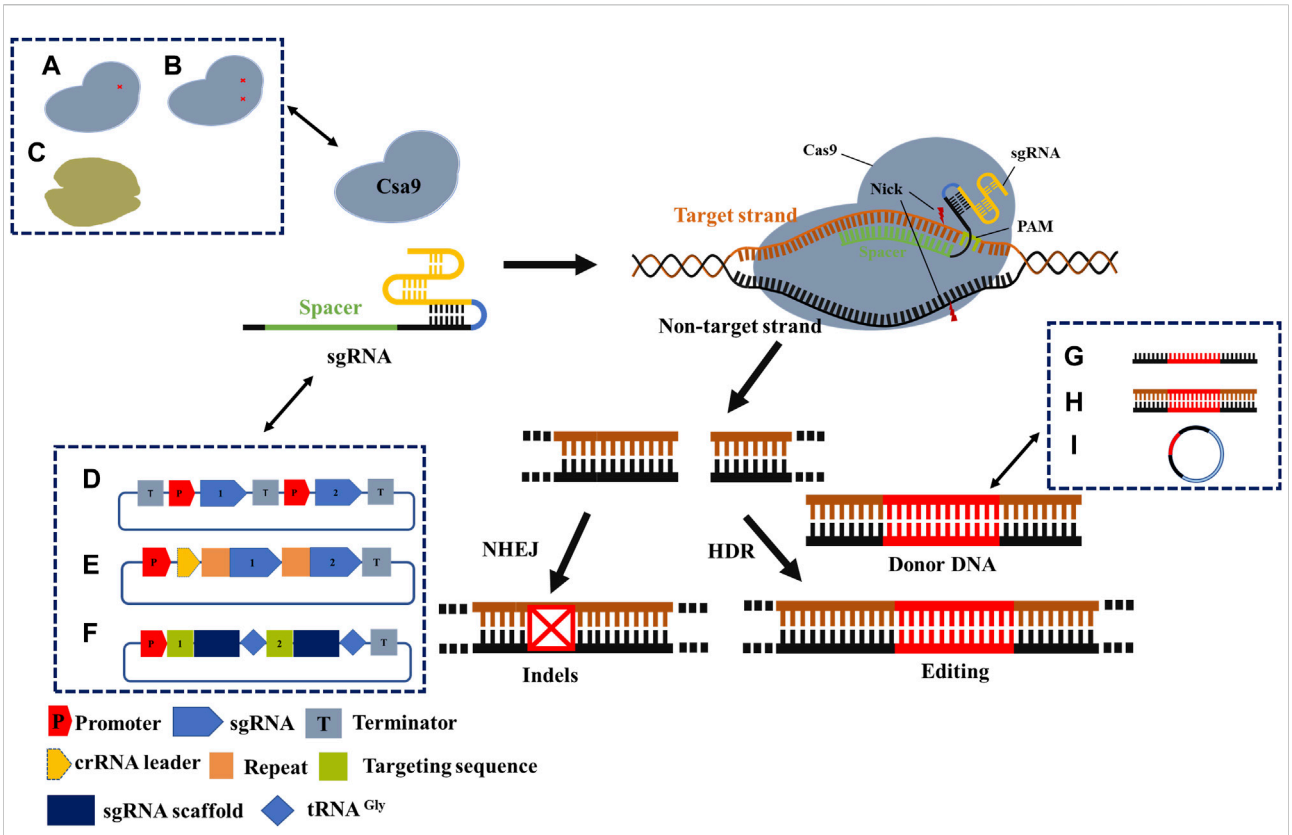


FIGURE 3
Mechanism and optimization strategy of CRISPR/Cas9 gene-editing system. Cas9 protein cleaves double-stranded DNA under the guidance of sgRNA to obtain DSB. Precise editing by the introduction of donor DNA in cells with HDR machinery, or formation of indels in cells with NHEJ repair machinery. A- C describes general optimization strategies. **(A)** Cas9 nickase; **(B)** Dead Cas9; **(C)** The other Cas proteins. **(D)** Independent expression of sgRNA; **(E)** Simultaneous expression of multiple sgRNAs; **(F)** tRNA process; **(G)** Donor DNA is provided as ssDNA; **(H)** Donor DNA is provided as dsDNA; **(I)** Donor DNA is provided as a plasmid. Recent applications of CRISPR/Cas engineering.

can improve the editing accuracy of PAM sequences, such as TTTC, TTTG, GTTG, or CTTC, reaching an accuracy of over 91.6%. Using the crRNA array, the CPF1-RECT system could simultaneously edit two and three genes, with an editing efficiency of 91.6% for two-gene editing. Using the CPF1-RECT system, the editing efficiency of the 1 kb DNA fragment was 79.6%, while the editing efficiency of 5 kb DNA fragments reached 91.3% (Nannan Zhao et al., 2020). Recently, Wu et al. (2021) discovered a micro-CRISPR-Cas12F system, which solved the problem caused by the large size of the CRISPR/Cas9 complex and promoted the application of viral approaches in therapeutic genome editing (Wu et al., 2021).

Improving the editing efficiency of the CRISPR/Cas system, especially the multiplexed genome engineering efficiency, is beneficial for expanding its application in synthetic biology. Thus, this is a primary focus area for researchers developing novel CRISPR/Cas systems. Multiple editing technologies of CRISPR/Cas have been developed for a wide range of conventional industrial microbes, including *E. coli*, *Streptomyces avermitilis,* and *Bacillus subtilis* (Banno et al., 2018; Zhong et al., 2019; Yu et al., 2020). For instance, by requiring a separate promoter/terminator for each sgRNA transcription unit (Figure 3D), Banno et al. achieved the simultaneous editing of six genes in *E. coli* with an efficiency of 87.5% (Banno et al., 2018). Similarly, Zhong et al. achieved the simultaneous base editing of five (60% efficiency) and nine genes (efficiency not shown) in *Streptomyces avermitilis* (Zhong et al., 2019). Moreover, Yu et al. achieved the simultaneous editing of 3–4 genes in *Bacillus subtilis*, with efficiencies of 100% and 50%, respectively (Yu et al., 2020). In conclusion, this strategy enables high efficiency and wide application in multiplexed base editing, despite shortcomings such as large structure, repeated sequences, and unstable plasmids. Another alternative method can be found in the original CRISPR/Cas system, wherein multiple spacers and repeat sequences can be transcribed through the crRNA array via a promoter and terminator (Figure 3E). Li et al. (2020) applied CRISPR/dCas12a and crRNA arrays in *Corynebacterium glutamate* to inhibit four genes of the lysine synthesis pathway concurrently

by CRISPR interference (CRISPRi), whose inhibition exhibited an efficiency of over 90%. Bao et al. applied CRISPR/Cas9 and crRNA arrays to concurrently edit three genes in *Saccharomyces cerevisiae* (Bao et al., 2015). The tRNA-processed type II CRISPR system did not require the introduction of exogenous toxic proteins or shorter sequences (approximately 70 bp) and was more stable in multiplexed base editing (Figure 3F). In another study, the efficiency of simultaneous editing of eight genes was as high as 87% based on multiple sgRNAs used by the tRNA processing system in *Saccharomyces cerevisiae* (Zhang et al., 2019). Specifically, the function of the Cas protein and whether the cognition of sgRNA is mature should be taken into account.

Unlike the application of the sgRNA arrays described earlier, a recent study exported a general, precise genome editing method that detected only a specific sgRNA, denoted as prime editing (PE), using two specific components: the prime guide RNA (pegRNA) and engineered reverse transcriptase. pegRNA can bind to the specific region of DNA that must be edited and obtain a modification template from which it synthesizes DNA with the correct sequence. Notably, DNA repair mechanisms in cells eventually automatically integrate this newly synthesized sequence into their genomes. In contrast to the commonly used CRISPR/Cas system, this system does not require a DSB or donor DNA. In principle, 89% of the known genetic variations associated with human diseases can be corrected by gene editing (Tong et al., 2021).

Moreover, the concentration and delivery method of donor DNA are also critical factors that influence editing efficiency (Table 1). Donor DNA templates can be generated as linear fragments or plasmids. Linear fragments can be divided into linear single-stranded DNA (ssDNA) (Figure 3G) and double-stranded DNA (dsDNA) (Figure 3H). Genome-editing efficiency varies depending on the template provision method. In the case of the λ-red recombination system, the system shows a higher editing efficiency when providing templates in the form of ssDNA, which is superior to dsDNA. However, because of the limitation of ssDNA length, long homologous arms cannot be provided; therefore, large fragments cannot be readily knocked out or inserted. In addition to the linear segments, several researchers have chosen the plasmid form to provide donor DNA (Figure 3I). Although the construction of plasmids increases the complexity of the experiments, in prokaryote microbiology, the plasmid conversion efficiency is much higher than that of the linear segment. At the same time, providing the donor DNA plasmid from templates may guarantee the stability of the fragment and extend the period of restructuring effect. Combined with the improvement in the operation process, the editing efficiency and operation time can be significantly enhanced (Lin et al., 2015). When Zhao et al. developed the CRISPR/Cas9 system in *Escherichia coli*, donor DNA was provided in the form of a plasmid (Zhao et al., 2016). After the donor DNA was transferred into the host in the form of a free plasmid, positive transformants were selected for culture, and Cas9 and gRNA expression were induced simultaneously. Cas9 and gRNA were fully expressed in this process, and the donor DNA also acted for a longer time. Ultimately, the process from plasmid construction to obtaining the correct strain was completed within 3 days. Thus, compared to most linear fragment methods in this strain, the process developed by Zhao et al. is more time- and cost-efficient.

Developing chemical-producing cell factories using the CRISPR/Cas system is an important research direction. Compared to traditional gene-editing strategies, CRISPR/Cas technology has been applied to rapidly develop cell factories for various chemicals, owing to its flexibility (Table 2). Owing to the low efficiency of single-point gene editing, many research teams are currently attempting to expand this system. Glucose and xylose are the two most abundant sugars in renewable lignocelluloses. However, they cannot be used simultaneously due to the inhibition of carbon catabolites. Wu et al. developed a set of CRISPRi systems to increase the production of N-acetylglucosamine by repressing xylose utilization genes (zwf, pfkA, and glmM) in *Bacillus* subtilis (Wu et al., 2018). One of the best xylose utilization strains among *E. coli* was obtained using the CRISPR/Cas9-facilitated multiplex pathway optimization (CFPO) technique by Zhu et al. (2017). Four genes (*xyl*A, *xyl*B, *tkt*A, and *tal*B) in the xylose assimilation pathway are regulated by this system. The plasmid, which includes the random RBS library, was first constructed and then co-transformed with pRedCas9, and the original RBS position in the chromosomes was replaced by red-assisted recombination. Next, pgRNA was introduced into the gene plasmid to produce DSBs, eliminating strains that contained native RBS regions. Finally, three transcriptional units were modulated with 70% efficiency (Zhu et al., 2017). Both laid the foundation for further constructing cell factories that efficiently produce chemicals using lignocellulosic resources. In metabolic engineering, a heterologous gene cassette for pyruvate decarboxylase (*pdc*) and alcohol dehydrogenase (*adh*B) from *Zymomonas mobilis* converts pyruvate to ethanol in *E. coli*. Dong et al. (2018) increased ethanol production by three-fold using CRISPR activation (CRISPR A, activation genes), which was optimized to contain a SOXS activator (Dong et al., 2018).

Multiplex CRISPRi via target gRNA arrays and a dCas9 protein have been widely applied to simultaneously repress the expression of multiple genomic DNA loci and has higher efficiency than the CRISPR single gene-editing system. Ye et al. achieved the upregulation of central metabolic enzymes by controlling proteolysis using a combination of the CRISPRi cascade system (Ye et al., 2021). In *E. coli*, Gao et al. used multiplexed CRISPRi to simultaneously inhibit the expression of three crucial enzyme-encoding genes involved in the malic acid biosynthesis pathway. Using this strategy, the titer of the

TABLE 2 Recent applications of CRISPR/Cas engineering in prokaryotes.

Strains	Stratges	Applications	Efficiency	Ref
B. subtilis	CRISPRi and 27 arrays containing sgRNAs with different repression capacities targeting the *zwf, pfkA* and *glmM*	N-acetylglucosamine	The production of N-acetylglucosamine increased 84.1% and a 3-L fed-batch bioreactor reached 103.1 ± 2.11 g/L and 1.17 ± 0.024 g/L/h	Wu et al. (2018)
E. coli	CRISPR Cascade system	Alanine	The overall yield in this process was 0.74 g ala- nine/g of glucose, with the yield in the production phase reaching 0.81 g alanine/g of glucose	Ye et al. (2021)
E. coli	A multi-gene CRISPRi/a control program	Ethanol	The production quadrupled	Dong et al. (2018)
E. coli	CRISPR/Cas9-facilitated multiplex pathway optimization technique (regulating the expression of multiple genes)	Xylose	The utilization rate increased by 3 times	Zhu et al. (2017)
E. coli	Multiplex CRISPRi	Malic acid	A 2.3-fold increase	Gao et al. (2018)
C. glutamicum	Multiplex CRISPRi	Squalene	3.4 times higher than that of the parental strains	Park et al. (2019)
C. glutamicum	Multiplex CRISPRi	L-pyrrolysine	Production improved by 39% compared with wild-type	Park et al. (2018)
Synechocystis sp	Multiplex CRISPRi	Octadecanol	Increased octadecanol productivity threefold	Kaczmarzyk et al. (2018)
K. pneumoniae	CRISPRi with trp operon	1, 3-propylene glycol and 3-hydroxypropionic acid (3-HP)	Produced 58.9 g/L 3- HP in a 5 L bioreactor	Zhao et al. (2021)

final product, malic acid, was 2.3 times higher than that of the initial strain (Gao et al., 2018). In addition, CRISPRi and sequence-specific sgRNAs successfully achieved single or dual gene inhibition, which increased l-lysine production by 39% compared to that of wild-type *Corynebacterium glutamicum* (Park et al., 2018). Wu et al. have been working on developing synthetic biology using CRISPR as a tool for *C. glutamicum* for several years. Recently, they successfully used CRISPRi to inhibit target genes and achieved a REDOX balance of the precursor, blocking the availability of isoprene diphosphate, resulting in improved *Corynebacterium* that produced squalene from glucose at a rate 3.4-fold higher than the parental strain (Park et al., 2019). Moreover, Kaczmarzyk et al. (2018) used multiplex CRISPRi to allow for the simultaneous partial inhibition of up to six genes, downregulating PlsX (SLR1510) and increasing octadecanol production in *Synechocystis* by three-fold (Kaczmarzyk et al., 2018). Beyond these types of strains, several researchers are opting for using CRISPR systems to drive the development of synthetic biology for unfamiliar strains. For example, Zhao et al. (2021) reported that the glycerol oxidation and reduction pathways of *Klebsiella pneumoniae* could be switched to produce 58.9 g/L of 1,3-propylene glycol and 3-hydroxypropionic acid (3HP) by coupling the CRISPRi system with the trp operon (Zhao et al., 2021). Thus, these studies highlight the advantages of the CRISPR/Cas system in gene editing and provide insights into the construction of metabolically engineered strains and the development of cell factories via efficient and accurate modifications of large fragment integration and multi-site editing.

Exploitation strategy of CRISPR/Cas systems in new species

Currently, the CRISPR/Cas system has been widely developed and applied in commonly studied microorganisms. However, due to some inherent advantages of this system, its development in unconventional hosts is also being carried out based on the three main components (e.g., Cas, gRNA, and donor DNA) required by CRISPR/Cas system. In this context, the system construction strategies can be divided into single-vector expression, multiple-vector expression, and chromosome integration.

Single-vector expression is the most common construction strategy of the CRISPR/Cas gene-editing system, that is, the integration of Cas, gRNA, and donor DNA into the same vector. This strategy has been widely used in the early development of the system combined with GoldnGate molecular cloning techniques, such as CPEC, and has allowed for the shortening of editing cycles. However, its defects lie in the large size of the entire plasmid, which imposes a sizeable physiological burden on cells. Additionally, the plasmid must be imported repeatedly after each round of editing, which requires an effective competent cell. To address this, Huang et al. utilized two promoters (P-thl and P-araE) to express Cas9 and sgRNA in *Clostridium ljungdahlii*, with deletion efficiencies of up to 100% (Huang et al., 2016). Similarly, Xu et al. (2015) developed *Streptococcus pyogenes* CRISPR/Cas9 to edit the genome of *Clostridium cellulolyticum*, which contributed to a high editing efficiency, even when using 200 bp homologous arms.

The multi-vector expression system is used to express CRISPR/Cas components through multiple plasmids, which partly solves the problem of a single-vector expression strategy. Cas can exist in cells for a long time under the

control of an inducible promoter. Only gRNA and donor DNA must be constructed for each round of editing, thus addressing the problem of large plasmids in the single-vector strategy, thereby increasing the flexibility of the system. In addition to the two-plasmid system, a strain with many plasmids, such as *E. coli*, can be used to establish a three-plasmid system. Zhu et al. developed a three-plasmid system for Cas9, gRNA, and donor DNA expressed individually, facilitating plasmid construction because of the independent expression vectors. The system also showed high efficiency in multi-site editing because of the continuous expression of the donor DNA (Zhu et al., 2017). Recently, Kozaeva et al. constructed a group of plasmids based on CRISPR and dCas9 in *Pseudomonas putida* that could increase the expression of *glt*A encoding citrate synthase. As a result, *acc*A, an essential gene encoding the subunit of acetyl-CoA carboxylase complex A, was dynamically reduced. The content of acetyl-CoA reconnection was increased by eight times (Kozaeva et al., 2021). Because such multiple plasmids editing systems are limited by the number of plasmids available in the strains, their development in unconventional microorganisms has yet to be fully developed.

The chromosome integrated strategy shows potential for developing strains that do not have multiple available plasmids. However, this strategy involves integrating the Cas gene and gRNA into chromosomes, which requires a traceless operation because of its complex process. In particular, if an iterative cycle system is formed, it is generally necessary to introduce a traceless operating system, leading to a lower prevalence. However, the chromosome integration strategy also has unique advantages, including more stable Cas gene expression, which is less likely to be lost, and the fact that the physiological burden on bacteria is small. In a classic experiment, Westbrook et al. constructed a CRISPR/Cas9 gene-editing box strategy in *Bacillus subtilis* and integrated spCas9 and gRNA expression frames into lacA and thrC sites of chromosomes, respectively, followed by providing dsDNA as a donor DNA template to achieve genome editing. Notably, the researchers achieved high single-point editing efficiency and induced the simultaneous deletion of double genes in the genome, with an efficiency of up to 85% (Westbrook et al., 2016).

Concluding remarks and future perspectives

In recent years, genome-editing relying on CRISPR/Cas and its derivative technologies, such as precise editing, multi-gene editing, and precise regulation, has experienced rapid innovation, giving rise to comprehensive solutions and greater flexibility in the development of synthetic biology. In turn, this has allowed for large-scale studies of genes and mutations. However, its development has been hampered by issues with editing efficiency due to off-target effects and cytotoxicity, as well as multiple editing efficiency.

The CRISPR/Cas system may be improved via optimization. To this end, a range of techniques have been applied to detect unexpected mutations to improve in-target efficiency and reduce or avoid off-target effects. Additionally, the development of mutants and analogs of Cas9 also represents an effective method to avoid off-target effects, reduce toxicity, and improve editing efficiency. Using ACR competitive binding, the Cas protein can regulate CRISPR expression. Furthermore, selecting ssDNA rather than dsDNA as the donor DNA can ensure the stable existence of fragments and enable a higher editing efficiency. Furthermore, the rise of DNA base editors without DSBs has allowed researchers to solve the problems associated with the original system.

Researchers have continued to focus on improving the CRISPR/Cas system. A recent report found that CRISPR/Cas9 editing produces micronuclei, and chromosomes are structurally deficient, thus initiating a mutation process known as chromothripsis. Chromothripsis is an extensive chromosomal rearrangement limited to one or more chromosomes, causing congenital diseases and cancer in humans. This suggests that chromothripsis is another manifestation of the off-target effects of CRISPR/Cas9 (Leibowitz et al., 2021). Gene-editing technology has been widely applied in microbial synthetic biology and can quickly, efficiently, and accurately create more chassis organisms suitable for the production of high value-added products. CRISPR gene-editing technology may also contribute to the end of the coronavirus pandemic. Synthetic biology and CRISPR/Cas technology are emerging disciplines, albeit with several unresolved issues. However, with the rapid development and in-depth research of these two complementary technologies, improved applications are likely to be obtained in the near future.

Author contributions

WL and CH conceived of the idea. They wrote and revised the manuscript together under the guidance of JC. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Beijing Municipal Natural Science Foundation (L202045).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Albitar, A., Rohani, B., Will, B., Yan, A., and Gallicano, G. I. (2018). The application of CRISPR/Cas technology to efficiently model complex cancer genomes in stem cells. *J. Cell. Biochem.* 119 (1), 134–140. doi:10.1002/jcb.26195

Banno, S., Nishida, K., Arazoe, T., Mitsunobu, H., and Kondo, A. (2018). Deaminase-mediated multiplex genome editing in *Escherichia coli*. *Nat. Microbiol.* 3 (4), 423–429. doi:10.1038/s41564-017-0102-6

Bao, Z., Xiao, H., Liang, J., Zhang, L., Xiong, X., Sun, N., et al. (2015). Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in *Saccharomyces cerevisiae*. *Acs Synth. Biol.* 4 (5), 585–594. doi:10.1021/sb500255k

Barrangou, R. (2014). Cas9 targeting and the CRISPR revolution. *Science* 344 (6185), 707–708. doi:10.1126/science.1252964

Casini, A., Olivieri, M., Petris, G., Bianchi, A., Montagna, C., Reginato, G., et al. (2018). A highly specific SpCas9 variant is identified by *in vivo* screening in yeast. *Nat. Biotechnol.* 29 (12), 265–271. doi:10.1038/nbt.4066

Chang Li, C., Psatha, N., Gil, S., Wang, H. J., Papayannopoulou, T., and Lieber, A. (2018). HDAd5/35(++) adenovirus vector expressing anti-CRISPR peptides decreases CRISPR/Cas9 toxicity in human hematopoietic stem cells. *Mol. Ther. - Methods & Clin. Dev.* 9, 390–401. doi:10.1016/j.omtm.2018.04.008

Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., et al. 2017. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature*, 550(7676): 407-410. doi:10.1038/nature24268

Cong, L., Ran, F. A., Cox, D., Lin, S. L., Barretto, R., Habib, N., et al. (2013). Multiplex genome engineering using CRISPR/Cas Systems. *Science* 339 (6121), 819–823. doi:10.1126/science.1231143

Crosetto, N., Mitra, A., Silva, M. J., Bienko, M., Dojer, N., Wang, Q., et al. (2013). Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* 10 (4), 361–365. doi:10.1038/nmeth.2408

Debin Zhang, D., Hussain, A., Manghwar, H., Xie, K. B., Xie, S. S., Zhao, S. H., et al. (2020). Genome editing with the CRISPR-cas system: An art, ethics and global regulatory perspective. *Plant Biotechnol. J.* 18 (8), 1651–1669. doi:10.1111/pbi.13383

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y. J., Pirzada, Z. A., et al. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471 (7340), 602–607. doi:10.1038/nature09886

Dong, C., Fontana, J., Patel, A., Carothers, J. M., and Zalatan, J. G. (2018). Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat. Commun.* 9, 2489. doi:10.1038/s41467-018-04901-6

Fu, Y. F., Sander, J. D., Reyon, D., Cascio, V. M., and Joung, J. K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32 (3), 279–284. doi:10.1038/nbt.2808

Gao, C., Wang, S. H., Hu, G. P., Guo, L., Chen, X. L., Xu, P., et al. (2018). Engineering *Escherichia coli* for malate production by integrating modular pathway characterization with CRISPRi-guided multiplexed metabolic tuning. *Biotechnol. Bioeng.* 115 (3), 661–672. doi:10.1002/bit.26486

Grunewald, J., Zhou, R. H., Iyer, S., Lareau, C. A., Garcia, S. P., Aryee, M. J., et al. 2019. CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.*, 37(9): 1041-1048. doi:10.1038/s41587-019-0236-6

Haeussler, M., and Concordet, J. P. (2016). Genome editing with CRISPR-cas9: Can it get any better? *J. Genet. Genomics* 43 (5), 239–250. doi:10.1016/j.jgg.2016.04.008

Haeussler, M. (2020). CRISPR off-targets: A question of context. *Cell Biol. Toxicol.* 36 (1), 5–9. doi:10.1007/s10565-019-09497-1

Heigwer, F., Kerr, G., and Boutros, M. (2014). E-CRISP: Fast CRISPR target site identification. *Nat. Methods* 11 (2), 122–123. doi:10.1038/nmeth.2812

Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W. X., Chen, L. W., Sun, N., et al. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556 (7699), 57–63. doi:10.1038/nature26155

Huang, H., Chai, C. S., Li, N., Rowe, P., Minton, N. P., Yang, S., et al. (2016). CRISPR/Cas9-Based Efficient genome editing in *Clostridium ljungdahlii*, an autotrophic gas-fermenting bacterium. *Acs Synth. Biol.* 5 (12), 1355–1361. doi:10.1021/acssynbio.6b00044

Huang, C. Y., Ding, T. T., Wang, J. G., Wang, X. Q., Guo, L. W., Wang, J. L., et al. (2019). CRISPR-Cas9-assisted native end-joining editing offers a simple strategy for efficient genetic engineering in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 103 (20), 8497–8509. doi:10.1007/s00253-019-10104-w

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide-sequence of the iap gene, responsible for alkaline-phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene-product. *J. Bacteriol.* 169 (12), 5429–5433. doi:10.1128/jb.169.12.5429-5433.1987

Ishino, Y., Krupovic, M., and Forterre, P. (2018). History of CRISPR-Cas from encounter with a mysterious repeated sequence to genome editing technology. *J. Bacteriol.* 200 (7), e00580-17. doi:10.1128/JB.00580-17

Jiang, Y., Qian, F. H., Yang, J. J., Liu, Y. M., Dong, F., Xu, C. M., et al. (2017). CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* 8, 15179. doi:10.1038/ncomms15179

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337 (6096), 816–821. doi:10.1126/science.1225829

Jing Zhao, J., Fang, H., and Zhang, D. W. (2020). Expanding application of CRISPR-Cas9 system in microorganisms. *Synth. Syst. Biotechnol.* 5 (4), 269–276. doi:10.1016/j.synbio.2020.08.001

Jun Zhang, J., Zhang, D., Zhu, J., Liu, H., Liang, S., and Luo, Y. (2020). Efficient multiplex genome editing in *Streptomyces* via engineered CRISPR-Cas12a systems. *Front. Bioeng. Biotechnol.* 8, 726. doi:10.3389/fbioe.2020.00726

Kaczmarzyk, D., Cengic, I., Yao, L., and Hudson, E. P. (2018). Diversion of the long-chain acyl-ACP pool in Synechocystis to fatty alcohols through CRISPRi repression of the essential phosphate acyltransferase PlsX. *Metab. Eng.* 45, 59–66. doi:10.1016/j.ymben.2017.11.014

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z. L., et al. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529 (7587), 490–495. doi:10.1038/nature16526

Klompe, S. E., Vo, P. L. H., Halpin-Healy, T. S., Sternberg, S. H., and Nguyen, N. T. (2019). Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature* 571 (7764), 219–225. doi:10.1038/s41586-019-1323-z

Kozaeva, E., Volkova, S., Matos, M. R. A., Mezzina, M. P., Wulff, T., Volke, D. C., et al. (2021). Model-guided dynamic control of essential metabolic nodes boosts acetyl-coenzyme A-dependent bioproduction in rewired *Pseudomonas putida*. *Metab. Eng.* 67, 373–386. doi:10.1016/j.ymben.2021.07.014

Lee, J., Jung, M. H., Jeong, E., and Lee, J. K. (2019). Using Sniper-Cas9 to minimize off-target effects of CRISPR-Cas9 without the loss of on-target activity via directed evolution. *J. Vis. Exp.* 144, 8. doi:10.3791/59202

Lei Li, L., Wei, K. K., Zheng, G. S., Liu, X. C., Chen, S. X., Jiang, W. H., et al. (2018). CRISPR-Cpf1-Assisted multiplex genome editing and transcriptional repression in *Streptomyces*. *Appl. Environ. Microbiol.* 84 (18), e00827-18. doi:10.1128/AEM.00827-18

Leibowitz, M. L., Papathanasiou, S., Doerfler, P. A., Blaine, L. J., Sun, L. L., Yao, Y., et al. (2021). Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *Nat. Genet.* 53 (6), 895–905. doi:10.1038/s41588-021-00838-7

Li, M., Chen, J., Wang, Y., Liu, J., Sun, J., Chen, N., et al. (2020). Efficient multiplex gene repression by CRISPR-dCpf1 in *Corynebacterium glutamicum*. *Front. Bioeng. Biotechnol.* 8, 357. doi:10.3389/fbioe.2020.00357

Liang, F., Han, M. G., Romanienko, P. J., and Jasin, M. (1998). Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* 95 (9), 5172–5177. doi:10.1073/pnas.95.9.5172

Lieber, M. R. (2010). "The mechanism of Double-Strand DNA break repair by the nonhomologous DNA End-Joining pathway," in *Annu. Rev. Biochem.* Editors R. D. Kornberg, C. R. H. Raetz, J. E. Rothman, and J. W. Thorner, 79, 181–211. doi:10.1146/annurev.biochem.052308.093131

Lin, Z., Wang, Z., Tang, Y-J., Huang, C., Zhang, Y., Chen, T., et al. (2015). Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 meditated genome editing. *Metab. Eng.* 31, 13–21. doi:10.1016/j.ymben.2015.06.006

Liu, H., Ding, Y. D., Zhou, Y. Q., Jin, W. Q., Xie, K. B., and Chen, L. L. (2017). CRISPR-P 2.0: An improved CRISPR-cas9 tool for genome editing in plants. *Mol. Plant* 10 (3), 530–532. doi:10.1016/j.molp.2017.01.003

Luo, M. L., Mullis, A. S., Leenay, R. T., and Beisel, C. L. (2015). Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. *Nucleic Acids Res.* 43 (1), 674–681. doi:10.1093/nar/gku971

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J. J., Charpentier, E., Horvath, P., et al. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9 (6), 467–477. doi:10.1038/nrmicro2577

Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., et al. (2020). Evolutionary classification of CRISPR-cas systems: A burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18 (2), 67–83. doi:10.1038/s41579-019-0299-x

Martin, F., Sánchez-Hernández, S., Gutiérrez-Guerrero, A., Pinedo-Gomez, J., and Benabdellah, K. (2016). Biased and unbiased methods for the detection of off-target cleavage by CRISPR/Cas9: An overview. *Int. J. Mol. Sci.* 17 (9), 1507. doi:10.3390/ijms17091507

McNeely, K., Xu, Y., Bennette, N., Bryant, D. A., and Dismukes, G. C. (2010). Redirecting reductant flux into hydrogen production via metabolic engineering of fermentative carbon metabolism in a cyanobacterium. *Appl. Environ. Microbiol.* 76 (15), 5032–5038. doi:10.1128/AEM.00862-10

Mojica, F. J. M., Juez, G., and Rodriguezvalera, F. (1993). Transcription at different salinities of haloferax-mediterranei sequences adjacent to partially modified psti sites. *Mol. Microbiol.* 9 (3), 613–621. doi:10.1111/j.1365-2958.1993.tb01721.x

Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60 (2), 174–182. doi:10.1007/s00239-004-0046-3

Mojica, F. J., and Rodriguez-Valera, F. (2016). The discovery of CRISPR in archaea and bacteria. *FEBS J.* 283 (17), 3162–3169. doi:10.1111/febs.13766

Nannan Zhao, N. N., Li, L., Luo, G. J., Xie, S., Lin, Y., Han, S. Y., et al. (2020). Multiplex gene editing and large DNA fragment deletion by the CRISPR/Cpf1-RecE/T system in *Corynebacterium glutamicum*. *J. Ind. Microbiol. Biotechnol.* 47 (8), 599–608. doi:10.1007/s10295-020-02304-5

Oliveros, J. C., Franch, M., Tabas-Madrid, D., San-Leon, D., Montoliu, L., Cubas, P., et al. (2016). Breaking-Cas-interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes. *Nucleic Acids Res.* 44 (W1), W267–W271. doi:10.1093/nar/gkw407

Park, J., Shin, H., Lee, S. M., Um, Y., and Woo, H. M. (2018). RNA-guided single/double gene repressions in *Corynebacterium glutamicum* using an efficient CRISPR interference and its application to industrial strain. *Microb. Cell Fact.* 17, 4. doi:10.1186/s12934-017-0843-1

Park, J., Yu, B. J., Cho, J. I., and Woo, H. M. (2019). Heterologous production of squalene from glucose in engineered *Corynebacterium glutamicum* using multiplex crispr interference and high-throughput fermentation. *J. Agric. Food Chem.* 67 (1), 308–319. doi:10.1021/acs.jafc.8b05818

Pattanayak, V., Lin, S., Guilinger, J. P., Ma, E. B., Doudna, J. A., and Liu, D. R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* 31 (9), 839–843. doi:10.1038/nbt.2673

Pattharaprachayakul, N., Lee, M., Incharoensakdi, A., and Woo, H. M. (2020). Current understanding of the cyanobacterial CRISPR-Cas systems and development of the synthetic CRISPR-Cas systems for cyanobacteria. *Enzyme Microb. Technol.* 140, 109619. doi:10.1016/j.enzmictec.2020.109619

Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., et al. (2016). Naturally occurring off-switches for CRISPR-Cas9. *Cell* 167 (7), 1829–1838.e9. doi:10.1016/j.cell.2016.11.017

Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology-Sgm* 151, 653–663. doi:10.1099/mic.0.27437-0

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., et al. (2013). Repurposing CRISPR as an RNA-Guided platform for Sequence-Specific control of gene expression. *Cell* 152 (5), 1173–1183. doi:10.1016/j.cell.2013.02.022

Qin, Z. J., Yang, Y. T., Yu, S. Q., Liu, L., Chen, Y., Chen, J., et al. (2021). Repurposing the endogenous Type I-E CRISPR/Cas system for gene repression in *Gluconobacter oxydans* WSH-003. *Acs Synth. Biol.* 10 (1), 84–93. doi:10.1021/acssynbio.0c00456

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., et al. (2013a). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154 (6), 1380–1389. doi:10.1016/j.cell.2013.08.021

Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., et al. (2013b). Double nicking by RNA-Guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154 (6), 1380–1389. doi:10.1016/j.cell.2013.08.021

Ren, X. J., Yang, Z. H., Xu, J., Sun, J., Mao, D. C., Hu, Y. H., et al. (2014). Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in drosophila. *Cell Rep.* 9 (3), 1151–1162. doi:10.1016/j.celrep.2014.09.044

Sovova, T., Kerins, G., Demnerova, K., and Ovesna, J. (2017). Genome editing with engineered nucleases in economically important animals and plants: State of the art in the research pipeline. *Curr. Issues Mol. Biol.* 21, 41–62. doi:10.21775/cimb.021.041

Stanley, S. Y., and Maxwell, K. L. (2018). "Phage-Encoded anti-CRISPR defenses,". *Annu. Rev. Genet.* Editor N. M. Bonini (Palo Alto: Annual Reviews), 52, 445. doi:10.1146/annurev-genet-120417-031321

Strecker, J., Ladha, A., Gardner, Z., Schmid-Burgk, J. L., Makarova, K. S., Koonin, E. V., et al. (2019). RNA-guided DNA insertion with CRISPR-associated transposases. *Science* 365 (6448), 48–53. doi:10.1126/science.aax9181

Su, T. Y., Liu, F. P., Chang, Y. Z., Guo, Q., Wang, J. S., Wang, Q., et al. (2019). The phage T4 DNA ligase mediates bacterial chromosome DSBs repair as single component non-homologous end joining. *Synth. Syst. Biotechnol.* 4 (2), 107–112. doi:10.1016/j.synbio.2019.04.001

Sun, B. B., Yang, J. J., Yang, S., Ye, R. D., Chen, D. J., and Jiang, Y. (2018). A CRISPR-cpf1-assisted non-homologous end joining genome editing system of *Mycobacterium smegmatis*. *Biotechnol. J.* 13 (9), 1700588. doi:10.1002/biot.201700588

Tang, Y. Y., and Fu, Y. (2018). Class 2 CRISPR/Cas: An expanding biotechnology toolbox for and beyond genome editing. *Cell Biosci.* 8, 59. doi:10.1186/s13578-018-0255-x

Tong, Y. J., Jorgensen, T. S., Whitford, C. M., Weber, T., and Lee, S. Y. (2021). A versatile genetic engineering toolkit for *E. coli* based on CRISPR-prime editing. *Nat. Commun.* 12 (1), 5206. doi:10.1038/s41467-021-25541-3

Ungerer, J., and Pakrasi, H. B. (2016). Cpf1 is a versatile tool for crispr genome editing across diverse species of *Cyanobacteria*. *Sci. Rep.* 6, 39681. doi:10.1038/srep39681

Wang, Y., Zhang, Z. T., Seo, S. O., Lynn, P., Lu, T., Jin, Y. S., et al. (2016). Bacterial genome editing with CRISPR-cas9: Deletion, integration, single nucleotide modification, and desirable "clean" mutant selection in *Clostridium beijerinckii* as an example. *Acs Synth. Biol.* 5 (7), 721–732. doi:10.1021/acssynbio.6b00060

Westbrook, A. W., Moo-Young, M., and Chou, C. P. (2016). Development of a CRISPR-Cas9 tool kit for comprehensive engineering of *Bacillus subtilis*. *Appl. Environ. Microbiol.* 82 (16), 4876–4895. doi:10.1128/AEM.01159-16

Wu, Y. K., Chen, T. C., Liu, Y. F., Lv, X. Q., Li, J. H., Du, G. C., et al. (2018). CRISPRi allows optimal temporal control of N-acetylglucosamine bioproduction by a dynamic coordination of glucose and xylose metabolism in *Bacillus subtilis*. *Metab. Eng.* 49, 232–241. doi:10.1016/j.ymben.2018.08.012

Wu, Z. W., Zhang, Y. F., Yu, H. P., Pan, D., Wang, Y. J., Wang, Y. N., et al. (2021). Programmed genome editing by a miniature CRISPR-Cas12f nuclease. *Nat. Chem. Biol.* 17 (11), 1132–1138. doi:10.1038/s41589-021-00868-6

Xu, T., Li, Y. C., Shi, Z., Hemme, C. L., Li, Y., Zhu, Y. H., et al. (2015). Efficient genome editing in Clostridium cellulolyticum via CRISPR-cas9 nickase. *Appl. Environ. Microbiol.* 81 (13), 4423–4431. doi:10.1128/AEM.00873-15

Yan, M. Y., Li, S. S., Ding, X. Y., Guo, X. P., Jin, Q., and Sun, Y. C. (2020). A CRISPR-assisted nonhomologous end-joining strategy for efficient genome editing in *Mycobacterium tuberculosis*. *Mbio* 11 (1), e02364-19. doi:10.1128/mBio.02364-19

Ye, Z. X., Li, S., Hennigan, J. N., Lebeau, J., Moreb, E. A., Wolf, J., et al. (2021). Two-stage dynamic deregulation of metabolism improves process robustness & scalability in engineered *E. coli*. *Metab. Eng.* 68, 106–118. doi:10.1016/j.ymben.2021.09.009

Yeo, W. L., Heng, E., Tan, L. L., Lim, Y. W., Lim, Y. H., Hoon, S., et al. (2019). Characterization of Cas proteins for CRISPR-Cas editing in streptomycetes. *Biotechnol. Bioeng.* 116 (9), 2330–2338. doi:10.1002/bit.27021

Yu, S., Price, M. A., Wang, Y., Liu, Y., Wang, M., Ni, X., et al. (2020). CRISPR-dCas9 mediated cytosine deaminase base editing in *Bacillus subtilis*. *ACS Synth. Biol.* 9 (7), 1781–1789. doi:10.1021/acssynbio.0c0015110.1021/acssynbio.0c00151

Zeballos, C. M. A., and Gaj, T. (2021). Next-Generation CRISPR technologies and their applications in gene and cell therapy. *Trends Biotechnol.* 39 (7), 692–705. doi:10.1016/j.tibtech.2020.10.010

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-cas system. *Cell* 163 (3), 759–771. doi:10.1016/j.cell.2015.09.038

Zhang, Y., Wang, J., Wang, Z., Zhang, Y., Shi, S., Nielsen, J., et al. (2019). A gRNA-tRNA array for CRISPR-Cas9 based rapid multiplexed genome editing in *Saccharomyces cerevisiae*. *Nat. Commun.* 10 (1), 1053. doi:10.1038/s41467-019-09005-3

Zhao, D., Yuan, S., Xiong, B., Sun, H., Ye, L., Li, J., et al. (2016). Development of a fast and easy method for *Escherichia coli* genome editing with CRISPR/Cas9. *Microb. Cell Fact.* 15 (1), 205. doi:10.1186/s12934-016-0605-5

Zhao, P., Li, Q. Y., Tian, P. F., and Tan, T. W. (2021). Switching metabolic flux by engineering tryptophan operon-assisted CRISPR interference system in *Klebsiella pneumoniae*. *Metab. Eng.* 65, 30–41. doi:10.1016/j.ymben.2021.03.001

Zhaohui, Z., Simon, S., Qiurong, R., Lijia, Y., Yu, B., Caiyan, Q., et al. (2020). Improving Plant genome editing with high-fidelity xCas9 and non-canonical PAM-targeting Cas9-NG. *Mol. Plant* 12 (7), 1027–1036. doi:10.1016/j.molp.2019.03.011

Zhong, Z., Guo, J., Deng, L., Chen, L., Sun, Y., Li, S., et al. (2019). Base editing in *Streptomyces* with cas9-deaminase fusions. *J. BioRxiv* 2019, 630137. doi:10.1101/630137

Zhu, X. N., Zhao, D. D., Qiu, H. N., Fan, F. Y., Man, S. L., Bi, C. H., et al. (2017). The CRISPR/Cas9-facilitated multiplex pathway optimization (CFPO) technique and its application to improve the *Escherichia coli* xylose utilization pathway. *Metab. Eng.* 43, 37–45. doi:10.1016/j.ymben.2017.08.003

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Mini Review
PUBLISHED 05 September 2022
DOI 10.3389/fbioe.2022.920639

OPEN ACCESS

EDITED BY
Jean Marie François,
Institut Biotechnologique de Toulouse (INSA), France

REVIEWED BY
Danielle Maass,
Federal University of São Paulo, Brazil
Ross Carlson,
Montana State University, United States

*CORRESPONDENCE
Radhakrishnan Mahadevan,
krishna.mahadevan@utoronto.caca

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Synthetic Biology,
a section of the journal
Frontiers in Bioengineering and Biotechnology

RECEIVED 14 April 2022
ACCEPTED 08 August 2022
PUBLISHED 05 September 2022

CITATION
Chen J, Liu Y, Diep P and Mahadevan R (2022), Harnessing synthetic biology for sustainable biomining with Fe/S-oxidizing microbes.
Front. Bioeng. Biotechnol. 10:920639.
doi: 10.3389/fbioe.2022.920639

Harnessing synthetic biology for sustainable biomining with Fe/S-oxidizing microbes

Jinjin Chen[1†], Yilan Liu[1†], Patrick Diep[1] and Radhakrishnan Mahadevan[1,2*]

[1]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada, [2]Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

Biomining is a biotechnological approach where microorganisms are used to recover metals from ores and waste materials. While biomining applications are motivated by critical issues related to the climate crisis (*e.g.*, habitat destruction due to mine effluent pollution, metal supply chains, increasing demands for cleantech-critical metals), its drawbacks hinder its widespread commercial applications: lengthy processing times, low recovery, and metal selectivity. Advances in synthetic biology provide an opportunity to engineer iron/sulfur-oxidizing microbes to address these limitations. In this forum, we review recent progress in synthetic biology-enhanced biomining with iron/sulfur-oxidizing microbes and delineate future research avenues.

KEYWORDS

synthetic biology, Fe/S-oxidizing microbes, biomining, design-build-test-learn (DBTL) cycle, CRISPR

Introduction

Biomining is a technique that utilizes microorganisms to recover target metals from ores and waste materials, and potentially capture them for recovery and downstream purification. Chemolithotrophic iron/sulfur (Fe/S)-oxidizing microbes have attracted attention in biomining research since they thrive in extremely acidic metal-rich environments, produce a bioleaching lixiviant containing Fe^{3+} and H^+ by oxidizing ferrous iron and/or reduced inorganic sulfur compounds (RISCs) as an energy source, and fix CO_2 from the atmosphere as the carbon source (Chen et al., 2022a). Considerable efforts have been devoted to applying synthetic biology to improve the biomining performance of Fe/S-oxidizing microbes, and significant progress has been made in recent years. This mini review presents the up-to-date developments of synthetic biology in biomining with Fe/S-oxidizing microbes and provides perspectives on this research area.

Fe/S-oxidizing microbes in biomining processes

A typical biomining microbial community consists of autotrophic Fe/S-oxidizers, mixotrophs and heterotrophs (Li and Wen, 2021) (Figure 1). Neutrophilic Fe-oxidizers

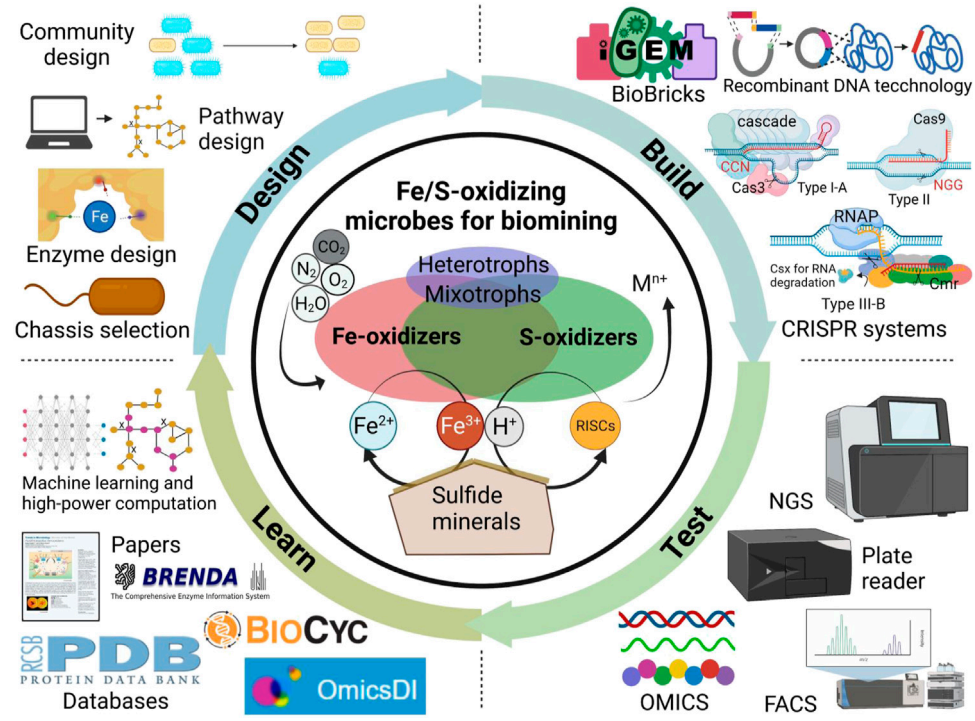


FIGURE 1
Harnessing synthetic biology for sustainable biomining with Fe/S-oxidizing microbes. Technology advances accelerate the design-build-test-learn cycle of synthetic biology-assisted biomining. NGS, next-generation sequencing; FACS, fluorescence-activated cell sorting, OMICS refers to genomics, proteomics, transcriptomics, metabolomics, and fluxomics.

(eg. *Rhizobiales* spp. and *Ralstonia* spp.) may play a role to kick start the mineral oxidation at circumneutral pH (Percak-Dennett et al., 2017). The acidophilic Fe-oxidizers (*e.g.*, *Leptospirillum* spp. and *Ferroplasma* spp.) obtain electrons from dissolved Fe^{2+} for energy generation, thus generating Fe^{3+} that can attack the mineral's lattice structure to release additional iron, target metals and sulfur compounds. However, S^0 or jarosite ($KFe_3(SO_4)_2(OH)_6$) can accumulate to form a passivation layer covering the mineral surface thus inhibiting further leaching. S-oxidizers (*e.g.*, *Acidianus* spp., some species in *Acidithiobacillus* and *Sulfolobus*) can oxidize sulfur compounds like S^0 to sulfuric acid. This process generates additional H^+ that assists Fe^{3+} in further degrading the mineral lattice structure to release the target metal. Combined, Fe-oxidizers and S-oxidizers maintain extremely acidic conditions with a high oxidation-redox potential (ORP) to leach target metals from the mineral without precipitating them. Moreover, mixotrophs and heterotrophs found in genera such as *Acidiphilium, Metallosphaera, Sulfobacillus* and *Sulfolobus* can also degrade organic compounds mitigating potential toxicity issues to autotrophic Fe/S-oxidizers.

Several species in the genera *Acidithiobacillus* are capable of oxidizing both Fe^{2+} and RISCs. In particular, *Acidithiobacillus ferrooxidans* is the most widely studied chassis for biomining applications. Mixed cultures of *A. ferrooxidans* with other Fe-oxidizing and S-oxidizing microbes, mixotrophs and heterotrophs have been shown to be effective for bioleaching (Wang et al., 2020). Hence, *A. ferrooxidans* with its metabolic versatility and ability to fix CO_2 is a prime target for synthetic biology efforts to enhance biomining kinetics, to produce value-added products from CO_2 and to confer advanced functionality like selective metal recovery. In addition, new Fe/S-oxidizing microbes found in extreme environments will also contribute valuable genetic diversity, such as the novel archaeal group found in Yellowstone National Park and *Zetaproteobacteria* from marine hydrothermal vents (Kozubal et al., 2013; McAllister et al., 2019).

Design-build-test-learn applied to engineering Fe/S-oxidizing microbes

A hallmark of synthetic biology is the application of principles outlined in the design-build-test-learn (DBTL) cycle commonly found in traditional engineering disciplines to the field of biology. This DBTL cycle integrates rational design, genetic engineering, results testing, and data analysis to learn and resolve bottlenecks in biomanufacturing. Such a DBTL cycle

could be specifically applied to engineering Fe/S-oxidizing microbes for biomining applications (Figure 1). For *Design*, computational modelling tools combined with biological databases (e.g., genome sequences, OMICS datasets) enable the rational design of enzymes, metabolic pathways, and communities of biomining microbes. A genome-scale metabolic model (GEM) of *A. ferrooxidans* (iMC507) is available for *in silico* analysis and design, and has been used to predict gene deletions for growth-coupled extracellular polymeric substance (EPS) production (Campodonico et al., 2016).

For *Build*, there are genetic engineering tools to engineer Fe/S-oxidizing microbes with desirable biomining features. Some conjugative/mobilizable plasmids and viruses have been reported in species of *Acidianus*, *Acidiphilium*, *Leptospirillum*, *Metallosphaera* and *Sulfobacillus* to allow for heterologous gene expression (Beard et al., 2021). However, significant advances have been made in *Acidithiobacillus* and archaea *Sulfolobus* since prior efforts were mostly devoted to studying these model organisms for acidophiles. For *Sulfolobus* spp., the endogenous type I-A and type III-B CRISPR/Cas systems have been harnessed for efficient targeted genome editing and gene silencing (Peng et al., 2017). Comparatively, genetic engineering of *Acidithiobacillus* spp. is time-consuming due to its generally slower growth rate, thus making it harder to engineer. Nonetheless, studies have shown it is still possible to use conventional genome editing methods like transposon-based gene interruption or suicide plasmid-based gene deletion (Chen et al., 2022a). Recently, Moya-Beltrán *et al.* investigated type IV CRISPR-Cas systems in the class *Acidithiobacillia*, which may enable genome editing like *Sulfolobus* spp. (Moya-Beltrán et al., 2021). Recently, there was a report of a successful gene knockdown in *A. ferrooxidans* using the CRISPR-dCas9 system derived from *S. pyogenes* (Yamada et al., 2022). At the same time, our group first achieved seamless genome editing using the CRISPR-Cas9 system in *Acidithiobacillus ferridurans*, which significantly accelerates the genetic engineering process (Chen et al., 2022b). However, additional studies are needed to apply CRISPR systems in *Acidithiobacillus* strains and other Fe/S-oxidizing microbes for biomining applications.

The *Test* and *Learn* phases of the DBTL involve validating and analyzing data generated from engineered strains for subsequent engineering. These components of the DBTL cycle are most underdeveloped for biomining microbes. The development of high-throughput technologies such as next-generation sequencing (NGS), automated liquid handling/plate reader workflows, fluorescence-activated cell sorting (FACS), and microfluidics have facilitated the screening of large strain libraries (Sarnaik et al., 2020). These big datasets combined with existing databases like OmicsDI, BioCyc, BRENDA, and PDB can enable machine learning methods to generate critical insights to inform design iterations. While for example high-throughput sequencing of 16S rRNA gene has been used in biomining studies to monitor microbial community dynamics, efforts are still needed in high-throughput engineering, culturing, testing, screening of Fe/S-oxidizing microbes with desirable biomining traits.

Advances of synthetic biology-enhanced biomining

While a comprehensive DBTL methodology is yet to be applied for engineering biomining, there are several examples where synthetic biology has been deployed in this context. Gumulya et al. (2018) reviewed the manipulation of genes responsible for acid tolerance, metal tolerance, osmotolerance, thermotolerance, Fe/S-oxidation, and carbon fixation to improve microbe's suitability for industrial use. In *Acidithiobacillus* strains, overexpression of *rus* and *cyc2* genes was found to improve Fe^{2+} oxidation activity, whereas overexpression of the quorum sensing (QS) operon *afeI-orf-afeR* improved sulfur oxidation rates and promoted EPS synthesis. Glutathione biosynthesis gene *gshB* overexpression significantly increased intracellular reactive oxygen species (ROS) levels and expanded the halotolerance of the cells. Other genes related to Fe/S-metabolisms (e.g., *fur*, *tetH*, *sdo1, sdo2, rsrR, rsrS,* and *sor*), metal resistance (e.g., *cueO*, *ars* and *mer*), c-di-GMP signaling (e.g., *dgc* and *pelD*) and quorum sensing (e.g., *aar*, *ado* and *act*) were also investigated (Jung et al., 2021).

A deeper understanding of Fe/S-oxidizing microbes' genes is nonetheless still needed for more detailed metabolic engineering towards desirable biomining properties, compared to the model chassis *Escherichia coli* and *Saccharomyces cerevisiae*. Schmitz et al. (2021) harnessed high-throughput genome editing and sequencing approaches in the study of heterotrophic bacterium *Gluconobacter oxydans* to produce organic acids for rare earth elements biomining. They created a library of single-gene transposon mutants in *G. oxydans* and found the bioleaching rates of rare earth elements increased up to 18% when phosphate-specific transport systems genes were disrupted. Therefore, high-throughput engineering, culture, sequencing, screening processes may facilitate faster development of engineered Fe/S-oxidizing microbes.

The outlook of synthetic biology-enhanced biomining applications

Biomining processes are commercially applied around the world in two broad categories: irrigation-type and stirred tank-type. Irrigation-type examples include heap bioleaching processes carried out by companies like Newmont Mining and Phelps Dodge, which are suitable for metal recovery from low-grade ores because these systems are easier and cheaper to construct, although the process can be slow with inefficient

recoveries (Petersen, 2016). In contrast, the more widely used processes are of stirred tank-type, such as BIOX™ technology from Metso-Outotec. The control of operating parameters contribute significantly to the fast and efficient metal recovery, but the operating cost is higher (Rawlings et al., 2003). Recently, Kremser et al. (2020) compared the potential of heap and stirred-tank bioreactors for metal recovery from shredder-light-fractions. Pure and co-culture of *A. ferrooxidans* and *L. ferrooxidans* were used for Cu, Zn and Ni recovery in batch and up-scale experiments, which highlights the potential for future commercial biomining applications with engineered Fe/S-oxidizing microbes. However, for large-scale applications, both heap and stirred-tank biomining provide an open, non-sterile environment. This open environment can lead to difficulties in optimizing a stable microbial consortia that can outcompete the native consortia and match the scale and complex nature of the ore feedstocks (Rawlings and Johnson, 2007). The design and construction of synthetic and mixed microbial consortia is an emerging area in synthetic biology, and may help to address these issues for better industrial biomining applications (Brune and Bayer, 2012; McCarty and Ledesma-Amaro, 2019).

Concluding remarks and future perspectives

Synthetic biology-enhanced biomining can be a sustainable, eco-friendly, and cost-effective technology with applications in leaching and recovering base metals, precious metals, and rare earth elements from a variety of feedstocks like low-grade ores, mine wastes, and electronic wastes. Though engineering Fe/S-oxidizing microbes is still at its earliest stage, we believe that the following developments can accelerate synthetic biology-enhanced biomining research:

(1) Fast and efficient genome editing methods and tools are needed for Fe/S-oxidizing microbes (i.e., CRISPR/Cas9 system). Genetic tools are available for *Sulfolobus* strains that are adapted to high temperature (80°C) and acidic conditions (pH < 3), highlighting their suitability for engineering (Lewis et al., 2021). We summarized the diverse genetic tools that are feasible for *Acidithiobacillus* species (Chen et al., 2022a). Different CRISPR systems have been successfully used for genome editing in *Sulfolobus* and *Acidithiobacillus* species. We believe genome editing methods for other Fe/S-oxidizing microbes can offer us a larger solution space for designing synthetic biomining consortia with desirable leaching properties.

(2) Development of high-throughput technologies to culture and screen of Fe/S-oxidizing microbes will accelerate the engineering of strains with more desirable industrial properties such as heavy metal tolerance, robustness under large bioreactor conditions. Example studies could include designing synthetic histidine kinases that can sense heavy-metals (Sarnaik et al., 2020), creating a whole-genome single-gene mutant library to enable screens for desired properties (Schmitz et al., 2021), or using CRISPRi system as an efficient platform for rapid identification of target genes (Wu et al., 2020).

(3) The leaching of metals from the mineral body is typically not a selective process as all metals contained in the mineral lattice are released into solution. Such a complex leachate complicates downstream processing. To mitigate this complexity, metal-sensing regulators (e.g., riboswitches) could be used to construct circuits that enable selective and sequential metal recovery by Fe/S-oxidizing microbes by regulating pathways responsible for specific metal uptake, storage, and efflux during biomining process (Diep et al., 2018).

(4) Engineering microbial communities using synthetic biology tools, such as CRISPR/Cas systems, riboswitches, quorum sensing or syntrophic exchanges (McCarty and Ledesma-Amaro, 2019) may promote synergistic effects among Fe-oxidizers, S-oxidizers and heterotrophs and further enhance biomining efficiency, especially in heap and stirred-tank biomining that carried out in the non-sterile environment.

(5) Rather than directly genetically engineering Fe/S-oxidizing bacteria, directed evolution may also be a good research direction since stakeholders are more likely to accept industrial bacteria with more natural origins. For example, the EvolvR system can help evolve industrial microbes with desired phenotypes, such as fast growth rate, high tolerance to complex metals, acidity, salinity, and more (Halperin et al., 2018).

Author contributions

JC, YL, and RM conceived and designed the structure of the minireview; JC and YL wrote the first draft; PD contributed to illustration and re-writing the manuscript; all authors ratified the submitted and revised versions.

Funding

This work was funded by Elements of Biomining Grant from the Province of Ontario through the ORF Research Excellence funding program, and the Postdoctoral fellowship of 2022-2023 Climate Positive Energy Rising Stars in Clean Energy from University of Toronto.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Beard, S., Ossandon, F. J., Rawlings, D. E., and Quatrini, R. (2021). The flexible genome of acidophilic prokaryotes. *Curr. Issues Mol. Biol.* 40, 231–266. doi:10.21775/cimb.040.231

Brune, K. D., and Bayer, T. (2012). Engineering microbial consortia to enhance biomining and bioremediation. *Front. Microbiol.* 3, 203. doi:10.3389/fmicb.2012.00203

Campodonico, M. A., Vaisman, D., Castro, J. F., Razmilic, V., Mercado, F., Andrews, B. A., et al. (2016). *Acidithiobacillus ferrooxidans*'s comprehensive model driven analysis of the electron transfer metabolism and synthetic strain design for biomining applications. *Metab. Eng. Commun.* 3, 84–96. doi:10.1016/j.meteno.2016.03.003

Chen, J., Liu, Y., Diep, P., and Mahadevan, R. (2022a). Genetic engineering of extremely acidophilic *Acidithiobacillus* species for biomining: Progress and perspectives. *J. Hazard. Mat.* 438, 129456. doi:10.1016/j.jhazmat.2022.129456

Chen, J., Liu, Y., and Mahadevan, R. (2022b). Genetic engineering of Acidithiobacillus ferridurans with CRISPR-Cas9/dCas9 systems. bioRxiv.

Diep, P., Mahadevan, R., and Yakunin, A. F. (2018). Heavy metal removal by bioaccumulation using genetically engineered microorganisms. *Front. Bioeng. Biotechnol.* 6, 157. doi:10.3389/fbioe.2018.00157

Gumulya, Y., Boxall, N. J., Khaleque, H. N., Santala, V., Carlson, R. P., and Kaksonen, A. H. (2018). In a quest for engineering acidophiles for biomining applications: Challenges and opportunities. *Genes* 9, 116. doi:10.3390/genes9020116

Halperin, S. O., Tou, C. J., Wong, E. B., Modavi, C., Schaffer, D. V., and Dueber, J. E. (2018). CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature* 560, 248–252. doi:10.1038/s41586-018-0384-8

Jung, H., Inaba, Y., and Banta, S. (2021). Genetic engineering of the acidophilic chemolithoautotroph *Acidithiobacillus ferrooxidans*. *Trends Biotechnol.* 40, 677–692. doi:10.1016/j.tibtech.2021.10.004

Kozubal, M. A., Romine, M., Jennings, R. D., Jay, Z. J., Tringe, S. G., Rusch, D. B., et al. (2013). Geoarchaeota: A new candidate phylum in the archaea from high-temperature acidic iron mats in Yellowstone national Park. *ISME J.* 7, 622–634. doi:10.1038/ismej.2012.132

Kremser, K., Thallner, S., Schoen, H., Weiss, S., Hemmelmair, C., Schnitzhofer, W., et al. (2020). Stirred-tank and heap-bioleaching of shredder-light-fractions (SLF) by acidophilic bacteria. *Hydrometallurgy* 193, 105315. doi:10.1016/j.hydromet.2020.105315

Lewis, A. M., Recalde, A., Bräsen, C., Counts, J. A., Nussbaum, P., Bost, J., et al. (2021). The biology of thermoacidophilic archaea from the order Sulfolobales. *FEMS Microbiol. Rev.* 45, fuaa063. fuaa063. doi:10.1093/femsre/fuaa063

Li, M., and Wen, J. (2021). Recent progress in the application of omics technologies in the study of bio-mining microorganisms from extreme environments. *Microb. Cell Fact.* 20, 178. doi:10.1186/s12934-021-01671-7

Mcallister, S. M., Moore, R. M., Gartman, A., Luther, G. W., Iii, Emerson, D., and Chan, C. S. (2019). The Fe (II)-oxidizing *Zetaproteobacteria*: Historical, ecological and genomic perspectives. *FEMS Microbiol. Ecol.* 95, fiz015. fiz015. doi:10.1093/femsec/fiz015

Mccarty, N. S., and Ledesma-Amaro, R. (2019). Synthetic biology tools to engineer microbial communities for biotechnology. *Trends Biotechnol.* 37, 181–197. doi:10.1016/j.tibtech.2018.11.002

Moya-Beltrán, A., Makarova, K. S., Acuña, L. G., Wolf, Y. I., Covarrubias, P. C., Shmakov, S. A., et al. (2021). Evolution of type IV CRISPR-cas systems: Insights from CRISPR loci in integrative conjugative elements of *Acidithiobacillia*. *CRISPR J.* 4, 656–672. doi:10.1089/crispr.2021.0051

Peng, N., Han, W., Li, Y., Liang, Y., and She, Q. (2017). Genetic technologies for extremely thermophilic microorganisms of Sulfolobus, the only genetically tractable genus of crenarchaea. *Sci. China Life Sci.* 60, 370–385. doi:10.1007/s11427-016-0355-8

Percak-Dennett, E., He, S., Converse, B., Konishi, H., Xu, H., Corcoran, A., et al. (2017). Microbial acceleration of aerobic pyrite oxidation at circumneutral pH. *Geobiology* 15, 690–703. doi:10.1111/gbi.12241

Petersen, J. (2016). Heap leaching as a key technology for recovery of values from low-grade ores–A brief overview. *Hydrometallurgy* 165, 206–212. doi:10.1016/j.hydromet.2015.09.001

Rawlings, D. E., Dew, D., and Du Plessis, C. (2003). Biomineralization of metal-containing ores and concentrates. *Trends Biotechnol.* 21, 38–44. doi:10.1016/s0167-7799(02)00004-5

Rawlings, D. E., and Johnson, D. B. (2007). The microbiology of biomining: Development and optimization of mineral-oxidizing microbial consortia. *Microbiology* 153, 315–324. doi:10.1099/mic.0.2006/001206-0

Sarnaik, A., Liu, A., Nielsen, D., and Varman, A. M. (2020). High-throughput screening for efficient microbial biotechnology. *Curr. Opin. Biotechnol.* 64, 141–150. doi:10.1016/j.copbio.2020.02.019

Schmitz, A. M., Pian, B., Medin, S., Reid, M. C., Wu, M., Gazel, E., et al. (2021). Generation of a *Gluconobacter oxydans* knockout collection for improved extraction of rare earth elements. *Nat. Commun.* 12, 6693–6711. doi:10.1038/s41467-021-27047-4

Wang, L., Yin, S., Wu, A., and Chen, W. (2020). Synergetic bioleaching of copper sulfides using mixed microorganisms and its community structure succession. *J. Clean. Prod.* 245, 118689. doi:10.1016/j.jclepro.2019.118689

Wu, J., Cheng, Z., Min, D., Cheng, L., He, R., Liu, D., et al. (2020). CRISPRi system as an efficient, simple platform for rapid identification of genes involved in pollutant transformation by *Aeromonas hydrophila*. *Environ. Sci. Technol.* 54, 3306–3315. doi:10.1021/acs.est.9b07191

Yamada, S., Suzuki, Y., Kouzuma, A., and Watanabe, K. (2022). Development of a CRISPR interference system for selective gene knockdown in *Acidithiobacillus ferrooxidans*. *J. Biosci. Bioeng.* 133, 105–109. doi:10.1016/j.jbiosc.2021.10.012

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Opinion
PUBLISHED 30 September 2022
DOI 10.3389/fbioe.2022.1029446

OPEN ACCESS

EDITED BY
Jean Marie François,
Institut Biotechnologique de Toulouse (INSA), France

REVIEWED BY
Allen Liu,
University of Michigan, United States

*CORRESPONDENCE
Pasquale Stano,
pasquale.stano@unisalento.it

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 27 August 2022
ACCEPTED 13 September 2022
PUBLISHED 30 September 2022

CITATION
Stano P (2022), A four-track perspective for bottom-up synthetic cells. *Front. Bioeng. Biotechnol.* 10:1029446. doi: 10.3389/fbioe.2022.1029446

A four-track perspective for bottom-up synthetic cells

Pasquale Stano*

Department of Biological and Environmental Sciences and Technologies (DiSTeBA), University of Salento, Lecce, Italy

KEYWORDS

artificial cells, synthetic cells, protocells, bottom-up synthetic biology, cell-free systems, sciences of the artificial, wetware artificial life

"Synthesizing life", 20 years later

A crucial move for the development of the bottom-up synthetic biology (SB) branch took place about 20 years ago, when Jack W. Szostak, David Bartel and Pier Luigi Luisi co-authored a *Nature* paper entitled "Synthesizing Life" (Szostak et al., 2001), which can be considered a sort of foundational paper for (or even the manifesto of) the modern approaches for constructing living artificial cells from scratch. Possibly as a sign of the *Zeitgeist*, the article was published almost simultaneously to two other foundational papers in SB (Elowitz et al., 2000; Gardner et al., 2000).

The very idea of synthesizing life—the Faustian dream of all times—is not new. Several (unsuccessful) attempts to build cell-like systems of minimal complexity fill the annals of science (Hanczyc, 2009). These attempts share a common anti-vitalistic viewpoint: synthesizing (cellular) life from scratch should be possible, and it would demonstrate that the biological phenomenology follows a "continuity principle" with respect to physics and chemistry. That is, life is an emergent property of some molecular systems characterized by a very peculiar type of structural and dynamical (self-) organization. However trivial and generally taken for granted by scientists, the emergence of life from inanimate matter has never been demonstrated experimentally and it is still one of the big targets in science.

What is, then, the remarkable and novel element that has been put forward in the "Synthesizing Life" article, and that can be considered as a foundational concept for bottom-up approaches in SB? The Authors actually focused on the hypothetical construction of primitive cell (protocell) models, made of catalytically active RNAs (ribozymes) (Bartel and Szostak, 1993; Eckland et al., 1995), encapsulated inside fatty acid vesicles (Hargreaves and Deamer, 1978; Bachman et al., 1992; Walde et al., 1994). The claim is that such structures would display minimal life-like behavior (reproductive and potentially evolutive) if the intravesicle ribozymes catalyze their own replication and the production of membrane molecules at the expenses of certain precursors available in the environment. The whole process would lead to a spontaneous growth-division of protocells in an allegedly primitive Earth scenario.

Leaving aside, for the moment, the mechanistic details and their plausibility, the fundamental and explicit message of that paper goes beyond the apparently narrow focus on the origin of life. To a closer inspection, in fact, the Authors put forward an operational methodology for the construction of chemical reacting systems that would show the difficult-to-define property of being alive just by fulfilling a specific structural and

dynamic organization. The latter is described by features as: (1) self-bounding, to let the system autonomously constitutes a unity, distinct from the surroundings (topological closure); (2) ability to completely specify, by molecularly embodied internal rules and operations, the construction and the degradation of all components, without the need of being heterodirected or instructed (organizational closure); (3) ability of exchanging of matter and energy with the surroundings, keeping itself thermodynamically open and continuously functioning out of the equilibrium; (4) possibility to adapt to the external conditions by plastic modification the network of their internal processes while remaining organizationally closed; (5) and possible evolution by the principles of diversification and selection. Such an organization is called *autopoietic* (self-constructing), and it was identified by H. Maturana and F. Varela in the 1970s (Varela et al., 1974). Autopoiesis thus becomes a convenient and elegant theoretical framework to guide the variegate experimental efforts to fabricate synthetic (artificial) cells (SCs/ACs), especially with respect to studies oriented at the origins and emergence of life. Other systemic theories are available, such as the chemoton theory (Ganti, 1975, 2003) and others (Cornish-Bowden and Cárdenas, 2020), but autopoiesis stands out, in our opinion, for its broader and deeper implications.[1]

To date, SC research has definitely put down roots. The number of scholars working in the field is constantly increasing, as well as the number of published articles, often placed in renowned journals. There is an enthusiastic involvement of researchers coming from different backgrounds. New centers, networks, consortia, and initiatives are currently driving the field forward (Schwille et al., 2018; Frischmon et al., 2021; Staufer et al., 2021). Importantly, the construction of *living* SC (considered the "Holy Grail" of the field), the understanding of the non-life to life transition, and the determination of the minimal complexity of living beings, have been flanked by other relevant goals. Pragmatic approaches aiming at the construction of *non-living* SCs are interesting too, having advantages such as easier realization and a potentially early use for basic understanding of physiological processed and in biotechnology. Whether or not the goal of SC research refers to living or non-living SCs, such a new "technology" represents a genuine novelty in modern science and constitutes an original and promising platform for investigating theoretical issues as well (Damiano and Stano, 2020; Magarini and Stano, 2021; Stano, 2022).

The construction of different types of cell-like systems with non-trivial complexity is now within the experimental reach. Since several excellent reviews on technical advancements in SCs research have been published recently (Cho and Lu, 2020; Gaut and Adamala, 2021; Ivanov et al., 2021; Lussier et al., 2021), the discussion below will focus, we hope, on less explored subjects, and aims at inspiring future investigation scenarios. Deepening the knowledge and broadening the range of interest on SCs may proceed, according to our viewpoint, according to four tracks: the theoretical, scientific, technological and educational ones.

The theoretical track: What SCs actually are

Irrespective of the scientific goal behind the construction of SCs (or protocells), the very fact that a cell-like system can be built in controlled laboratory conditions elicits a question about our theoretical/epistemological understanding of what they actually are. Several interesting analyses have discussed the place of SB and SCs in current scientific understanding of life, from historical and epistemological perspectives (Deplazes and Huppenbauer, 2009; Morange, 2009; Moya et al., 2009; Deplazes-Zemp, 2016; Zwart, 2019). The question is not only whether or not SCs can ever be "alive" *hic et nunc*, but whether or not the category of synthetic life is the same of natural life. Moreover, the very rich landscape of approaches, materials, systems that are currently explored makes difficult to define "what SCs are", and what is the role given to structure, organization, or function in order to evaluate and compare SCs. Attempting to address these and other questions is *per se* a stimulating intellectual journey.

An intriguing interpretation considers SB (and in particular SC research) as the wetware branch of the "*Sciences of the Artificial*" (Cordeschi, 2002; Damiano et al., 2011; Damiano and Stano, 2018). It means that SCs shares with robotics (the hardware branch) and artificial intelligence (AI, the software branch) a common set of scopes, perspectives, and theoretical analyses. These three approaches aim at constructing models that reproduce the biological phenomenology and/or organization, often following the "understanding by building" strategy (Kaneko, 2006). Classical as well as newer concepts related to information and communication theories, computation, self-organization, emergence and complexity can be explored in an extraordinary innovative way by means of SB. These concepts, when properly developed and understood in the SB molecular domain, become new tools for facing long discussed issues like machine/organism dichotomy (Deplazes and Huppenbauer, 2009; Nicholson, 2013) and the related computer/mind one (of course, here we mean minimal

1 Autopoiesis provides a description of "what life is" entirely in terms of causally concatenated relations of processes and components that (*i*) recursively generate themselves (processes and components) and (*ii*) determine a physical unity, which is physically distinct (and distinguishable) from the environment. Moreover, the autopoietic dynamics self-regulate to compensate the perturbations exerted by the environment, in so far as the induced change can be subordinated to the maintenance of the autopoietic organization (i.e., they are "autonomous"). The allowed perturbations constitute the "cognitive domain" of the unity. Concepts as autonomy, cognition, phenomenology, identity, coherence, structural coupling, plasticity, mind-likeness have roots in the autopoietic dynamics. Autopoiesis contributes to constructivism. Interested readers can further refer to (Varela et al., 1974; Varela, 1979; Maturana and Varela, 1980; Luisi, 2003; Ruiz-Mirazo and Moreno, 2004).

organisms with mind-like cognitive features). For example, the functioning of currently studied SCs can be simulated by an algorithm: their behavior is Turing computable. On the other hand, Turing computability of autopoietic (and thus living) systems has been questioned (Letelier et al., 2003; McMullin, 2004). Theoretical investigations related to current and future SCs are quite interesting indeed. If modern SB tools were available to early cyberneticians, the latter would have been certainly interested in them (Wiener et al., 1943; MacKay, 1969).

Let us focus here on the possible contribution of SC research to cognitive sciences, just to make an example. Embodied cognition is one of the three main branches of cognitive sciences, together with classical and connectionistic approaches (Dawson, 2013); it emphasizes the causal perception-action loop that a cognitive agent realizes by interacting with its body (and through its body) in an environment, where it is *situated* (Varela et al., 1992; Shapiro, 2011). This is made possible by sensorimotor capacities embedded in the agent body. As we have recently argued in a dedicated article, SB provides an excellent platform for investigations on "chemical embodied AI" via the development of properly designed wetware models (i.e., SCs) (Damiano and Stano, 2021). For example, in order to model minimal cognition, SCs should cope with environmental perturbation by adaptive mechanisms of self-regulation. It has been proposed to graft *chemical neural networks* in SCs (e.g., based on protein phosphorylation; Gentili and Stano, 2022) that respond to physico-chemical stimuli coming from the surrounding. However, to be adaptive, such networks must be able to self-regulation, and this is not at all trivial to achieve. Nevertheless, the latter seems an easier goal if compared to the rather challenging "whole-SC" autopoiesis (Damiano and Stano, 2018; see also Di Paolo, 2003; Kiverstein et al., 2022).

The scientific track: Integrate functions to reach higher complexity

This is, perhaps, the most obvious and important direction to look at. Imminent developments in SC research and technology must necessarily face the challenge of constructing systems with higher degree of organization and complexity. In this respect, the *integration* of the several different "modules" available so far in more complex SCs becomes a crucial milestone. A rich and ever increasing repertoire of functional "modules" for SCs operations have been developed in isolated way (e.g., protein synthesis, growth-division, DNA duplication, sending-receiving signals, *etc.*). The integration of these modules can be additive or synergic. For example, constructing SCs made of several "orthogonal" or "insulated" modules would correspond to an additive (linear) increase of complexity, while the combination of interrelated and causally dependent modules would bring about SCs of higher complexity, especially when self-regulatory properties emerge, because the embedded functions are more difficult to disentangle (higher "wholeness"). The first approach leads to an engineered system that can be decomposed into blocks, resembling top-down designed machine mechanisms; the second approach appears more bio-inspired as it points to interwoven processes and organism-like organization.

To face the difficulty of achieving high degrees of integration, an evolutionary approach has been proposed (Abil and Danelon, 2020). Directed evolution strategies should be considered as well (Sakatani et al., 2018; Okauchi and Ichihashi, 2021), especially when connected to adaptive responses. From the *architectural* viewpoint, complexification can be achieved *via* multiple compartmentalization. The latter can be hierarchical, i.e., according to a nested design (Altamura et al., 2021), or referred to 2D or 3D tissue-like systems (Bayley et al., 2019; Dupin et al., 2022): in both cases the behavior of the resulting "whole" will depend on the number, type, and function of constituent compartments.

The technological track: Looking for practical applications

As mentioned, most of the research on the construction of cell-like systems generally refers to basic scientific questions. However, SC technology is so genuinely innovative that can provide more, and demonstrate its practical utility. A decisive forward leap must come from considering SCs as a biotechnological platform. What are the practical uses of SCs? Who would produce or buy SCs, and why? These questions are often asked when SC research is presented to applied science-oriented audience, and require urgent answers.

The well-established liposome technology for drug delivery and the recent introduction of anti-SARS-CoV-2 vaccines based on RNA-lipid nanoparticles suggest a possible role of SCs as a kind of "smart" drug delivery (or drug producing) agents. The idea of using *ad hoc* designed enzyme-filled particles for enzyme replacement therapy, for instance, is not new at all (Chang, 1972). More recently, LeDuc and collaborators lucidly illustrated a scenario that resonates with SC philosophy (LeDuc et al., 2007). The advancements made on SC communicative properties (Lentini et al., 2017) let us imagine SCs that perceive their environment, and behave in programmable way in biological surroundings (Sato et al., 2022). The pioneer investigations on SCs producing a cancer-killing toxin (Krinsky et al., 2018), or on bacteria-killing SCs that operate upon a bacterial stimulus (Ding et al., 2018) provide a couple of illustrative examples. A realistic discussion about these developments should include however a consideration: the recent trend in SC research focuses on large structures (tens of micrometers), while therapeutic particles planned to be used for systemic administration must be rather small (typically <200 nm). The construction of sophisticated cell-like systems with such small size has been rarely reported (Pereira de Souza et al., 2009; Pols et al., 2019).

Further (and possibly nearer) applications can be devised when SC-like systems are conceived as tools for biotechnological research, exploiting the superior interfacing features between SC and other biological entities, and the possibility of designing SCs with a programmable behavior. For example, SCs could (i) mimic biological cells in viral research; (ii) host membrane sensors and/or reconstituted internal processes which are the target of drug action, to screen drug libraries; (iii) be hybridized with exosomes to complement and/or tailor their properties; (iv) be immobilized in form of biochip in order to respond in complex cell-like manner to several effectors, e.g., to run sophisticated tests; (v) be engineered as virus-like particles for treating cellular cultures or for special transfections; (vi) constitute—together with biological cells—hybrid organoids, or other sort of organized 2D/3D structures, or gel-embedded ensembles.

The educational track: SCs as a learning topic in "system thinking" programs

Whether or not SCs are designed as primitive cell models, or as non-living biotechnological tools, or as artificial autopoietic systems, it is evident that the very practice of their fabrication must embrace a *systemic* perspective. Systems are those entities or wholes, made of distinct parts, where the relations between the parts count as much as, if not more of, the parts themselves. Static and dynamic orders, patterns, and qualities become central to understanding biological phenomenology and complexity (Capra, 1996). In the case of SCs, the systemic perspective include both structure and organization, in the sense that SC properties, behavior, and features depend on how their components are assembled as a physical unity in space (e.g., due to containment in self-bounding compartments), and on the relations undergoing between the components (e.g., the *in situ* produced α-hemolysin chains self-assemble as a heptameric pore on the membrane, allowing small molecules enter or leave the SC lumen (Noireaux and Libchaber, 2004)). A systemic perspective is required for understanding, designing, constructing systems of all types.

While biology students are relatively well acquainted with systemic thinking (e.g., thanks to biochemistry, physiology, and ecology courses), it is not uncommon that students of other disciplines are less familiar with subjects as feedback, homeostasis, autonomy, compartmentation, multiple levels of organization, emergent phenomena, and circular organization[2]. The theory and the practice of SC construction is a convenient and valuable topic for courses on system thinking, as it can provide an opportunity to introduce systemic concepts at any educational level.

Another fecund intersection comes in mind, especially for chemistry students, when we consider the area of *systems chemistry* (Ruiz-Mirazo et al., 2014; Ashkenasy et al., 2017). The focus of systems chemistry goes beyond the mere building of chemical structures, and points to design chemical processes and systems that display features as autocatalysis, self-regulation, reaction-diffusion dynamics and oscillations, out-of-equilibrium dynamics, often exploiting the advantages of micro-compartmentalization. SCs are *de facto* major targets not only for SB, but for systems chemistry too.

Concluding remarks

This contribution aims at addressing the call made in the Research Topic "Insights in Synthetic Biology 2021*: Novel Developments, Current Challenges, and Future Perspectives*", that solicited forward-looking contributions describing the future challenges in SB. In particular, the subject of bottom-up SCs has been presented, highlighting its position in SB and its scientific relevance. The four "tracks" described above mirror the interests of the author and do not claim to be exhaustive.

Author contributions

PS conceived and wrote the article.

Acknowledgments

The Author thanks Luisa Damiano (Università IULM, Milan, Italy) for discussions about autopoiesis, autonomy, and the "Sciences of the Artificial".

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

2 By "circular organization" we mean that the system dynamics is constituted by concatenated processes which form, as a whole, a network of causally codependent processes. Such a dynamic organization is constantly regenerated owing to the existence of all processes, as in the case of cellular metabolism.

References

Abil, Z., and Danelon, C. (2020). Roadmap to building a cell: An evolutionary approach. *Front. Bioeng. Biotechnol.* 8, 927. doi:10.3389/fbioe.2020.00927

Altamura, E., Albanese, P., Mavelli, F., and Stano, P. (2021). The rise of the nested multicompartment model in synthetic cell research. *Front. Mol. Biosci.* 8, 750576. doi:10.3389/fmolb.2021.750576

Ashkenasy, G., Hermans, T. M., Otto, S., and Taylor, A. F. (2017). Systems chemistry. *Chem. Soc. Rev.* 46, 2543–2554. doi:10.1039/C7CS00117G

Bachmann, P., Luisi, P., and Lang, J. (1992). Autocatalytic self-replicating micelles as models for prebiotic structures. *Nature* 357, 57–59. doi:10.1038/357057a0

Bartel, D. P., and Szostak, J. W. (1993). Isolation of new ribozymes from a large pool of random sequences [see comment]. *Science* 261, 1411–1418. doi:10.1126/science.7690155

Bayley, H., Cazimoglu, I., and Hoskin, C. E. G. (2019). Synthetic tissues. *Emerg. Top. Life Sci.* 3, 615–622. doi:10.1042/ETLS20190120

Capra, F. (1996). *The web of life: A new scientific understanding of living systems*. New York: Anchor Books.

Chang, T. M. S. (1972). *Artificial cells*. Springfield, IL: Charles C Thomas.

Cho, E., and Lu, Y. (2020). Compartmentalizing cell-free systems: Toward creating life-like artificial cells and beyond. *ACS Synth. Biol.* 9, 2881–2901. doi:10.1021/acssynbio.0c00433

Cordeschi, R. (2002). *The discovery of the artificial. Behavior, mind and machines before and beyond cybernetics*. Netherlands: Springer.

Cornish-Bowden, A., and Cárdenas, M. L. (2020). Contrasting theories of life: Historical context, current theories. In search of an ideal theory. *Biosystems* 188, 104063. doi:10.1016/j.biosystems.2019.104063

Damiano, L., Hiolle, A., and Cañamero, L. (2011). "Grounding synthetic knowledge," in *Advances in artificial life, ECAL 2011*. Editors T. Lenaerts, M. Giacobini, H. Bersini, P. Bourgine, M. Dorigo, and R. Doursat. (Cambridge, MA: MIT Press), 200–207.

Damiano, L., and Stano, P. (2021). A wetware embodied AI? Towards an autopoietic organizational approach grounded in synthetic biology. *Front. Bioeng. Biotechnol.* 9, 724023. doi:10.3389/fbioe.2021.724023

Damiano, L., and Stano, P. (2020). On the "life-likeness" of synthetic cells. *Front. Bioeng. Biotechnol.* 8, 953. doi:10.3389/fbioe.2020.00953

Damiano, L., and Stano, P. (2018). Synthetic Biology and Artificial Intelligence. Grounding a cross-disciplinary approach to the synthetic exploration of (embodied) cognition. *Compl. Syst.* 27, 199–228. doi:10.25088/ComplexSystems.27.3.199

Dawson, M. R. W. (2013). *Foundations of cognitive science*. Edmonton, AB (Canada): Athabasca University Press.

Deplazes, A., and Huppenbauer, M. (2009). Synthetic organisms and living machines : Positioning the products of synthetic biology at the borderline between living and non-living matter. *Syst. Synth. Biol.* 3, 55–63. doi:10.1007/s11693-009-9029-4

Deplazes-Zemp, A. (2016). Artificial cell research as a field that connects chemical, biological and philosophical questions. *Chim. (Aarau)* 70, 443–448. doi:10.2533/chimia.2016.443

Di Paolo, E. A. (2003). "Organismically-inspired robotics: Homeostatic adaptation and teleology beyond the closed sensorimotor loop," in *Dynamical systems approach to embodiment and sociality*. Adelaide, Australia: Advanced knowledge international). Editors K. Murase and T. Asakura, 19–42.

Ding, Y., Contreras-Llano, L. E., Morris, E., Mao, M., and Tan, C. (2018). Minimizing context dependency of gene networks using artificial cells. *ACS Appl. Mat. Interfaces* 10, 30137–30146. doi:10.1021/acsami.8b10029

Dupin, A., Aufinger, L., Styazhkin, I., Rothfischer, F., Kaufmann, B. K., Schwarz, S., et al. (2022). Synthetic cell-based materials extract positional information from morphogen gradients. *Sci. Adv.* 8, eabl9228. doi:10.1126/sciadv.abl9228

Ekland, E. H., Szostak, J. W., and Bartel, D. P. (1995). Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science* 269, 364–370. doi:10.1126/science.7618102

Elowitz, M. B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335–338. doi:10.1038/35002125

Frischmon, C., Sorenson, C., Winikoff, M., and Adamala, K. P. (2021). Build-a-Cell: Engineering a synthetic cell community. *Life* 11, 1176. doi:10.3390/life11111176

Gánti, T. (1975). Organization of chemical reactions into dividing and metabolizing units: The chemotons. *BioSystems* 7, 15–21. doi:10.1016/0303-2647(75)90038-6

Gánti, T. (2003). *The principles of life*. New York: Oxford University Press.

Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature*. New York: New York: Oxford University Press 403, 339–342. doi:10.1038/35002131

Gaut, N. J., and Adamala, K. P. (2021). Reconstituting natural cell elements in synthetic cells. *Adv. Biol. (Weinh).* 5, e2000188. doi:10.1002/adbi.202000188

Gentili, P. L., and Stano, P. (2022). Chemical neural networks inside synthetic cells? A proposal for their realization and modeling. *Front. Bioeng. Biotechnol.* 10, 927110. doi:10.3389/fbioe.2022.927110

Hanczyc, M. M. (2009). The Early History of Protocells – the search for the recipe of life. *Protocells: Bridging Nonliving and Living Matter*. Cambridge MA: MIT Press, 3–18.

Hargreaves, W. R., and Deamer, D. W. (1978). Liposomes from ionic, single-chain amphiphiles. *Biochemistry* 17, 3759–3768. doi:10.1021/bi00611a014

Ivanov, I., Castellanos, S. L., Balasbas, S., Otrin, L., Marušič, N., Vidaković-Koch, T., et al. (2021). Bottom-up synthesis of artificial cells: Recent highlights and future challenges. *Annu. Rev. Chem. Biomol. Eng.* 12, 287–308. doi:10.1146/annurev-chembioeng-092220-085918

Kaneko, K. (2006). *Life: An introduction to complex systems biology*. Heidelberg: Springer.

Kiverstein, J., Kirchhoff, M. D., and Froese, T. (2022). The problem of meaning: The free energy principle and artificial agency. *Front. Neurorobot.* 16, 844773. doi:10.3389/fnbot.2022.844773

Krinsky, N., Kaduri, M., Zinger, A., Shainsky-Roitman, J., Goldfeder, M., Benhar, I., et al. (2018). Synthetic cells synthesize therapeutic proteins inside tumors. *Adv. Healthc. Mat.* 7, e1701163. doi:10.1002/adhm.201701163

Leduc, P. R., Wong, M. S., Ferreira, P. M., Groff, R. E., Haslinger, K., Koonce, M. P., et al. (2007). Towards an *in vivo* biologically inspired nanofactory. *Nat. Nanotechnol.* 2, 3–7. doi:10.1038/nnano.2006.180

Lentini, R., Martín, N. Y., Forlin, M., Belmonte, L., Fontana, J., Cornella, M., et al. (2017). Two-way chemical communication between artificial and natural cells. *ACS Cent. Sci.* 3, 117–123. doi:10.1021/acscentsci.6b00330

Letelier, J. C., Marín, G., and Mpodozis, J. (2003). Autopoietic and (M, R) systems. *J. Theor. Biol.* 222, 261–272. doi:10.1016/s0022-5193(03)00034-1

Luisi, P. L. (2003). Autopoiesis: A review and a reappraisal. *Naturwissenschaften* 90, 49–59. doi:10.1007/s00114-002-0389-9

Lussier, F., Staufer, O., Platzman, I., and Spatz, J. P. (2021). Can bottom-up synthetic biology generate advanced drug-delivery systems? *Trends Biotechnol.* 39, 445–459. doi:10.1016/j.tibtech.2020.08.002

MacKay, D. M. (1969). *Information, mechanism and meaning*. Cambridge, MA: MIT Press.

Magarini, M., and Stano, P. (2021). Synthetic cells engaged in molecular communication: An opportunity for modelling shannon- and semantic-information in the chemical domain. *Front. Comms. Net.* 2, 48. doi:10.3389/frcmn.2021.724597

Maturana, H. R., and Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht: D. Reidel Publishing Company.

McMullin, B. (2004). Thirty years of computational autopoiesis: A review. *Artif. Life* 10, 277–295. doi:10.1162/1064546041255548

Morange, M. (2009). A new revolution? The place of systems biology and synthetic biology in the history of biology. *EMBO Rep.* 10 (1), S50–S53. doi:10.1038/embor.2009.156

Moya, A., Krasnogor, N., Peretó, J., and Latorre, A. (2009). Goethe's dream. Challenges and opportunities for synthetic biology. *EMBO Rep.* 10 (1), S28–S32. doi:10.1038/embor.2009.120

Nicholson, D. J. (2013). Organisms ≠ machines. *Stud. Hist. Philosophy Sci. Part C Stud. Hist. Philosophy Biol. Biomed. Sci.* 44, 669–678. doi:10.1016/j.shpsc.2013.05.014

Noireaux, V., and Libchaber, A. (2004). A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci. U. S. A.* 101, 17669–17674. doi:10.1073/pnas.0408236101

Okauchi, H., and Ichihashi, N. (2021). Continuous cell-free replication and evolution of artificial genomic DNA in a compartmentalized gene expression system. *ACS Synth. Biol.* 10, 3507–3517. doi:10.1021/acssynbio.1c00430

Pereira de Souza, T., Stano, P., and Luisi, P. L. (2009). The minimal size of liposome-based model cells brings about a remarkably enhanced entrapment and protein synthesis. *ChemBioChem* 10, 1056–1063. doi:10.1002/cbic.200800810

Pols, T., Sikkema, H. R., Gaastra, B. F., Frallicciardi, J., Śmigiel, W. M., Singh, S., et al. (2019). A synthetic metabolic network for physicochemical homeostasis. *Nat. Commun.* 10, 4239. doi:10.1038/s41467-019-12287-2

Rosenblueth, A., Wiener, N., and Bigelow, J. (1943). Behavior, purpose and teleology. *Philos. Sci.* 10, 18–24. doi:10.1086/286788

Ruiz-Mirazo, K., Briones, C., and de la Escosura, A. (2014). Prebiotic systems chemistry: New perspectives for the origins of life. *Chem. Rev.* 114, 285–366. doi:10.1021/cr2004844

Ruiz-Mirazo, K., and Moreno, A. (2004). Basic autonomy as a fundamental step in the synthesis of life. *Artif. Life* 10, 235–259. doi:10.1162/1064546041255584

Sakatani, Y., Yomo, T., and Ichihashi, N. (2018). Self-replication of circular DNA by a self-encoded DNA polymerase through rolling-circle replication and recombination. *Sci. Rep.* 8, 13089. doi:10.1038/s41598-018-31585-1

Sato, W., Zajkowski, T., Moser, F., and Adamala, K. P. (2022). Synthetic cells in biomedical applications. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* 14, e1761. doi:10.1002/wnan.1761

Schwille, P., Spatz, J., Landfester, K., Bodenschatz, E., Herminghaus, S., Sourjik, V., et al. (2018). MaxSynBio: Avenues towards creating cells from the bottom up. *Angew. Chem. Int. Ed.* 57, 13382–13392. doi:10.1002/anie.201802288

Shapiro, L. (2011). *Embodied cognition*. New York: Routledge.

Stano, P. (2022). Exploring information and communication theories for synthetic cell research. *Front. Bioeng. Biotechnol.* 10, 927156. doi:10.3389/fbioe.2022.927156

Staufer, O., De Lora, J. A., Bailoni, E., Bazrafshan, A., Benk, A. S., Jahnke, K., et al. (2021). Building a community to engineer synthetic cells and organelles from the bottom-up. *Elife* 10, e73556. doi:10.7554/eLife.73556

Szostak, J. W., Bartel, D. P., and Luisi, P. L. (2001). Synthesizing life. *Nature* 409, 387–390. doi:10.1038/35053176

Varela, F. J., Maturana, H. R., and Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems* 5, 187–196. doi:10.1016/0303-2647(74)90031-8

Varela, F. J. (1979). *Principles of biological autonomy*. New York: Elsevier/North Holland.

Varela, F. J., Thompson, E. T., and Rosch, E. (1992). *The embodied mind: Cognitive science and human experience*. Revised Edition. Cambridge, MA: MIT Press.

Walde, P., Goto, A., Monnard, P., Wessicken, M., and Luisi, P. (1994). Oparin's reactions revisited: Enzymic synthesis of poly(adenylic acid) in micelles and self-reproducing vesicles. *J. Am. Chem. Soc.* 116, 7541–7547. doi:10.1021/ja00096a010

Zwart, H. (2019). From primal scenes to synthetic cells. *Elife* 8, e46518. doi:10.7554/eLife.46518

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 10 November 2022
DOI 10.3389/fbioe.2022.1046723

OPEN ACCESS

EDITED BY
Jean Marie François,
Institut Biotechnologique de Toulouse (INSA), France

REVIEWED BY
Georgios Efthimiou,
University of Hull, United Kingdom
Adison Choonkit Wong,
Singapore Institute of Technology, Singapore

*CORRESPONDENCE
Mads Kærn,
mkaern@uottawa.ca

SPECIALTY SECTION
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 17 September 2022
ACCEPTED 25 October 2022
PUBLISHED 10 November 2022

CITATION
Gill H, Ahsan M, Khalil Y, Feng V, Pearce J, Sharma T, Radwan M, Boucinha A and Kærn M (2022), The BioExperience Research and Entrepreneurship Challenge: An iGEM-inspired applied research program for BIOSTEM talent and skills development. *Front. Bioeng. Biotechnol.* 10:1046723. doi: 10.3389/fbioe.2022.1046723

The BioExperience Research and Entrepreneurship Challenge: An iGEM-inspired applied research program for BIOSTEM talent and skills development

Hertek Gill[1,2,3], Mahdi Ahsan[1,2,3], Yara Khalil[1,2,4], Victoria Feng[1,2,5], Jessie Pearce[1,2,4], Tarasha Sharma[1,2,6], Mohamad Radwan[1,2,4,7], Austin Boucinha[8,9] and Mads Kærn[1,2,10]*

[1]Ottawa Institute of Systems Biology, University of Ottawa, Ottawa, ON, Canada, [2]Department of Cellular and Molecular Medicine, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada, [3]Translational and Molecular Medicine Program, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada, [4]Biochemistry Program, Faculty of Science, University of Ottawa, Ottawa, ON, Canada, [5]Biomedical Science Program, Faculty of Science, University of Ottawa, Ottawa, ON, Canada, [6]Biopharmaceutical Sciences Program, Faculty of Science, University of Ottawa, Ottawa, ON, Canada, [7]Chemical Engineering Program, Faculty of Engineering, University of Ottawa, Ottawa, ON, Canada, [8]Canadian Synthetic Biology Educational Research Group (CSBERG), Centre for Applied Bioscience and Bioengineering Research (BioZone), University of Toronto, Toronto, ON, Canada, [9]Dufferin-Peel Catholic District School Board (DPCDSB), Mississauga, ON, Canada, [10]BioGroupe Canada Inc., Ottawa, ON, Canada

Post-secondary education is falling behind in delivering the talent and skills development needed to support the growth of biology-based economies and the demands of professional and research-based graduate degree programs. Here, we describe an applied research program, the BioExperience Research and Entrepreneurship Challenge, launched in May 2020 to mitigate the impact of the COVID-19 pandemic on undergraduate experiential learning programs at the University of Ottawa, Ontario, Canada. The program provided undergraduates with meaningful talent and skills development opportunities by implementing a student-centred, project-based learning approach inspired by the International Genetically Engineered Machine (iGEM) competition. We present evidence from participant exit surveys suggesting that the program delivers a high-quality learning environment and improves learning outcomes compared to traditional work-integrated learning. Notably, 84% of respondents reported an excellent or exceptional learning experience and significant or profound improvements in skills, such as leadership (72% of respondents), problem-solving (42% of respondents) and research preparedness (52%) that are difficult to develop in conventional academic programs. Remarkably, 60% of respondents report that the job-readiness training provided by the program is better or much better than traditional work-integrated learning. Our study demonstrates that a cost-effective and scalable alternative to the iGEM competition can improve talent and skills development in BIOSTEM fields.

KEYWORDS

experiential and work-integrated learning, engineered and synthetic biology, talent, skills development, undergraduate applied research

Introduction

The practical application of bioscience, biotechnology, bioengineering and biomedicine (BIOSTEM) research discoveries have had innumerable socio-economic benefits (Ganguly et al., 2014). This is evident by the rapid development of safe and effective RNA-based vaccines against the SARS-CoV2 virus (Dodd et al., 2021) and of CRISPR-based genetic engineering technology, which has the potential to cure serious human diseases (Jinek et al., 2012; Williams, 2014; Chan et al., 2021), facilitate equitable global economic development (Chui et al., 2020; Lange et al., 2021), reduce food insecurity (Shelton et al., 2020), replace unsustainable manufacturing and resource extraction practices (Kumar and Kumar, 2017; Degli Esposti et al., 2021), and mitigate the impact of environmental degradation and climate change (El Enshasy et al., 2020).

A highly educated workforce is essential to derive socio-economic benefits from technological advancements, and public post-secondary institutions have a social responsibility to ensure that this workforce can support economic development (Kirby, 2007). Unfortunately, post-secondary institutions are not adequately providing BIOSTEM graduates with the skills and competencies they need for successful careers in the bioeconomy (Wetzel et al., 2006; BioTalent Canada, 2021).

Several reports have documented deficiencies in post-secondary talent and skills over the past decade. A study from 2012 by the European Commission highlights insufficient skills development and a lack of industry collaboration as significant barriers to the European bioeconomy (New Skills for a European Bioeconomy-Conference Report, 2012). A 2013 survey of US college graduates revealed that while the majority believe they are well-prepared to enter the workforce, less than half of employers agreed (Mourshed et al., 2013). More recent studies by BioTalent Canada have revealed that many recent graduates lack skills and abilities that employers in the bioeconomy value the most, including non-technical essential skills such as problem-solving, critical thinking, communication, and collaboration skills (BioTalent Canada, 2021).

The expansion of work-integrated learning (WIL) has been proposed as a solution to address a growing labour shortage in the bioeconomy (Gamble et al., 2010). This type of learning can assist students by complementing their academic and technical skills with non-technical skills, such as collaboration, communication, and intrapersonal skills that are in high demand among employers but are challenging to develop in a classroom setting (Jackson, 2014; Edwards et al., 2015). WIL allows students to gain experiences relevant to their field of study through learning activities by incorporating experiences and practices from a professional setting (Billett, 2009; Sattler and Peters, 2013). For example, undergraduates pursuing research-oriented career paths often seek studentships in academic labs, while those pursuing careers in business seek corporate internships. Newer forms of WIL include community-based research and independent molecular biology (Cameron and Rexe, 2022), as well as a diverse range of extracurricular and community-organized competitions, hackathons, incubators, and accelerators, to meet increased demand for innovation and entrepreneurship skills development (de Villiers Scheepers et al., 2018).

Competition-based learning (CBL) is an appealing alternative to forms of traditional WIL, including studentships, internships, work-study and cooperative placements, because they deliver similar or enhanced learning opportunities to a higher number of learners with fewer human resources, particularly in project settings (Desai et al., 2014). The appeal of CBL arises because traditional WIL situates the learning in a highly structured environment under an authoritative figure responsible for defining what, when, and how tasks are to be completed. Because of this, the learner lacks the opportunity to set work objectives or participate in decision-making. In CBL, the learning is positioned in an environment created in collaboration with peers instead of a supervisor. This positioning shifts the learning responsibility to the students, who must develop strategies to acquire the knowledge and skills needed to compete. Accordingly, conventional WIL involves a structured work environment and continuous one-on-one engagement that are not required for CBL. Moreover, like other forms of student-centred project-based learning (Hoidn and Klemenčič, 2020), CBL creates opportunities for learners to engage in teamwork in ways that contribute directly to the development of intra- and interpersonal skills, such as communication, problem-solving, critical thinking, collaboration, and leadership skills (Abushakra et al., 2019).

The International Genetically Engineered Machine (iGEM) competition is a highly effective CBL program created almost 20 years ago to accelerate innovation in DNA-based biotechnology and bioengineering by applying engineering principles to molecular biology (Smolke, 2009; Vilanova and Porcar, 2014). The competition has facilitated the emergence of Synthetic/Engineering Biology as an applied science discipline and has contributed to numerous biotechnology start-up companies (Wright, 2020). It also helps students develop skills and abilities that bioeconomy employers have identified as lacking in recent graduates (Diep et al., 2021). However, the iGEM competition has a relatively narrow focus and must be complemented by talent and skills programs that can deliver similar learning opportunities more broadly across the bioeconomy.

The BioExperience Research and Entrepreneurship Challenge was created to replicate the exceptional learning outcomes of the iGEM experience and the entrepreneurial and opportunity-seeking mindsets of competition-based learning (Abushakra et al., 2019). The program was developed in 2020 at the University of Ottawa, Ontario, Canada, to mitigate the impact of public health restrictions imposed early in the COVID-19 pandemic. These restrictions caused a widespread loss of work-integrated learning opportunities that many undergraduate students rely on to earn a living while gaining hands-on experience over the summer. The program involves students working in teams to design and complete an applied research project defined by an industry or community partner that involves either a design project, consulting project, or research project.

More than 100 students have completed the program. Most used their participation in place of a cooperative learning work placement or a student research position and received a salary or a bursary. Others participated as unpaid volunteers. Participants working more than 10 h per week were included in most learning activities, and those receiving cooperative learning credits were graded based on the peer assessments. Students participating as research interns and volunteers did not earn academic credits for their work.

The analysis of program evaluations acquired through exit surveys shows that students had an overwhelmingly superb learning experience. In addition to developing technical knowledge, students report significant gains in skills and abilities associated with job readiness and research preparedness, including project planning and management, leadership, team management and collaboration, creative thinking, group thinking, adaptability, time management, organization, interpersonal relations, community engagement and entrepreneurship.

Our findings suggest that the BioExperience program offers a model for situated learning in post-secondary BIOSTEM education that requires relatively few resources. Although further research is necessary, evidence indicates that the program delivers high-quality talent and skills development in areas critical for biotechnology and biological engineering innovation and the growth of the bioeconomy.

Materials and methods

Program evaluation survey

Qualitative and quantitative data was collected through an online survey that participants were asked to complete immediately after the program ended. Google Forms hosted the online survey for the 2020 survey and SurveyMonkey for the 2021 and 2022 surveys. They consisted of questions with predefined Likert scale answers and questions with open-ended text answers. Likert Scale questions were chosen in consultation with members of the Faculty of Education at the University of Ottawa. Questions with open-ended answers asked participants to explain their ratings. The survey was mandatory for full-time participants and optional for part-time participants. Participants were predominantly University of Ottawa science and engineering undergraduate students. Other participants were University of Ottawa undergraduates in medicine and business, undergraduates from science, engineering, and business management at Carleton University and Western University, and secondary students.

Data collection

Survey responses were exported and downloaded to the University of Ottawa IT network for filtering and analysis. There were complete responses from 59 unique respondents in 2020 (95% response rate), 34 respondents in 2021 (62% response rate) and nine respondents in 2022 (45% response rate). Participants who did not complete the survey were predominantly students who were volunteers with no formal association with the program. The anonymized but otherwise complete datasets are available upon request.

Cohort compositions

Survey respondents were mainly science and engineering students (91%) in their third or fourth year. The most represented academic programs are biochemistry (17%), biotechnology (15%), biomedical mechanical or mechanical engineering (13%), biomedical science (11%), chemical engineering (9%) and software and computer engineering (9%). The percentage of female respondents is 59% (2020 cohort), 67% (2021 cohort), and 56% (2022 cohort). The rate of visible minority respondents is 56% (2020 cohort), 67% (2021 cohort), and 56% (2022 cohort).

Data analysis

Participant responses to Likert scale questions were analyzed to determine the count and the percentages of responses for each option on the scale. The five Likert scale options were converted into numerical values from 0 to 4, with 0 corresponding to the most negative response option and four corresponding to the most positive response option. The resulting data were analyzed using R (version 4.3.1) and statistical functions included in the ggplot2 (version 3.3.3) and rstatixs (version 0.7.0) packages available from the Comprehensive R Archive Network. Scripts were written and executed using RStudio (Spotted Wakerobin release).

TABLE 1 Participant responses to open-ended questions were analyzed by associating each answer to one or more recurring themes assigned a unique code, synonym, and description. The recurrent themes are associated with the learning experience (EXP), the learning environment (ENV), and the learning outcomes (OUT).

	Theme code	Theme synonym	Theme description
ENV	COM	Communication	References to interactions among team members
	WFH	Location	References to the working from home
	COL	Collaboration	References to work within a team or other teams
	IND	Independence	Reference to student-led activities
	TIME	Organization	Reference to the timing of program components
EXP	GUI	Guidance	References mentorship, advice, feedback
	RWC	Relevance	Reference to practical or real-world application
	DIV	Opportunities	References to a variety of learning opportunities
	PCP	Impact	Reference to personal contributions
OUT	SOFT	Non-technical skills	References to "soft" skills development
	HARD	Technical skills	References to "hard" skills development
	DEV	Tangible outcomes	References to the work products and deliverables

Qualitative analysis was conducted on participant responses to open-ended questions using an inductive approach.Table 1 is the codebook to guide the investigation. It developed and refined as suggested by MacQueen et al(2016) until the reliability of Fleiss's kappa score ≥0.80 was achieved (McHugh, 2012). The refinement was conducted individually by five members of the analysis team (Vieira et al., 2010; Falotico and Quatto, 2015). Three randomly chosen research assistants were assigned participant response codes. Kappa values were calculated for each code for each question using formulas provided by Nichols et al(2011) and an overall kappa value for each question (De Vries et al., 2008). The final code(s) assigned to each response was restricted to code(s) appearing at least twice. A Python script was written to help determine if a code was present or not.

Study limitations

The design of the program evaluation survey imposes certain limitations. Notably, the quality of the responses could have been improved from the first iteration of the study by using more concise questions and detailed instructions. However, we decided against making changes to the initial survey questions to ensure consistency across the dataset. The study is also limited by insufficient research funding to organize and conduct the follow-up interviews and focus groups needed to confirm the thematic analysis results.

Results

Program development

The BioExperience program was launched to mitigate the impact of the COVID-19 pandemic on traditional undergraduate WIL programs at the University of Ottawa in Ontario, Canada. It was developed in April 2020 by faculty members from science, medicine, and engineering in collaboration with the University of Ottawa Co-operative (CO-OP) Programs Office and BioTalent Canada, a Canadian national non-profit human resources association for the biotechnology industry.

The rapid spread of the SARS-CoV-2 virus in early 2020 led to the closure of Canada's international borders to most foreign nationals, the suspension of classes at Canadian post-secondary institutions, and the closure of all non-essential workplaces. Although the duration of these measures was unclear, they were expected to significantly impact undergraduate students' ability to access the WIL programs that many rely on for meaningful employment and skills development during the summer term.

We designed the BioExperience program to replicate the student-centred approach of the iGEM competition by giving students a direct role in project design and team building. The freedom to self-manage is a central element of iGEM and is associated with essential intrapersonal and interpersonal skills (Farny, 2018). We also sought to design a program structure that addresses challenges experienced by students participating in the iGEM competition (Diep et al., 2021), including students choosing unrealistic and overly ambitious projects, participant

TABLE 2 Examples of Team Challenges from the 2020 BioExperience program. Full project descriptions and reports are available at https://biogroupe.ca/bioexperience/.

DIY 3D bioprinter systems for tissue engineering
Tissue engineering, which seeks to regenerate the tissues and organs in our bodies from the combination of cells, scaffolds, and bioactive signals, has shown promise for the treatment of injuries and diseases and for the development of improved *in vitro* models to study physiological and pathological cellular processes. This project aims to generate comprehensive knowledge of these technologies and develop a detailed plan to build a novel, cost-effective, and versatile 3D bioprinting system that can generate complex multicellular and anisotropic tissue structures for various applications
Experimental Strategies for the Craft Beer and Vodka Community
To meet the increasing demand for highly skilled workers in the microbrewery and craft distillery, the Faculty of Engineering is developing a pilot-scale microbrewery and craft distillery. Once established, this facility will provide an experiential learning environment where students can acquire knowledge and hands-on expertise to increase their job readiness and meet the recruiting needs of a rapidly growing sector of the bioeconomy. To better prepare students for employment in the microbrewery and craft distillery sector, the proposed project will investigate economic small-scale production alternatives for local craft businesses and develop experiential learning activities and experimental strategies for small-scale brewing and distillation systems
Biodigital Convergence—Strategic Foresight in an Era of Disruption
Advanced DNA-based biotechnology is radically changing our economy, ecosystems, and society. Biodigital Convergence—the merging of digital and biological technologies and systems, has the potential to change the way we work, live, and play. This project aims to examine how future biodigital technologies could be woven into our lives and potentially transform our understanding of ourselves and the natural world, the meaning of human connection and the essence of humanity itself
Emerging Biotechnologies for COVID-19 Point-of-care Testing
Diagnostic testing of viral infection often involves processes prone to error and require specialized equipment. Several biotechnologies have recently been developed to address these issues and enable cost-effective and reliable point-of-care testing. This project aims to generate comprehensive knowledge of the biological principles behind SARS-CoV-2 diagnostic testing methods and to develop a design implementation of rapid COVID-19 point-of-care testing

dissatisfaction with mandatory tasks deemed irrelevant, and the breakdown of social cohesion within teams.

Students were given a direct role in project design by creating a context where a team of students is commissioned by a "client" to complete a research or entrepreneurship "challenge" on their behalf. The challenge is a short statement defining each team's focus in terms of a problem, task, or question. Examples of Team Challenges are provided in Table 2. They must:

- Address an issue directly or indirectly related to an area of the bioeconomy,
- Require students to plan and engage in applied research while developing their skills and knowledge,
- Allow students to define their objectives and problem-solving strategies.

It is helpful to think of a challenge as a task outsourced to a student team, much like an organization might outsource work to a consulting company. In both scenarios, the "client" defines general expectations and parameters and delegates the details to "consultants." This approach transfers the project's ownership and responsibility to the students, who must collaborate to identify project objectives and develop the appropriate research plans and problem-solving strategies. It also gives teams the flexibility to ensure that their project accommodates the interests, skills and learning needs of all team members. We envisioned adopting this approach would translate into an engaging and motivating learning experience by allowing participants to focus on activities that align with their interests and long-term aspirations.

We initially focused on projects that students could complete without access to a laboratory or other physical workspaces, including design projects, consulting projects, business development projects and projects involving literature reviews. We did this in anticipation that students would not be able to return to in-person learning. However, challenges requiring field- or laboratory work or other in-person activities are allowed if the team client or academic advisor provides the required facilities, resources, and training.

We also focused on projects endorsed by a faculty member committed to providing guidance and project management support. We did this to ensure that teams could set achievable goals, develop realistic research plans, and maintain a safe and productive learning environment. Projects can be initiated by a non-academic client, including, for example, private companies, business development organizations, and government agencies, or by faculty members. In the latter case, the faculty member acts as the team client and academic advisor.

We implemented a stacked recruitment process that gives students a key role in team building and project development. First, an initial cohort of 10 "Team Leads" was recruited to identify the skills and abilities needed to complete their challenges. Students in the second cohort were recruited to build a core team, while students in the third cohort were recruited to fill any remaining gaps. The idea was for faculty advisors to interview and assign Team Leads to specific projects and for Team Lead to work with the advisor to identify

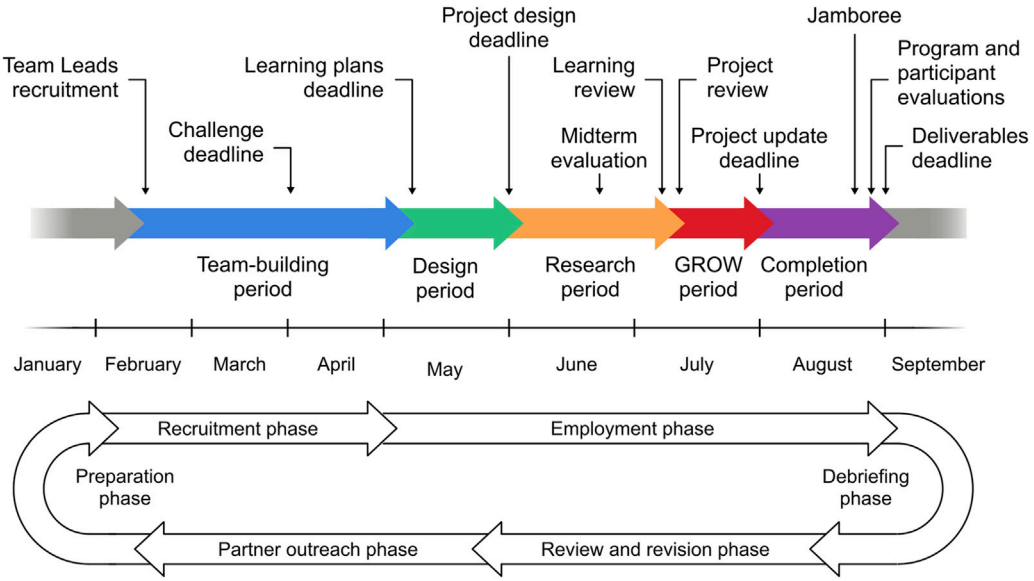


FIGURE 1
Program overview. Schematic illustration of milestones and timelines spanning a recruitment period from January to the end of April and an employment or work period from May to August. The recruitment phase and the first week of the work period are dedicated to team building. The remaining employment period consists of a design period, a research review, a review/GROW period, and a completion period. The program ends with a debriefing and a review and revision period. The planning for the following summer starts with a partner outreach period and is followed by a recruitment preparation period.

recruitment needs, write job postings, and conduct applicant interviews with support from their advisors. We also anticipated that this would help teams develop and maintain social cohesion by clarifying each team member's purpose and roles, and responsibilities.

The launch of the program was made possible by a commitment in mid-April 2020 by BioTalent Canada to pre-approve 50 University Ottawa CO-OP students for up to CAD 7,500 for their Student Work Placement Program (SWPP). Shortly after, the University of Ottawa agreed to provide administrative support and complement the BioTalent Canada funding with CAD 4,500 per student through the University of Ottawa Work-Study program.

A job posting advertising the 10 Cohort I Team Lead positions was released on 29 April 2020. It emphasized that ideal candidates should be able to work harmoniously with others in a team-based and project-oriented environment, be self-directed and self-motivated, comfortable with initiative and leadership, and have an interest in project development and management. Among 56 applicants, 11 Team Leads (one more than initially expected) were recruited to develop the applied research projects. Subsequent Cohort II and Cohort III recruitment, which was extended to accommodate students who became unemployed in mid-June, resulted in the enrolment of 59 students. Remarkably, all students completed the program, and only one project failed to proceed beyond the initial development phase. This failure was caused by insufficient support from a faculty member. The orphaned student was successfully integrated into another project where they became responsible for a sub-project.

Core components

The BioExperience program has evolved since its first iteration but has retained a structure composed of five core components, including team building, project design, research, review, and completion. The order of the components and their associated milestones are depicted in Figure 1. They are envisioned to span 8 months, with a low-intensity recruitment phase from January to April and a high-intensity period from May to August. The figure also depicts program phases that do not involve students directly. They include a debriefing phase involving discussions with participating faculty members and non-academic partners, and a program review and revision phase to identify areas where the program can be improved. The solicitation of challenges for the next summer starts in November faculty members and non-academic partners,

and preparations for the recruitment of Team Leads start in December.

Team-building period

The team building period begins with recruiting one or more Team Leads and ends with a learning activity where students share personal learning plans with their teammates. As mentioned in the previous section, Team Leads, or Cohort I students, are recruited to work with the Team Client to develop a project outline that can guide the team-building process. The project outline helps the Team Leads create job postings and conduct student interviews by identifying the skills and competencies needed to complete the project.

This team-building process allows Team Leads to practice their leadership and human resources skills. A faculty member supports Team Leads during the hiring process to ensure that program expectations and the roles and responsibilities of Cohort II and Cohort III recruits are communicated accurately. Notably, it is essential to firmly establish that the responsibility of the Team Leads is to facilitate the development of a supportive and inclusive learning environment where students engage in work that aligns with their interests, skills, and learning needs. They are not responsible for the project's overall success and are not permitted to assume the role of an employer. In other words, they cannot instruct or supervise the work of other team members.

The first week of the employment period focuses on developing personal learning plans. The goal is to encourage students to take ownership of their learning and facilitate constructing a learning environment that accommodates all team members' interests, goals, and needs. The personal learning plan is a mandatory component of BioTalent Canada's SWPP program and consists of answers to four open-ended questions:

1. Where am I now, and where do I want to go?
2. Which skills and competencies do I need to succeed?
3. What learning activities will help me acquire these skills?
4. How can my teammates and advisors support me?

Students participating in the program's first iteration found these questions challenging, and many could not articulate clear goals or identify the skills and abilities they would need to achieve them. Accordingly, a team-based learning activity called "Own Your Learning" was created to get students PRIMED for self-directed learning through self-reflection and group discussions with team peers. Preparing students PRIMED for learning involves:

- Making it more apparent to students what they want to accomplish and why (Purpose),
- Letting students take charge and ownership of their learning (Responsibility),
- Encouraging students to be their own and each other's teachers (Independence),
- Creating a sense of belonging, meaning and accomplishment (Motivation),
- Facilitating personal growth and professional development (Evolution), and
- Helping students stay focused as they progress through the program (Direction).

The question "where do I want to go?" was challenging for many students. Accordingly, in 2021, students were asked to create a list of specific, measurable, attainable, relevant, and timely (SMART) learning goals (O'Neill and Conzemius, 2006) before completing their learning plan. Students were specifically asked to answer the following questions:

- What will achieving your learning goals allow you to do? Why is this ability important to you (Specific)
- How will you know that you have successfully reached your goals? What are ways to assess your progress (Measurable)
- Are your goals realistic? What are the skills and competencies you need to achieve them? Do you already have the necessary expertise? Are resources available to access or acquire them (Attainable)
- How are your goals related to your personal and professional aspirations? Do they make sense in the context of the BioExperience program and your Team Challenge (Relevant)
- Is this the right time to pursue these specific goals? Why is it important to attain them now rather than later? Could your time be spent on something more substantial (Timely)

Clarifying the purpose of the personal learning plan using the PRIMED concept and introducing SMART learning goals was further augmented with surveys to help students identify and prioritize specific skills and abilities and to monitor their progress throughout the program.

Design period

The project design phase occurs after developing personal learning plans to ensure that students have an opportunity to create a project that supports the learning of all team members. During the design phase, teams are expected to use the project outline created by the Team Lead as the foundation of a detailed account describing how they plan to complete their challenge.

Students are asked to answer three questions individually:

1. Why is addressing this challenge significant?
2. Who will benefit from the work done by the Team?
3. How will they benefit?

The goal is to help the student connect to the project on a personal level before working with their peers to create a shared vision of what they would like to achieve as a team. The shared vision is expected to support collaboration and decision-making and strengthen social cohesion within the Team. It also helps teams adopt a backward design strategy (Emory, 2014) that focuses on the project's overall purpose and encourages students to think creatively, take risks and explore opportunities as they emerge.

The development of a purpose-driven project is facilitated by teams answering a series of sequential questions in a project design guide. The questions, in order, are:

1. What do you aspire to accomplish by completing your project (Project Goal)
2. Why is reaching the goal important? Who will benefit? How (Project Purpose)
3. How will you know that you are successful? What will you create to demonstrate to your client that you have achieved your goal (Tangible Outcomes)
4. What are the significant steps to be completed before you can generate these outcomes (Specific Aims).

To further support research planning, teams are asked to produce a step-by-step research plan describing what they intend to do to complete each specific aim, including a Gantt chart (Geraldi and Lechter, 2012) to visualize timelines. They are also asked to describe the knowledge and skills the Team needs to acquire, how they will be developed, what risks might prevent the Team from succeeding, and how it will mitigate them.

Research period

The research period is when teams execute their research plan. Teams are expected to meet with their faculty advisors at least once per week and to provide brief progress updates at a weekly meeting of all participants. This meeting aims to connect students from different teams and create team collaboration opportunities.

Students must work closely with their teammates and advisors. They are expected to have a daily team meeting to coordinate their work and share their findings, support the learning of others, and seek assistance from teammates when needed. They are also told that they are expected to collaborate to define, coordinate, and delegate tasks consistent with individual team members' interests and learning goals. This expectation creates rich opportunities for students to engage in negotiation, mediation, and conflict resolution.

Team leads are tasked with maintaining a positive and productive learning environment and are supported in this responsibility through weekly meetings with the program director and other team leads. The program director is required to mediate conflict, resolve differences of opinion, or restore social cohesion. However, it is crucial to convey to all participants that the program facilitates professional development and that challenges are opportunities to improve interpersonal and intrapersonal skills.

The midpoint of the research period includes a peer- and self-assessment survey. The self-assessment asks students to reflect on their learning progress by rating their proficiency in the skills and abilities they were asked to prioritize earlier in the program. They are also invited to identify up to three skills and abilities they feel they have developed and describe how they demonstrated high proficiency. The peer-assessment questionnaire is identical to the self-assessment questionnaire and asks students to rate their teammates' proficiencies the same way they rated themselves. They are also invited to identify up to three skills and abilities that stand out and to describe how the teammate demonstrated proficiency. The peer assessments provide valuable feedback and give participants a complete picture of their learning journey by identifying areas where progress has been made and areas to develop further.

Review (GROW) period

The research period ends with a review period where teams assess their progress and reassess their project design. The review was done using a GROW project planning model (Kang et al., 2021), which asks teams to answer the following questions:

- What is it that you are trying to achieve (Goals)?
- What progress have you made? What is impeding your progress (Reality)?
- What is the ideal solution? What are realistic solutions (Options)?
- What will you do now? How will you do it (Will)?

Completing the GROW activity helps teams ensure that their goals are still achievable in the remaining time based on what they have learned during the research period and gain an increased understanding of the time it takes to complete various tasks. It also allows teams to revise their tangible outcomes in case the initial plan was overly ambitious or unexpected complications emerged during the research period.

Completion period

The final period of the program begins after the review period. Teams have now assessed their plans and made the necessary changes to ensure their objectives are achievable in the time remaining in the program term. This period is termed the "Go Period," where teams focus on completing their tangible outcomes.

The Completion Phase culminates in a final celebratory Jamboree when each Team presents their achievements. This presentation acts as an opportunity for teams to showcase their work, celebrate each other's success, and give members of the public a window into the work achieved by the teams. The presentations are not evaluated as the focus is on

TABLE 3 Leadership skills development. Participants were asked to rate the improvement in their ability to complete tasks associated with leadership. The rating options were none, minor, moderate, significant, and profound. (A) Fraction of respondents who rated their improvement in specific abilities as significant or profound. (B) Fraction of respondents who rated their overall skill improvement as high (significant or profound) or low (none or minor). The overall rating *r* of each respondent was determined by assigning to each rating option a numeric value from zero (none) to four (profound) and averaging these values. The number of participants providing a high and low overall rating is the number of partitions with $r \geq 2.5$ and $r < 1.5$, respectively.

(A) Description	N	Count	Percent (%)
Identify and set realistic project goals, tasks, and priorities	89	63	71
Manage, delegate and coordinate project tasks to match the interests and competencies of team members	89	65	73
Recognize the capabilities of individual team members through inclusion in decision-making	89	65	73
Build and maintain team cohesion, morale, and discipline	89	66	74
Provide support and guidance to team members	89	66	74
Ensure that team members can contribute in ways which support their learning	89	60	67
Resolve conflicts and negotiate differences of opinion among team members	88	45	51
(B) Overall rating fractions			
High overall rating (significant or profound improvement)	89	64	72
Low overall rating (none or minor improvement)	89	5	6

understanding how obstacles were overcome and what future steps for the projects might be.

Auxiliary components

In 2021, we began offering weekly instructor-led learning activities covering various topics, from researching literature using databases to writing a personal learning plan and creating compelling pitch presentations. We did this because it was apparent that many students could not reach program milestones independently. We also worked with BioTalent Canada to provide participants with access to their "Skills for Success" online workshop series. These modules cover essential interpersonal skills such as communication, collaboration and problem-solving as well as technical skills such as quality assurance/quality control and good lab practices. This auxiliary component is not easily reproduced.

Skills improvements

The BioExperience program aims to help participants improve skills essential for employment in the bioeconomy and graduate studies in a BIOSTEM field. To assess if this objective was achieved, we examined quantitative Likert-scale ratings by the participants of their advancement in eight higher-order non-technical skills, including self-management, problem-solving, critical thinking, research, collaboration, teamwork, communication, and leadership. Participants were asked to rate the degree to which the program helped them improve relevant proficiency on a five-point scale from none to profound.

To assess participant skills and competency development, we calculated the percentage of respondents reporting high improvement in their ability to complete specific tasks and activities. Table 3 provides an example for the rating of improvement of leadership skills. This skill is associated with seven sets of abilities, such as setting team goals and resolving conflicts among team members. In the example, 74% of respondents rate their ability to "build and maintain team cohesion, morale and discipline" improved significantly or profoundly. In comparison, 51% give this high rating for their ability to "resolve conflicts and negotiate differences of opinion among team members".

To get a sense of the overall rating of individual respondents, we also calculated an aggregate rating to better assess the average improvement rating across the sets of abilities used to assess higher-order skills. In this calculation, we used a linear numerical scale from zero (no improvement) to four (profound improvement) and calculated an overall rating score *r* for respondents by averaging. The fraction of high and low ratings was then computed by counting the respondents with an overall rating of $r \geq 2.5$ or $r < 1.5$, respectively. In the example in Table 3, 64 respondents (72%) rated their improvement in leadership skills as high (i.e., significant, or profound improvement). In comparison, five respondents (6%) rated their improvement as low (i.e., no, or minor improvement). The remaining respondents rated their improvement as moderate.

The ratings for seven higher-orders skill groups suggest that most participants saw skill improvements. Figure 2 depicts the distribution of all respondent rating scores for

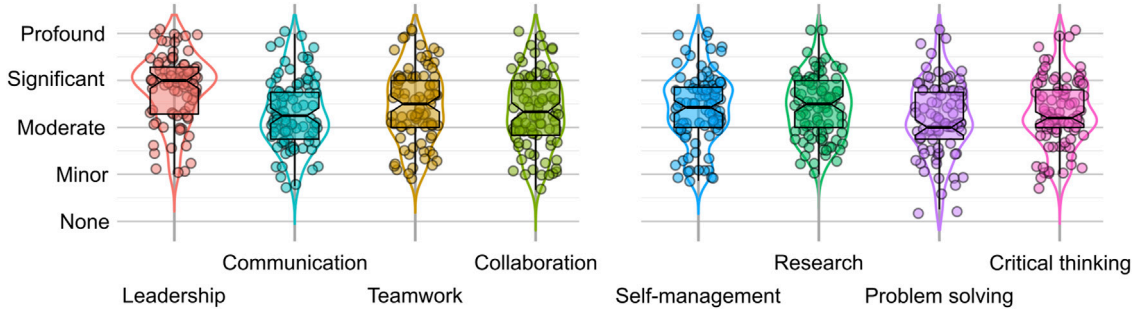


FIGURE 2
Higher-order skills improvement. Participant response distributions for higher-order skill groups. Data points indicate the average respondent rating for each group (see Table 1 for details). Curves represent the distribution of data points (violin plots), and the notched boxes represent the 2nd and 3rd quartiles of the data (box plots).

TABLE 4 High-order skills development. The percentage of respondents rating their improvement as high (significant or profound) or low (none or minor). The most (least) frequently improved abilities are the abilities that most (least) respondents identified as significantly or profoundly improved.

Skill	High rating (%)	Low rating (%)	Most frequently improved	Least frequently improved
Self-management	49	16	Be reliable and consistent in the completion of tasks; Set and manage personal schedules and priorities (57%)	Manage mental and physical health (30%)
Problem solving	42	14	Integrate multiple perspectives and disciplines in problem-solving strategies (51%)	Perform data analysis and hypothesis generation (28%)
Critical thinking	40	12	Integration of conflicting evidence or viewpoints (51%)	Be aware of and able to challenge biases, inferences, and assumptions (38%)
Research	52	6	Develop and manage research or product development projects (54%)	Conduct and organize literature or technical reviews (47%)
Collaboration	48	14	Negotiate and distribute tasks fairly (54%)	Resolve conflicts with or among others (35%)
Teamwork	52	11	Share knowledge and support the learning of others (58%)	Reach consensus through negotiation and compromise (40%)
Communication	44	8	Use plain language to communicate complex information (58%) clearly	Develop clear technical standards, procedures, protocols, or guidelines (38%)
Leadership	72	6	Build and maintain team cohesion, morale, and discipline; Provide support and guidance to individual team members (74%)	Resolve conflicts and negotiate differences of opinion (51%)

each skill group on a continuous scale from zero (no improvement) to four (profound improvement). The median rating scores lie between 2.0 (moderate) for problem-solving and 3.0 (significant) for leadership. While individual scores vary considerably within each group, the box- and violin plots used to represent the data clearly illustrate that most respondents reported a moderate improvement or higher. The lowest first quartile value is 1.75 for communication and problem-solving, meaning that at least 75% of respondents saw medium ($1.5 \leq r < 2.5$) or high ($r \geq 2.5$) overall improvement.

The percentage of respondents with high ($r \geq 2.5$) and low ($r < 1.5$) overall improvement ratings provide insight into the ability of the program to facilitate the development of higher-order skills and competencies. Table 4 summarizes the results of our analysis by presenting the fraction of responses indicating high and low overall improvement rates. The fraction of respondents rating their improvement as high is around 50% for most skills, with significant or profound improvement, most common in leadership skills (74%) and research skills (60%). The table also includes the survey questions with the highest and the lowest number of responses with an improvement rating of significant or profound. For example, in the research skills, the ability to "develop and manage research or product development projects" was the ability improved by the most participants (54%), while the ability to "conduct and organize literature or technical reviews" was the ability improved by the fewest participants (47%).

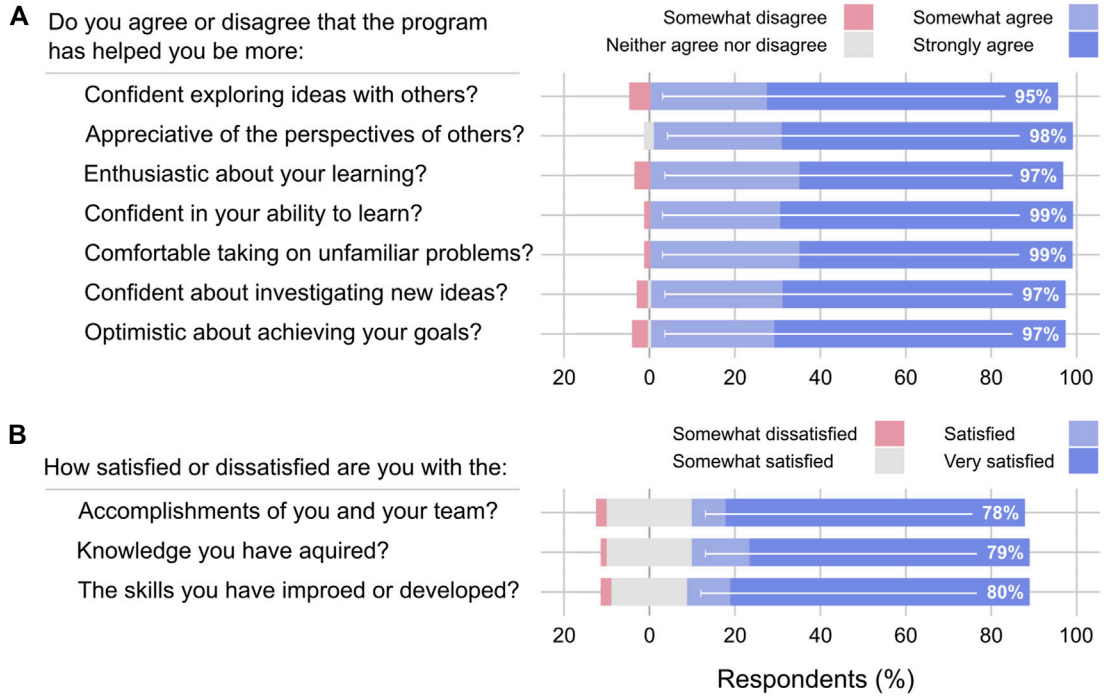


FIGURE 3
Learning outcomes. **(A)** Personal development. **(B)** Personal achievements. Each horizontal bar includes the percentage of respondents selecting the two highest options on a five-point Likert scale. None of the respondents chose the lowest option (unacceptable, strongly disagree, or very dissatisfied).

Personal development and achievements

The personal development of participants was assessed through quantitative and qualitative means. In the quantitative approach, we were particularly interested in gauging to what degree participants felt the program had helped them improve in areas important for self-motivation and self-direction and developing an inquisitive mindset. The results are summarized in Figure 3A.

Almost all respondents agree that the program helped their personal development in one area or another. For example, 95% of respondents agree or strongly agree that the program had made them more "confident about my ability to learn," 99% agree or strongly agree they are more "comfortable taking on unfamiliar problems," and 97% agree or strongly agree that they are more "positive about achieving my goals." To further assess personal development, we asked participants to describe up to three things they learned that would be valuable in their future (data not shown). The most frequently referenced themes were non-technical skills (75% of participants) and technical skills (65%). Recurring subjects were leadership, communication, teamwork, adaptability, self-management, entrepreneurship, project planning, programming, engineering design, and biotechnology and bioeconomy knowledge.

We also asked participants to rate their satisfaction with their achievements and to describe what they were most proud of having accomplished in the program. The results are summarized in Figure 3B. Almost all respondents reported that they were somewhat satisfied, satisfied or very satisfied with their accomplishments (98%), the knowledge they acquired (99%) and the skills they developed (98%).

Remarkably, roughly two of every three respondents expressed that they were very satisfied with their achievements in at least one area. In open-ended answers, 61% of respondents referenced a work product or project deliverable as their proudest accomplishment. They also identified non-technical skills (20%) and technical skills (17%), with communication, leadership, and collaboration often recurring in responses referencing non-technical skills. Software and web development, product development, and research abilities were the most frequently recurring technical skills.

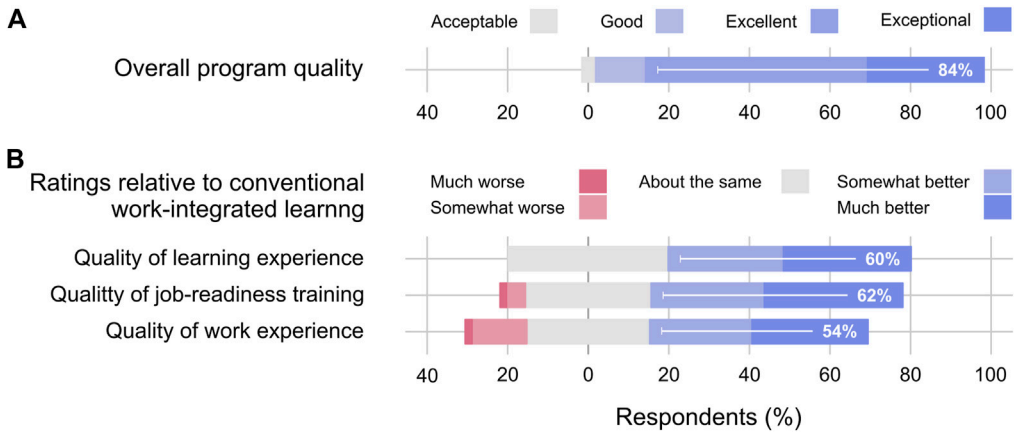


FIGURE 4
Program quality. **(A)** Distribution of responses to the survey question "How would you rate your overall experience?" **(B)** Distribution of ratings comparing the BioExperience to conventional WIL programs. The quality of the work experience refers to the meaningfulness of the work as perceived by respondents. Horizontal bars include the percentage of respondents selecting "Excellent" or "Exceptional" in **(A)** and "Somewhat better" or "Much better" in **(B)**.

Program quality

We asked participants to provide an overall rating of the program and to describe what worked well for them in the program. Remarkably, 84% of respondents rated the overall experience as "excellent" (49 respondents) or "exceptional" (26 respondents). The remaining respondents gave a rating of "good" (11 respondents) or "acceptable" (3 respondents). This data is presented in Figure 4A. In the answers to the open-ended question, the three most frequently identified themes were communication (43% of respondents), collaboration (40%) and technical skills development (21%). Recurring sub-themes were team meetings and software for communication, delegating tasks, setting common goals, peer support for collaboration, and writing and software use for technical skills.

We also asked participants to describe why they might complete the program in the future or recommend it to a peer. Respondents frequently referenced the ability to develop technical skills (24%), non-technical skills (20%), and learning opportunities (24%). Subjects mentioned frequently by respondents included leadership opportunities, the development of self-awareness, the management of research projects, and the ability to choose projects related to their academic studies and career goals. The most frequent reasons why respondents would recommend the program are the learning opportunities (35% of respondents) and the development of non-technical (29%) or technical skills (21%). Other common reasons participants recommend the program include the student-led learning environment (19%) and collaboration (11%). Recurring subjects were related to the participant's ability to learn about things they are interested in or passionate about in a diverse and inclusive environment.

Comparison to conventional WIL

We asked participants who had previously participated in WIL to provide a comparative rating regarding the quality of the learning experience, the value of the program for job-readiness training, and the ability to do meaningful work. The summary results are presented in Figure 4B. All respondents rated the quality of the learning experience as similar or better than a previous conventional WIL experience, and more than half rated it somewhat better or much better. Similar ratings were obtained for job-readiness training and work experience, where 62% and 54% rated the quality of the BioExperience program better than conventional WIL.

We also asked participants to explain their ratings. In the context of the learning experience, the most frequent themes were technical skills development (44% of respondents), non-technical skills development (32%), collaboration (27%) and independence (24%). In the technical skills area, recurring subjects were increased knowledge of the bioeconomy and biotechnology research, the development of entrepreneurial skills, and proficiency in research design and management. In the non-technical skills area, respondents frequently mentioned project development and management, teamwork, and self-directed learning, while working with like-minded students, problem-solving and self-direction were common subjects in the collaboration and independence themes.

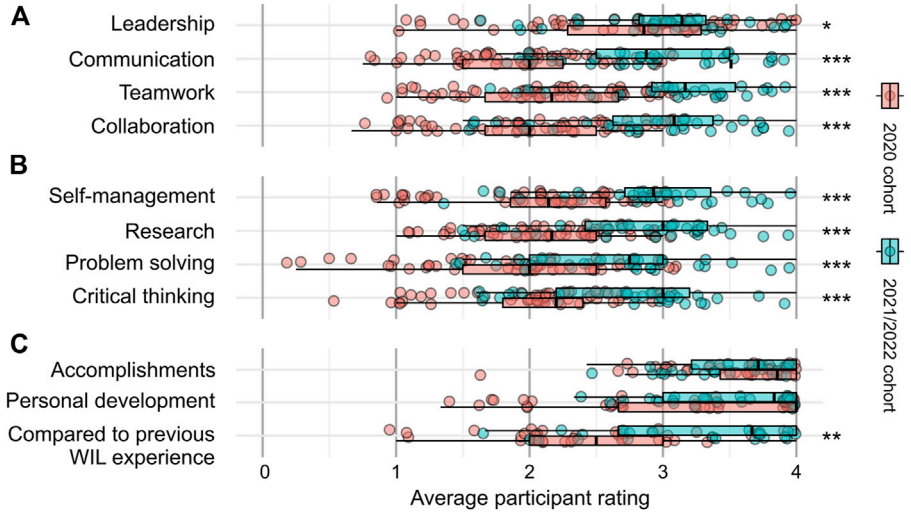


FIGURE 5
Program improvements. Response distributions grouped by participants from the first iteration of the program (2020 cohort) and participants from later iterations (2021/2022 cohort) grouped by **(A)** interpersonal skills, **(B)** cognitive and intrapersonal skills, and **(C)** personal experiences. Asterisks indicate the upper value of the *p*-value associated with the null hypothesis that the two distributions are not shifted relative to one another (Two-sample Wilcox test, one-sided). One asterisk corresponds to $p < 0.05$, two asterisks correspond to $p < 0.002$, and three asterisks correspond to $p < 0.0001$).

In terms of improved job-readiness training and work relevance, respondents frequently referenced the relevance of their work to real-world problems (26% of respondents) and the development of non-technical skills (18%) and technical skills (16%). Recurrent subjects include helping others and addressing societal challenges, teamwork, organization, self-motivation, communication, and leadership in the non-technical skills theme, research skills, technical knowledge, entrepreneurship, engineering design and scientific writing in the technical skills theme.

Areas for improvement

To identify areas of improvement, we asked students to describe something that did not work well for them, one thing they could change about their work-related interactions, and one thing they would add to the program if they could. The responses to the two first questions were consistent, with the most frequently referenced identified being communication, workplace location (i.e., working from home), and guidance. Guidance and learning opportunities were the most identified areas of improvement. However, there was a significant difference between the 2020 and 2021 program iterations. While 38 responses from the 2020 survey referenced guidance or opportunities as areas of improvement, corresponding to 45% and 36% of respondents, they were only referenced in one response from the 2021 survey. In the 2021 survey, the most frequently referenced areas of improvement were non-technical skills (38%), technical skills (37%) and program structure (38%).

The differences between program iterations were explored further by conducting a statistical comparison of participants' responses from 2020 to those collected in 2021 and 2022. The results in Figure 5 indicate a significant improvement in interpersonal and intrapersonal skills development. In terms of interpersonal skills (Figure 5A), the median rating increases from roughly two (moderate improvement) to three (significant improvement) in communication, teamwork, and collaboration skills. The same trend is observed for intrapersonal skills (Figure 5B) and certain program aspects (Figure 5C). Notably, more than 50% of 2021 and 2022 respondents reported significant or profound gains in self-management, research, problem-solving and critical thinking skills.

Interestingly, the rating of the BioExperience compared to conventional WIL programs also increased significantly after the first year. This effect is quite striking. The median rating was roughly 2.5 among 2020 respondents and close to 3.8 among 2021 and 2022 respondents. This increase reflects that 76% of respondents (19 of 25) participating in 2021 and 2022 reported that the program is better (24%) or much better (52%) than conventional WIL.

Discussions and conclusion

The BioExperience Research and Entrepreneurship challenge was created to mitigate the widespread disruption of conventional post-secondary WIL programs caused by the Covid-19 pandemic. The program provides meaningful talent

and skills development opportunities through a student-centred, project-based approach inspired by the iGEM competition. The competition has been a unifying force in the synthetic biology community for almost two decades and has allowed countless students to explore biology-based technology and engineering solutions to significant societal problems. We have demonstrated that critical talent and skills development elements of the iGEM competition can be replicated in a program that addresses deficiencies documented in recent research and requires significantly fewer resources than conventional WIL models.

The BioExperience program differs from the iGEM competition by having each team focus on a challenge posed by a program partner, such as a faculty member, a company, or a non-governmental organization. This partner, or team client, supports the team throughout the program in an advisory capacity and is not permitted to assume the role of an employer or supervisor. Students must apply their creativity and ingenuity and use their combined knowledge, skills, and abilities to produce tangible solutions. This independence allows students to develop numerous cognitive, intrapersonal, and interpersonal skills, simultaneously acquiring technical and field-specific competencies that align with their interests, personal needs, and long-term goals.

The responses from BioExperience participants are striking. We knew from experience that the iGEM-inspired approach would provide a learning experience of high quality. Still, we did not anticipate that ~90% respondents rated the program as excellent or outstanding, or that 19 of 20 respondents reported significant or profound skills improvements in at least one area. Remarkably, three of four respondents found the program to be better or much better than conventional WIL regarding the overall experience, the job-readiness training provided, and the ability to do meaningful work.

Our thematic analysis highlights factors contributing to the learning experience, environment, and outcomes central to the program. The variety of learning opportunities, the relevance to real-world problems, and the impact of personal contributions were highlighted as factors contributing to a high-quality learning experience. More specifically, the independence and self-management afforded by the student-centred learning approach and the program's collaborative nature were frequently mentioned as superior to conventional WIL.

The interpretation of the survey data results is obscured by the unclear impact of the COVID-19 pandemic on the BioExperience learning environment. For example, guidance and support from advisors were identified as improvement areas but were referenced predominantly by participants in the 2020 program. It is unclear if this change results from advisors being more consistently available in 2021 and 2022 (we recommend at least two hours per week), for example, or students becoming more accustomed to working independently. While many respondents reported that working from home had a negative effect, others highlighted the opportunity to work from home as a benefit of the program. Regardless, our results strongly suggest that the program can provide a high-quality experience without facilities, such as shared office space and classrooms, where participants can work in physical proximity. To improve communication and social cohesion, we now encourage teams to organize regular in-person meetings to use public facilities such as libraries, coffee shops or pubs, and clients to meet occasionally with the team at their workplace.

The analysis also highlights the work products and deliverables, and the development of technical and non-technical skills, as critical factors. Most respondents expressed that they were very satisfied with the project's tangible outcomes they produced. Although some teams were required to create specific products and deliverables for their client, including physical prototypes, protocols, and educational materials, most were only asked to produce the deliverables required by the program. These deliverables are the same for all teams: a 250-word non-technical project summary and a 2500-word non-technical project description, a two-minute pitch video, a 15-min presentation, and technical reports that are sufficiently detailed for future teams to continue the project. Because of the high satisfaction rate (98% of respondents were satisfied and 60% were very satisfied) and frequent positive references to work products and deliverables, we see no reason to suggest changes to the program.

Cognitive, interpersonal, and interpersonal skills are essential to many employers (BioTalent Canada, 2021) and success in professional and research-based graduate programs (Madan and Teitge, 2013). These skills are difficult to acquire in conventional academic programs, and BioTalent Canada recently called for expanding WIL opportunities (BioTalent Canada, 2020). BioExperience program participants reported that the program improved essential cognitive, interpersonal, and intrapersonal skills. While ~50% of the respondents reported significant or profound improvement in self-management, problem-solving, critical thinking, collaboration, teamwork, or communication skills, the program appears well-suited for developing research and leadership skills.

Remarkably, 98% of respondents were satisfied with their improvement in non-technical skills, and more than two-thirds were very satisfied. These improvements are also reflected in the roughly 95% of respondents who agree that the program contributed to their personal development by increasing their confidence and enthusiasm for learning, comfort with taking on unfamiliar challenges, and confidence in their ability to investigate and share new

ideas. Notably, intrapersonal skills development is vital to emotional wellness, optimism, self-esteem, effective leadership, and educational success (Hindes et al., 2008; Ferreira et al., 2020). It creates the foundation for developing an inquisitive and opportunity-seeking mindset that contributes to success in research-based graduate programs. Nonetheless, because of the importance of non-technical skills development, we recommend further augmenting the student-centred learning model with additional instructor-led learning activities. The goal of these activities is to further support participant skill development by providing them with structured opportunities to learn critical skills, such as creating learning plans and conducting self-reflections, that will make their experience in the BioExperience program more rewarding.

Data availability statement

The definitions used to define higher-order skills used in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

MK designed the study and the survey. MK did the quantitative analysis. HG and MA researched thematic analysis and created the codebook. HG, MA, JP, VF, YK, TS, and MR coded the answers to open-ended survey questions. HG, MA, AB, and MK wrote the manuscript. HG, YK, and MK edited the manuscript.

Funding

This work was supported by an NSERC Discovery Grant (for MK, grant #05659-2016), NSERC Student Research Awards (for HG and MR), the University of Ottawa Faculty of Science Undergraduate Research Scholarship program (for HG, MA, and JP), the University of Ottawa Work-Study program (for MA, YK, and TS), and the BioTalent Canada Student Work Placement Program (for MA, YK, and TS).

Acknowledgments

The authors would like to acknowledge contributions from Ndolo Njie-Mokonya, The Ottawa Hospital and Faculty of Health Sciences, University of Ottawa, to preliminary investigations and Trista Hollweck, the Atlantic Rim Collaboratory and the Faculty of Education, University of Ottawa, for assistance with the development of the survey. The BioExperience program would not have been possible without the support of Joanne Jordan at BioTalent Canada and the support of Gabi St-Pierre, Chantal Yelle, Amine Manceur, and Manon Racine at the University of Ottawa CO-OP office. MK would like to acknowledge the contribution to this work by Lisa Howell, Faculty of Education, University of Ottawa, for the ongoing support and editing.

Conflict of interest

MK is the co-founder and President of BioGroupe Canada Inc., a non-profit organization founded to make the BioExperience Research and Entrepreneurship Challenge and other talent and skills development programs more readily available and accessible to post-secondary students. Author AB is affiliated with the company Canadian Synthetic Biology Educational Research Group (CSBERG).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2022.1046723/full#supplementary-material

References

Abushakra, A., Firdouse, M., Khan, R., Wahhab, R. A., and Al Maqbali, H. (2019). Innovation in entrepreneurship education through competition based learning role: Students' perspective on the enhancement of soft skills. *Humanit. Soc. Sci. Rev.* 7, 862–869. doi:10.18510/hssr.2019.75112

Billett, S. (2009). Realising the educational worth of integrating work experiences in higher education. *Stud. High. Educ.* 34, 827–843. doi:10.1080/03075070802706561

BioTalent Canada (2021). *BioTalent Canada's report "Close-up on the bio-economy – national report"*. Ottawa: BioTalent Canada. Available at: https://www.biotalent.ca/wp-content/uploads/BioTalent-Canada-LMI-National-Report-13OCT2021-1.pdf (Accessed November 2, 2022).

BioTalent Canada (2020). *The Talent Differential the case for work-integrated learning in the bio-economy*. Ottawa: BioTalent Canada. Available at: https://www.biotalent.ca/wp-content/uploads/BioTalentCanada_LMI-The-Talent-Differential-09OCT2020-1.pdf (Accessed September 16, 2022).

Cameron, M., and Rexe, D. (2022). Community-based access to apprenticeship: An Indigenous work-integrated learning model. *Int. J. Work - Integr. Learn.* 23, 203–218.

Chan, B. K., Stanley, G., Modak, M., Koff, J. L., and Turner, P. E. (2021). Bacteriophage therapy for infections in CF. *Pediatr. Pulmonol.* 56 (1), S4–S9. doi:10.1002/PPUL.25190

Chui, M., Evers, M., Manyika, J., Zheng, A., and Nisbet, T. (2020). *The Bio Revolution Innovations transforming economies, societies, and our lives*. Washington, DC: McKinsey Global Institute. Available at: https://www.mckinsey.com/~/media/mckinsey/industries/life%20sciences/our%20insights/the%20bio%20revolution%20innovations%20transforming%20economies%20societies%20and%20our%20lives/may_2020_mgi_bio_revolution_report.pdf (Accessed November 2, 2022).

de Villiers Scheepers, M. J., Barnes, R., Clements, M., and Stubbs, A. J. (2018). Preparing future-ready graduates through experiential entrepreneurship. *Educ. Train.* 60, 303–317. doi:10.1108/et-11-2017-0167

De Vries, H., Europe, R., Elliott, M. N., Kanouse, D. E., Teleki, S. S., and Health, R. (2008). Using pooled kappa to summarize interrater agreement across many items. *Field Methods* 20 (3), 272–282. doi:10.1177/1525822X08317166

Degli Esposti, M., Morselli, D., Fava, F., Bertin, L., Cavani, F., Viaggi, D., et al. (2021). The role of biotechnology in the transition from plastics to bioplastics: An opportunity to reconnect global growth with sustainability. *FEBS Open Bio* 11, 967–983. doi:10.1002/2211-5463.13119

Desai, A., Tippins, M., and Arbaugh, J. B. (2014). Learning through collaboration and competition: Incorporating problem-based learning and competition-based learning in a capstone course. *Organ. Manag. J.* 11, 258–271. doi:10.1080/15416518.2014.973793

Diep, P., Boucinha, A., Kell, B., Yeung, B. A., Chen, X., Tsyplenkov, D., et al. (2021). Advancing undergraduate synthetic biology education: Insights from a canadian igem student perspective. *Can. J. Microbiol.* 67, 749–770. doi:10.1139/cjm-2020-0549

Dodd, R. H., Pickles, K., Nickel, B., Cvejic, E., Ayre, J., Batcup, C., et al. (2021). Concerns and motivations about COVID-19 vaccination. *Lancet Infect. Dis.* 21, 161–163. doi:10.1016/S1473-3099(20)30926-9

Edwards, D., Perkins, K., Pearce, J., and Hong, J. (2015). *Work integrated learning in STEM in Australian universities*. Melbourne: Australian Council for Educational Research. Available at: https://research.acer.edu.au/higher_education/44 (Accessed January 25, 2022).

El Enshasy, H., Joe Dailin, D., Abd Malek, R., Zahidah Nordin, N., Chin Keat, H., Eyahmalay, J., et al. (2020). "Biocement: A novel approach in the restoration of construction materials," in *Microbial biotechnology approaches to monuments of cultural heritage*. Editors A. Yadav, A. Rastegari, V. Gupta, and N. Yadav (Singapore: Springer). doi:10.1007/978-981-15-3401-0_10

Emory, J. (2014). Understanding backward design to strengthen curricular models. *Nurse Educ.* 39, 122–125. doi:10.1097/NNE.0000000000000034

Falotico, R., and Quatto, P. (2015). Fleiss' kappa statistic without paradoxes. *Qual. Quant.* 49, 463–470. doi:10.1007/s11135-014-0003-1

Farny, N. G. (2018). A vision for teaching the values of synthetic biology. *Trends Biotechnol.* 36, 1097–1100. doi:10.1016/J.TIBTECH.2018.07.019

Ferreira, M., Martinsone, B., and Talić, S. (2020). Promoting sustainable social emotional learning at school through relationship-centered learning environment, teaching methods and formative assessment. *J. Teach. Educ. Sustain.* 22, 21–36. doi:10.2478/JTES-2020-0003

Gamble, N., Patrick, C. J., and Peach, D. (2010). Internationalising work-integrated learning: Creating global citizens to meet the economic crisis and the skills shortage. *High. Educ. Res. Dev.* 29, 535–546. doi:10.1080/07294360.2010.502287

Ganguly, N. K., Croft, S., Singh, L., Sinha, S., and Balganesh, T. (2014). Biomedicine and biotechnology: Public health impact. *Biomed. Res. Int.* 2014, 1–2. doi:10.1155/2014/524785

Geraldi, J., and Lechter, T. (2012). Gantt charts revisited: A critical analysis of its roots and implications to the management of projects today. *Int. J. Manag. Proj. Bus.* 5, 578–594. doi:10.1108/17538371211268889

Hindes, Y. L., Thorne, K. J., Schwean, V. L., and McKeough, A. M. (2008). Promoting intrapersonal qualities in adolescents: Evaluation of rapport's teen leadership breakthrough program. *Can. J. Sch. Psychol.* 23, 206–222. doi:10.1177/0829573508327307

Hoidn, S., and Klemenčič, M. (2020). "The routledge international handbook of student-centered learning and teaching in higher education," in *ROUTLEDGE-INTERNATIONAL-HANDBOOK-STUDENT-CENTERED-LEARNING-TEACHING-HIGHER-EDUCATION-SABINE-HOIDN-MANJA-KLEMEN* (Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK: Routledge). doi:10.4324/9780429259371

Jackson, D. (2014). Employability skill development in work-integrated learning: Barriers and best practice. *Stud. High. Educ.* 40, 350–367. doi:10.1080/03075079.2013.842221

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi:10.1126/SCIENCE.1225829

Kang, K. I., Lee, N., and Joung, J. (2021). Nursing students' experience of online peer tutoring based on the grow model: A qualitative study. *Nurse Educ. Today* 107, 105131. doi:10.1016/J.NEDT.2021.105131

Kirby, D. (2007). Reviewing Canadian post-secondary education: Post-secondary education policy in post-industrial Canada. *Can. J. Educ. Adm. Policy*.

Kumar, R., and Kumar, P. (2017). Future microbial applications for bioenergy production: A perspective. *Front. Microbiol.* 8, 450. doi:10.3389/fmicb.2017.00450

Lange, L., Connor, K. O., Arason, S., Bundgård-Jørgensen, U., Canalis, A., Carrez, D., et al. (2021). Developing a sustainable and circular bio-based economy in EU: By partnering across sectors, upscaling and using new knowledge faster, and for the benefit of climate, environment & biodiversity, and people & business. *Front. Bioeng. Biotechnol.* 8, 4. doi:10.3389/fbioe.2020.00004

MacQueen, K. M., McLellan, E., Kay, K., and Milstein, B. (2016). Codebook development for team-based qualitative analysis. *CAM J.* 10, 31–36. doi:10.1177/1525822x980100020301

Madan, C. R., and Teitge, B. D. (2013). The benefits of undergraduate research: The student's perspective. *Mentor Innov. Scholarsh. Acad. advis.* 15, 1–3. doi:10.26209/MJ1561274

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochem. Med.* 22, 276–282. doi:10.11613/bm.2012.031

Mourshed, M., Farrell, D., and Barton, D. (2012). *Education to employment: Designing a system that works*. Washington, DC: McKinsey Center for Government and a director. Available at: https://www.mckinsey.com/industries/education/our-insights/education-to-employment-designing-a-system-that-works (Accessed September 16, 2022).

New Skills for a European Bioeconomy (Conference Report) (2012). *New Skills for a European Bioeconomy (Conference Report)*. Brussels: European Commission. Available at: https://www.bioways.eu/download.php?f=74&l=en&key=0fcf9f5d0547577191c03e0ec15dc5a7 (Accessed November 2, 2022).

Nichols, T. R., Wisner, P. M., Cripe, G., and Gulabchand, L. (2011). Putting the kappa statistic to use. *Qual. Assur. J.* 13, 57–61. doi:10.1002/QAJ.481

O'Neill, J., and Conzemius, A. (2006). *Power of SMART goals: Using goals to improve student learning*. Bloomington, IN: Solution Tree Press.

Sattler, P., and Peters, J. (2013). *Work-integrated learning in Ontario's postsecondary sector: The experience of Ontario graduates*. Toronto: Higher Education Quality Council of Ontario. Available at: https://heqco.ca/wp-content/uploads/2020/03/WIL_Grad-Follow-up-ENG.pdf (Accesssed November 2, 2022).

Shelton, A. M., Long, S. J., Walker, A. S., Bolton, M., Collins, H. L., Revuelta, L., et al. (2020). First field release of a genetically engineered, self-limiting agricultural pest insect: Evaluating its potential for future crop protection. *Front. Bioeng. Biotechnol.* 7, 482. doi:10.3389/fbioe.2019.00482

Smolke, C. D. (2009). Building outside of the box: iGEM and the BioBricks foundation. *Nat. Biotechnol.* 2712 (27), 1099–1102. doi:10.1038/nbt1209-1099

Vieira, S. M., Kaymak, U., and Sousa, J. M. C. (2010). "Cohen's kappa coefficient as a performance measure for feature selection," in Proceedings of the International Conference on Fuzzy Systems, Barcelona, Spain, 18-23 July 2010. doi:10.1109/FUZZY.2010.5584447

Vilanova, C., and Porcar, M. (2014). iGEM 2.0—refoundations for engineering biology. *Nat. Biotechnol.* 325 (32), 420–424. doi:10.1038/nbt.2899

Wetzel, S., Duchesne, L. C., and Laporte, M. F. (2006). "The way forward," in *Bioproducts from Canada's forests* (Dordrecht: Springer), 211–223. doi:10.1007/1-4020-4992-7_13

Williams, D. A. (2014). Curing genetic disease with gene therapy. *Trans. Am. Clin. Climatol. Assoc.* 125, 122–129.

Wright, W. (2020). The iGEM entrepreneurship program — iGEM blog. Available at: https://blog.igem.org/blog/2020/3/4/the-igem-entrepreneurship-program (Accessed September 16, 2022).

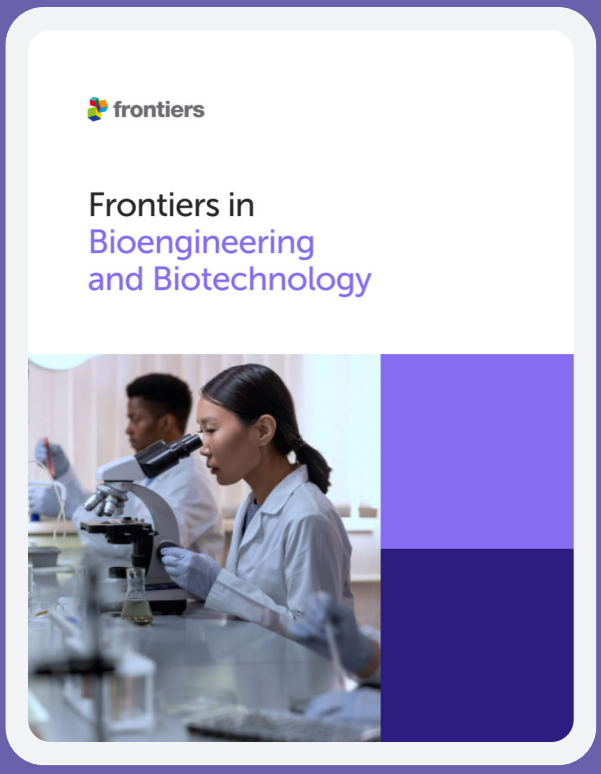
Frontiers in
Bioengineering and Biotechnology

Accelerates the development of therapies, devices, and technologies to improve our lives

A multidisciplinary journal that accelerates the development of biological therapies, devices, processes and technologies to improve our lives by bridging the gap between discoveries and their application.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact