Synthetic Biology, 2020, 5(1): ysaa019

doi: 10.1093/synbio/ysaa019
Advance Access Publication Date: 14 October 2020
Research Article

Modular cell-free expression plasmids to accelerate biological design in cells

Ashty S. Karim [1,2,3,†] Fungmin (Eric) Liew,[4,†] Shivani Garg,[4] Bastian Vögeli [1,2,3] Blake J. Rasor [1,2,3] Aislinn Gonnot,[4] Marilene Pavan,[4] Alex Juminaga,[4] Séan D. Simpson,[4] Michael Köpke,[4,*] and Michael C. Jewett[1,2,3,5,*]

[1]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA, [2]Chemistry of Life Processes Institute, Northwestern University, Evanston, IL 60208, USA, [3]Center for Synthetic Biology, Northwestern University, Evanston, IL 60208, USA, [4]LanzaTech Inc, Skokie, IL 60077, USA and [5]Simpson Querrey Institute, Northwestern University, Chicago, IL 60611, USA

[†]These authors contributed equally to this work.
*Corresponding authors: E-mails: m-jewett@northwestern.edu and Michael.Koepke@lanzatech.com

Abstract

Industrial biotechnology aims to produce high-value products from renewable resources. This can be challenging because model microorganisms—organisms that are easy to use like *Escherichia coli*—often lack the machinery required to utilize desired feedstocks like lignocellulosic biomass or syngas. Non-model organisms, such as *Clostridium*, are industrially proven and have desirable metabolic features but have several hurdles to mainstream use. Namely, these species grow more slowly than conventional laboratory microbes, and genetic tools for engineering them are far less prevalent. To address these hurdles for accelerating cellular design, cell-free synthetic biology has matured as an approach for characterizing non-model organisms and rapidly testing metabolic pathways *in vitro*. Unfortunately, cell-free systems can require specialized DNA architectures with minimal regulation that are not compatible with cellular expression. In this work, we develop a modular vector system that allows for T7 expression of desired enzymes for cell-free expression and direct Golden Gate assembly into *Clostridium* expression vectors. Utilizing the Joint Genome Institute's DNA Synthesis Community Science Program, we designed and synthesized these plasmids and genes required for our projects allowing us to shuttle DNA easily between our *in vitro* and *in vivo* experiments. We next validated that these vectors were sufficient for cell-free expression of functional enzymes, performing on par with the previous state-of-the-art. Lastly, we demonstrated automated six-part DNA assemblies for *Clostridium autoethanogenum* expression with efficiencies ranging from 68% to 90%. We anticipate this system of plasmids will enable a framework for facile testing of biosynthetic pathways *in vitro* and *in vivo* by shortening development cycles.

Key words: cell-free systems; *Clostridium*; plasmids; industrial biotechnology; metabolic engineering.

Submitted: 5 May 2020; **Received (in revised form):** 18 September 2020; **Accepted:** 22 September 2020

© The Author(s) 2020. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

1. Introduction

Industrial biotechnology seeks to produce chemical products from inexpensive and prevalent feedstocks, such as lignocellulosic biomass and syngas (1–3). While most synthetic biologists work with easy to use model organisms like *Escherichia coli* and *Saccharomyces cerevisiae*, these organisms can be limited by accessible feedstocks (e.g., they do not naturally possess metabolic pathways to access the carbon in syngas), products, and stable operating environments in which to work. Thus, researchers try to harness non-model organisms of diverse genera for commercial deployment that are more amenable to utilizing these types of feedstocks and capable of unique biochemical transformations (4). *Clostridium* represents one such genus, which includes the gas-fermenting, acetogenic *Clostridium autoethanogenum* (5–7). Unfortunately, these anaerobic organisms grow more slowly than model microbes, and genetic tools for engineering them are still developing.

Advances in cell-free synthetic biology allow for characterization of non-model organisms (8–10) and biosynthetic pathways *in vitro* (11–14). Without needing to construct new DNA elements for each pathway design, metabolic pathways can be built and assembled *in vitro* by using cell-free gene expression (CFE) (15) (13, 16–20). This approach benefits from the ability to test more enzyme variants and defined reaction conditions as compared to conventional workflows, leading to shorter engineering cycles to down-select beneficial pathways for cellular biochemical production (21). Cell-free pathways can be constructed by modular mixing of lysates enriched with pathway enzymes (13), while cellular expression requires DNA assembly of genes into operons. Additionally, the specialized plasmids for *Clostridium* expression and those for CFE are not inherently compatible, for example, requiring different promoters (cell-free expression typically relies on the orthogonal T7 system) and additional elements such as a Gram-positive origin of replication, specific antibiotic cassettes, and a low GC content (22). This means *Clostridium*-optimized DNA for successful pathway designs identified *in vitro* must be separately synthesized and cloned prior to transformation in *Clostridium*, adding several weeks of effort and considerable costs. Streamlining this process would increase the ability to engineer non-model organisms for metabolic engineering applications.

In this work, we present a modular plasmid system on the basis of standard cell-free vector pJL1 (10, 19, 23, 24) and universal *Clostridium* shuttle vector system pMTL80000 (22) to rapidly bridge cell-free prototyping efforts and strain engineering in *C. autoethanogenum* and reduce the overall engineering cycle time. Engineering compatible expression plasmids requires fine tuning to minimize impacts on the genetic context of open reading frames, particularly around the critically important ribosome binding site (RBS) (13, 19). First, we designed several plasmid architectures that resemble our top-performing cell-free expression vector, pJL1 (Addgene #69496), and would enable flexible arrangement of genes and promoters (Golden Gate assembly (25) compatibility) for expression in *C. autoethanogenum*. Next, we validated that these new vectors are sufficient for CFE of biosynthetic enzymes that are functionally active. Then, we demonstrated DNA assembly efficiency ranging from 68% to 90% when assembling up to six parts for *C. autoethanogenum* expression. We finally showed automation of the whole workflow on two different automation systems. This modular 'cell-free to *Clostridium*' vector system along with high-throughput and automatable workflows will accelerate strain development efforts for *C. autoethanogenum* or other *Clostridium* species. Principles learned here, or perhaps the vectors themselves, could also accelerate biological design in other organisms, by decreasing delays in the transition between cell-free prototyping and cellular validation.

2. Materials and methods

2.1 Strains and plasmids

For generation of the 'cell-free to *Clostridium*' vector system and cloning, *E. coli* strain TOP10 (Invitrogen) was used. First, the counter-selectable marker *ccdB* (flanked with BsaI recognition sites) was cloned into pMTL82251 and pMTL83151 (22) to generate the recipient *Clostridium* expression vectors (pCExpress). The construction of vectors pD2 and pD4 involved TOPO (Invitrogen) cloning of terminator and promoter parts (flanked with BsaI recognition sites) amplified or synthesized by Joint Genome Institute (JGI) into the plasmid, pCR-blunt (Invitrogen). The 'cell-free to *Clostridium*' vectors were derived from pJL1 plasmid (Addgene #69496), modified in the T7 promoter region to contain a BsaI recognition site between the RBS and START codon in three variations to generate pD1, pD3 and pD5. All recipient and donor vectors were verified by DNA sequencing.

DNA codon-optimized genes for *C. autoethanogenum* were generated using LanzaTech's in-house codon optimization software. *Escherichia coli* adapted sequences were generated using codon optimization tools from Twist Biosciences (CA, USA). Genes of interest were provided by JGI in the 'cell-free to *Clostridium*' vectors pD1, pD3 and pD5. All vector DNA sequences used in this study are listed in Supplementary Table S1, and all DNA parts are listed in Supplementary Table S2. The 58 modular vectors containing parts from Supplementary Table S2 are listed in Supplementary Table S3. The biosynthetic genes used in cell-free assays are listed in Supplementary Table S4, and those used in GG assembly are listed in Supplementary Table S5.

2.2 Cell-free assays

All cell extracts for CFE were prepared with *E. coli* BL21 Star(DE3) (NEB) (21). These cells were grown, harvested, lysed and prepared using previously described methods (19, 26). CFE reactions were performed to express each enzyme individually using a modified PANOx-SP system at 15- or 30-μl volumes in 2-ml Eppendorf tubes as described in previous publications (27, 28). Protein measurements were taken after 20 h. Active superfolder green fluorescent protein (sfGFP) protein yields were quantified by measuring fluorescence. To do this, 2 μl of the total CFE reaction was added in the middle of the flat bottom of 96-well half area black plates (Costar 3694; Corning Incorporated, Corning, NY, USA). sfGFP was excited at 485 nm while measuring emission at 528 nm with a 510 nm cutoff filter. The fluorescence of sfGFP was converted to concentration (μg/ml) according to a standard curve (29). All other proteins were measured using CFE reactions with radioactive ^{14}C-Leucine (10 μM) supplemented for incorporation during protein production. We used trichloroacetic acid (TCA) to precipitate radioactive protein samples. Radioactive counts from TCA-precipitated samples were measured by liquid scintillation to then quantify soluble and total yields of each protein produced as previously reported (MicroBeta2; PerkinElmer) (27, 30).

The cell-free activity assays were run at 15-μl volumes in 1.5-ml Eppendorf tubes. All enzyme-enriched lysates (via CFE reaction) were added at 0.4 μM concentration of enriched enzyme as determined by C^{14} measurements with the balance

being 'blank CFE' reaction (no DNA added) up to 50% of the total reaction volume. Small molecules were added to achieve final concentrations of 120 mM glucose, 3 mM NAD^+, 5 mM CoA, 100 mM BisTris pH 7, 8 mM magnesium acetate salts, 0.1 U/μl catalase, in the reaction. Reactions were run for 20 h, quenched with 5% trichloroacetic acid and measured via High-Performance Liquid Chromatography as described previously (13, 19).

2.3 Golden gate assembly using manual workflow

Two- to six-part DNA assemblies were performed using GeneArt Type IIs (BsaI) assembly kit (Invitrogen, CA, USA). Specifically, 75 ng of recipient vector was used. Other parts (pD1, pD2, pD3, pD4 and pD5) were added in 1:1 molar ratio with respect to the recipient vector along with the GeneArt Type IIs enzyme mix. The reaction was then incubated in a thermocycler (37°C for 1 min, 16°C for 1 min, cycled 30×, followed by cooling at 4°C). Afterwards, the assembly mixture was transformed into *E. coli* Top 10 chemically competent cells (ThermoFisher Scientific, CA, USA), and plated onto LB agar containing appropriate antibiotics. Resulting colonies were screened via PCR for presence of the parts cloned, followed by sequence confirmation via Next-Generation Sequencing.

2.4 Golden gate assembly using automated workflow

Two automated assembly workflows were developed, either using the liquid handling robot Hamilton STARLet or the Labcyte Echo 525 (31). The assembly reactions were carried out with the final concentration for each individual DNA part was 2 nM (25). The assembly reaction volume for Hamilton STARLet was 20 μl prepared as follows: 2 μl of each DNA part (10 nM), 10 μl GeneArt Type IIs Assembly Kit BsaI (Invitrogen A15917) and deionized water to a total of 20 μl. When using the Labcyte Echo 525, the reaction volume was downsized to a final volume of 2 μl. All DNA samples were quantified by absorbance at 280 nm, employing a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). Reactions were incubated in an INHECO heat block using the following parameters: 37°C for 2 h, 50°C for 5 min, 80°C for 10 min, then stored at −20°C until transformation. Transformations were also performed using the INHECO blocks: 2 μl of each reaction mix was added to 20 μl of Invitrogen One Shot Top10 chemically competent cells (C404003) and incubated for 20 min at 4°C. Cells were then heat-shocked at 60°C for 45 s, then recovered for 2 min at 4°C. Afterwards, 180 μl of Super Optimal Broth (Invitrogen 15544034) with catabolite repression (SOC) media was added to cell mixtures, and cells were recovered at 37°C for 2 h. Finally, 7 μl of each transformation reaction consisting of undiluted culture volume was plated on Lennox Lysogeny Broth (LB) + agar plates containing the appropriate antibiotic and incubated overnight at 37°C. We randomly chose two colonies to sequence throughout the assembled regions. NGS sequencing confirmed that more than 90% of the clones screened showed complete assemblies.

2.5 Data and materials availability

Data or unique materials presented in the manuscript may be available from the authors upon reasonable request and through a materials transfer agreement.

3. Results and discussion

3.1 Design of a modular 'cell-free to *Clostridium*' vector system

Our goal was to develop a DNA vector system that would enable easy exchange of DNA between cell-free plasmids and cellular plasmids. This cell-free to *Clostridium* framework would minimize repetitive DNA synthesis and subcloning and enable reuse of DNA, allowing for facile and rapid testing of biosynthetic pathways *in vitro* and *in vivo*. Traditional approaches required optimization and synthesis of candidate genes for cell-free expression and cell-free testing followed by separate optimization and synthesis of genes for *Clostridium* expression (**Figure 1a**, red), taking significant time and slowing down research efforts. The workflow presented here allows for a single round of codon optimization and synthesis for *Clostridium* that can be used for both cell-free testing and pathway assembly in *Clostridium* (**Figure 1a**, blue), reducing the time and cost of research and development in these complex organisms. Specifically, there is a 50% reduction in DNA synthesis time and cost as well as only one round of cloning into cell-free/donor vectors prior to *in vivo* vector assembly. To achieve this goal, CFE expression plasmids were modified via addition of Golden Gate (GG) sites so that these can be used directly for multiple DNA parts assembly directly into *Clostridium* expression vectors (**Figure 1b**). With such a system in place, genes could be ordered once with *Clostridium* codon-adapted sequences in these plasmids, prototyped in cell-free reactions using these plasmids, and then the best performing gene variants could be assembled from the same plasmids into *in vivo* expression plasmids via one-step GG assembly.

As a starting point, we designed six total vectors. Three vectors, pD1, pD3 and pD5, were constructed by adding GG sites (BsaI recognition sites) within the T7 promoter region of the pJL1 vector (Addgene #69496), a standard CFE expression plasmid. These vectors were designed to serve as gene donor vectors for assembly in a new recipient vector based on pMTL80000 universal *Clostridium* expression vectors (22) with the addition of two GG sites flanking a *ccdB* survival gene along with *Clostridium* promoter and terminator flanking the GG sites. We also constructed pD2 and pD4 to serve as promoter-terminator donor vectors. This system of six vectors (five donor vectors and one recipient vector) would allow for *in vitro* expression of genes using pD1, pD3 and pD5, followed by one-step assembly of up to six DNA parts (inserts supplied by pD1, pD2, pD3, pD4 and pD5) directly into our *Clostridium* expression vector. We note that different combinations of these vectors can be used to assemble one-gene insertions (Supplementary Figure S1A) or two-gene insertions (Supplementary Figure S1B) when fewer genes are desired. It is also possible to combine more than three genes in an expression operon by using multicistronic donor vectors (Supplementary Figure S1C).

Utilizing JGI's gene synthesis program, the GG system was expanded to create recipient and donor vectors with varying promoters, such as P_{fdx} (17), P_{pta} (18), P_{pfor} (19) and P_{wl} (18). Additionally, GG sites were varied in recipient vectors to allow assembly of anywhere between two to six parts with varying promoters, resulting in a total of 58 modular vectors (Supplementary Figure S1D and Table S3). The variety of assembly options using different DNA parts (Supplementary Table S2) increases the versatility of this vector system.

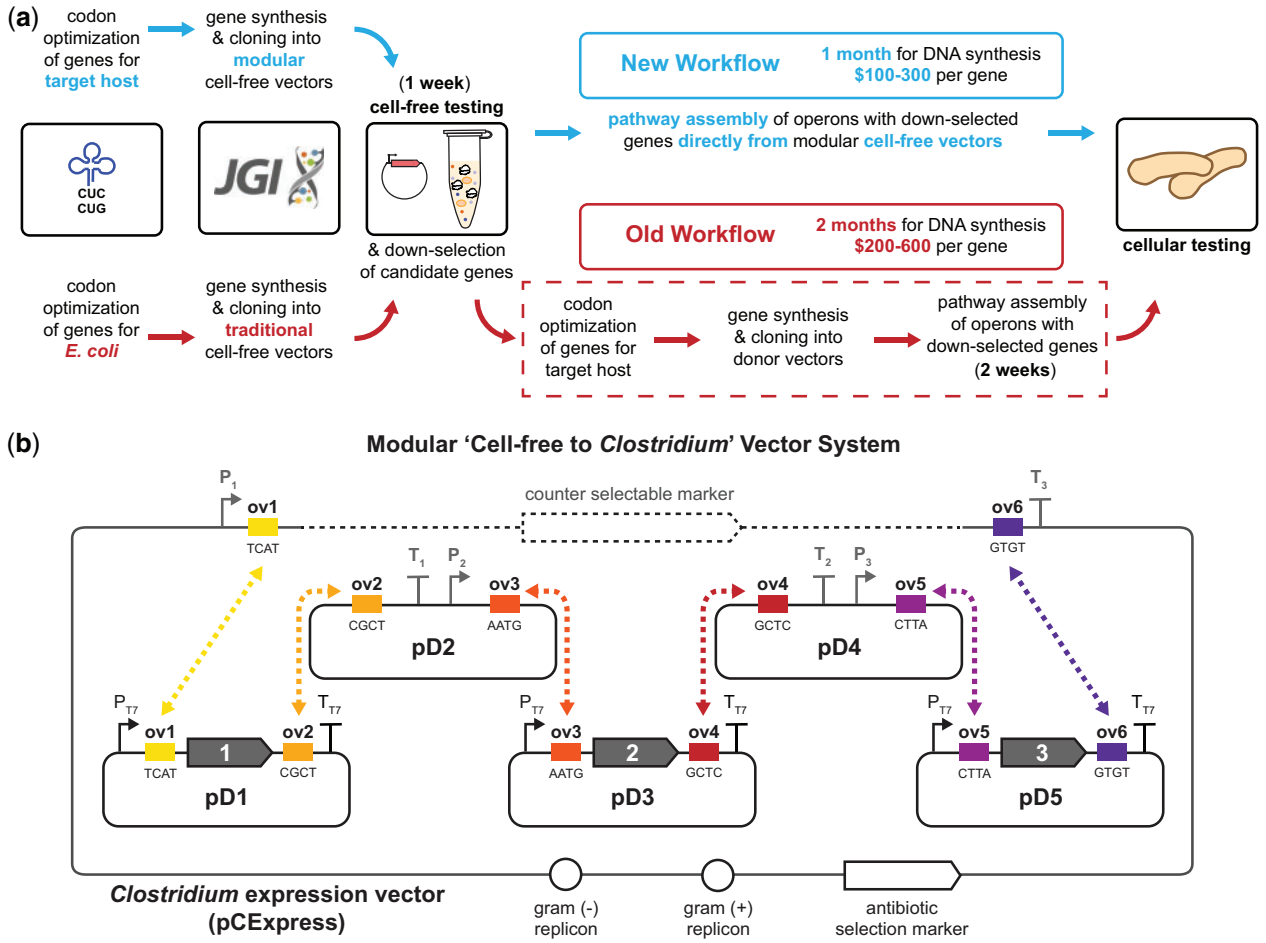

Figure 1. A framework for a modular 'cell-free to *Clostridium*' vector system that enables seamless assembly of cell-free vectors into a *Clostridium* expression vector. (**a**) A schematic representation of how information between *in vitro* and *in vivo* needs is used to design DNA sequences, JGI facilities can construct DNA designs and DNA materials can be used in both *in vitro* and *in vivo* experiments. Approximate times are noted for cell-free testing, *in vivo* construct assembly and DNA synthesis (for new and old workflows). Costs associated to DNA synthesis are calculated with assumptions of 0.1 USD/bp and 1–3 kb genes. (**b**) The architecture of the modular vector system is shown. Cell-free vectors are made compatible for assemblies by adding unique overhang (ov) sites generated from BsaI digests.

3.2 Evaluation of the vector system for CFE

The highest-yielding cell-free systems take advantage of T7 RNA polymerization and are substantially affected by changes in plasmid architecture (32, 33). For example, our previous work using the pJL1 vector that leverages T7 RNA polymerase to make mRNA achieved protein yields of ~2.7 g/l sfGFP (34). In order to adapt our robust pJL1 vector for GG compatibility, we chose to test the insertion of three BsaI site designs in pJL1 (Figure 2a; Supplementary Table S2). Specifically, GG sites with sequences TCAT, AATG or CTTA were introduced between the RBS and start codon, which increased the spacer length between these two elements by one to three nucleotides. This created three distinct donor vectors each with three possible BsaI cut sites. We evaluated each of these nine designs in a CFE reaction based on the PANOx-SP system to produce sfGFP to first assess the impact of GG sites on protein expression. After 20 h, cell-free reactions produced comparable or slightly greater concentrations of sfGFP than the unaltered pJL1 plasmid (**Figure 2b**). For the donor vectors to be compatible, we had to choose the same variant for each of the three vectors. Thus, we chose variant 2 as the highest-performing set. Variant 2 GG vectors were validated further with expression of the enzymes phosphotransbutyrylase (Ptb) and butyrate kinase (Buk) from butyric acid metabolism in *Clostridium acetobutylicum* ATCC 824 (**Figure 2c**). This experiment highlights the importance of genetic context for expression of different proteins, yielding variable amounts of protein for Ptb and Buk despite nearly identical expression of sfGFP.

We next evaluated whether *C. autoethanogenum*-optimized genes would be sufficiently expressed in our established cell-free assays. To test this, we constructed a panel of 16 *Clostridium*-optimized biosynthesis genes (Supplementary Table S4) related to acid/alcohol fermentation from a variety of organisms in pD1, pD3 and pD5. Several of these genes were used in a previous study for the production of butanol from acetyl-CoA, while the rest were identified through sequence similarity to these genes. After a 20-h CFE reaction, we observed a range of expression all significantly lower (10-fold) than what we saw for *E. coli*-codon-optimized sfGFP, Ptb or Buk (**Figure 2d**). However, expressing full-length enzymes at concentrations >1 μM provides at least 0.1 μM enzyme after dilution upon pathway assembly *in vitro* which we have found to be sufficient for prototyping (13). Thus, we proceeded to test whether *C. autoethanogenum*-optimized sequences were sufficient for prototyping biosynthetic pathways. First, we ran the 16 enzymes on a protein gel via SDS-PAGE followed by ^{14}C autoradiography to confirm that indeed those enzymes that expressed, expressed

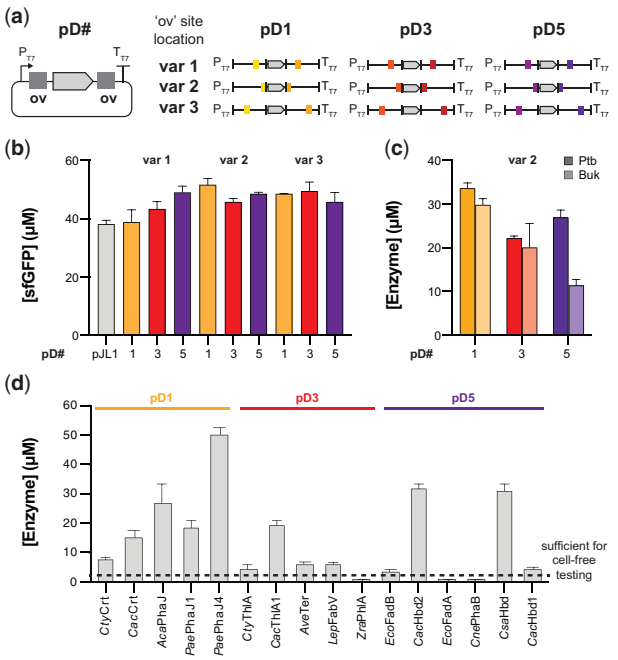

Figure 2. Cell-free expression of Golden Gate compatible vectors is sufficient for prototyping biosynthetic enzymes. (**a**) A schematic representation of three variants (change in where BsaI sites are located) of each of the three donor plasmids in CFE using a reporter sfGFP is shown. (**b**) sfGFP concentration was measured by fluorescence at 20 h after cell-free reaction start. Data are shown for $n = 2$ independent experiments with average error. (**c**) Protein concentration at 20 h for Ptb and Buk enzymes expressed from each of the three donor vectors was measured via C^{14}-leucine incorporation. Data are shown for n = 2 independent experiments with average error. (**d**) Sixteen enzymes of interest for acid and alcohol fermentation were codon optimized for *C. autoethanogenum* and cloned into pD1, pD2 and pD3. Protein expression was measured at 20 h for $n = 3$ independent experiments. All error bars represent 1 standard deviation.

at full length (Supplementary Figure S3A). Decreasing the reaction temperature to slow down protein translation and folding and altering the DNA template sequence (e.g. RBS, coding sequence) can be pursued when important homologs are not expressed in soluble form. Then, we used these enzymes to construct biosynthetic pathways to produce butanoic acid to test the activity of these enzymes. This pathway runs from glucose to acetyl-CoA using native metabolism present in *E. coli* crude lysates followed by four enzymatic steps to produce butanoic acid. Mixing 12 combinations of the enzymes from **Figure 2d** with glucose and cofactors, each combination produced >7 mM butanoic acid (Supplementary Figure S3B). These enzymes can be studied further using more detailed activity assays via purification and defined substrates. For enzymes that are not expressed well, separately expressing the *E. coli*-codon-optimized versions can improve enzyme expression (Supplementary Figure S2 and Table S4). Although the soluble protein yields are generally lower using *C. autoethanogenum* (31% GC content) codon-optimized sequences in *E. coli*-based (50% GC content) cell extract, our data suggest that *C. autoethanogenum*-optimized enzyme sequences are active in *E. coli* crude lysates. With GG-compatible vectors at hand, we next sought construct *in vivo* expression plasmids.

3.3 Six-part DNA assembly from CFE vectors into *Clostridium* expression plasmid

Once the pD vectors were successfully validated in CFE, these modified vectors were then used for testing the efficiency of multiple-part assembly directly into a *Clostridium* expression vector with a variety of biosynthetic genes. Specifically, we carried out a six-part GG assembly that contained: (i) a recipient vector based on pMTL8315 backbone containing a promoter (P1) and terminator (T3) flanking the two GG sites (pCExpress), (ii) pD2 and pD4, both containing terminator and promoter combinations (i.e. T1-P2 and T2-P3) and (iii) pD1, pD3 and pD5 containing gene 1, gene 2 and gene 3 (**Figure 3a**; Supplementary Table S1). The assembly mixture was transformed into our *E. coli* cloning strain and six colonies were picked and genotyped by PCR which indicated 90–100% of the picked colonies had plasmids with all six parts correctly assembled (**Figure 3b**). These were confirmed via sequencing. The six-part assembly was validated using a different set of genes and promoter-terminator combinations for at least five additional designs (Supplementary Table S5). In a total of 20 manual assemblies were carried out with efficiency ranging from 70% to 95%.

Assembled constructs can then be transformed into Clostridia to test for biosynthetic pathway activity. We previously showed that optimization of pathways in an *E. coli* cell-free system could inform cellular design in Clostridia (13). This work showed positive correlations between cell-free activity data and *in vivo* expression. Down-selecting a fraction of the >200 pathway combinations tested *in vitro* for constructing in Clostridia saved more than 6 months of research efforts. Combining this modular vector system with Clostridia-based cell-free expression (10) should lead to more positive correlations and a streamlined research pipeline.

3.4 Workflow automation

Workflow automation can improve throughput and reliability. CFE reactions can routinely be performed using liquid-handling robotics (35, 36). These reactions can be scaled down to 2 μl without significant changes in protein expression (37). In addition, GG assembly for *in vivo* expression can also be automated. After demonstrating successful assembly of up to six DNA parts using a manual workflow, we then developed an automated workflow to increase our DNA assembly throughput (**Figure 3a**). Due to the complexity of biological systems, it is often necessary to test a large number of enzyme homologs along with different promoters to obtain an optimal engineering solution. Indeed, testing just five homologs and three promoters for a three-gene operon would yield 3,375 different permutations. However, this experimental throughput is difficult and laborious when using manual techniques and procedures. Automated, well-informed designs help to increase the number of designs that can be generated, the speed these designs can be generated, and it helps to narrow down the design space prioritizing the best candidates to be built and tested, saving lab resources (38). In order to increase throughput, efficiency and accuracy of our strain engineering pipeline, free up researchers from repetitive tasks, and increase results reproducibility, we validated a Golden Gate DNA assembly automated protocol on two automation systems. Both the design of constructs and the worklists to run the experiments were generated by J5 software (39, 40). We assembled three- to six-part GG assemblies using both a Hamilton STARLet liquid-handling robot and a Labcyte Echo 525 acoustic liquid-handling robot with greater than 90% efficiencies.

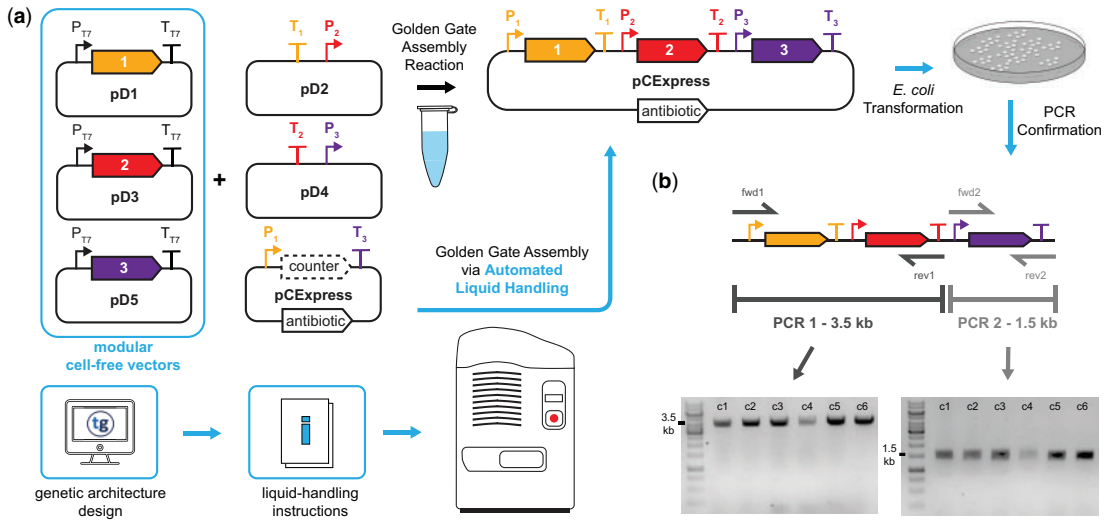

Figure 3. Golden Gate assembly of a three-gene construct using compatible cell-free vector system. **(a)** A schematic representation of our Golden Gate assembly workflow including automated assembly consisting of computational design of plasmids, liquid-handling instructions, plasmid assembly and plasmid confirmation. **(b)** PCR confirmation of plasmid assembly in six colonies containing the constructed *Clostridium* expression vector.

4. Conclusion

In this study, we describe a set of modular vectors for both CFE and cloning into *Clostridium* expression plasmids. This framework allows facile testing of biosynthetic pathways *in vitro* and *in vivo* for shorter engineering cycles and enables an improved workflow between our *in vitro* team, our *in vivo* team, and JGI, without lengthy and costly re-synthesis and/or subcloning. The 'cell-free to *Clostridium*' vector system is easy to use for Golden Gate assembly of up to six parts (three open reading frames with unique promoter and terminator sequences) at once with up to 90% efficiencies and feeds directly into JGI's Community Science Program platform. For longer operons, genes can be sequentially located on each of the CFE vectors (pD1, pD3 and pD5). These vectors along with laboratory automation have already increased the speed and efficiency of our workflows and will continue to facilitate the ability to prototype biosynthetic pathways *in vitro* followed by *in vivo* cloning pipelines. Standardization of these vector systems allows for new simplified workflows. The pJL1 cell-free vector and variants thereof are routinely used in multiple bacterial cell-free systems (i.e. *E. coli* (19), *Clostridium* (10), *Pseudomonas* (23), *Streptomyces* (24, 41), *Vibrio natriegens* (42, 43)). In addition, the pMTL vector system has been demonstrated in several Clostridia species (i.e. *autoethanogenum*, *ljungdahlii*, *acetobutylicum*, *beijerinckii*, *difficile*, *sporogeneses*, *perfringens*, *pasteurianum*, *tyrobutyricum*) as well as other Gram-negative and Gram-positive model organisms such as *E. coli* and *Bacillus* (22, 44). Taken together, the breadth of bacterial cell-free systems that can use the pJL1 vector and ubiquity of Golden Gate cloning suggests broad applicability of our plasmid vector system. Looking forward, we anticipate this system of vectors will allow researchers to integrate more *in vitro* prototyping practices into their existing workflows across multiple organisms to speed up metabolic engineering efforts.

Supplementary data

Supplementary Data are available at SYNBIO Online.

Author contributions

A.S.K., F.E.L., A.J., M.K. and M.C.J. conceived the study. A.S.K., B.V. and B.J.R. performed cell-free experiments. F.E.L., S.G., A.G. and M.P. performed in vivo experiments. S.D.S., M.K. and M.C.J. provided supervision. All authors wrote the manuscript.

Acknowledgements

We would like to acknowledge Jan-Fang Chen, Robert Evans, Yasuo Yoshikuni and Miranda Harmon-Smith from the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, for their DNA synthesis, DNA assembly and support in this work as part of the Community Science Program (CSP).

Funding

This work was supported by the U.S. Department of Energy (DOE) Biological and Environmental Research Division (BER), Genomic Science Program (GSP) for funding of this project under Contract No. DE-SC0018249. We also thank the Joint Genome Institute Community Science Program Project 503280. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We also thank the following investors in LanzaTech's technology: BASF, CICC Growth Capital Fund I, CITIC Capital, Indian Oil Company, K1W1, Khosla Ventures, the Malaysian Life Sciences, Capital Fund, L. P., Mitsui, the New Zealand Superannuation Fund, Novo Holdings A/S, Petronas Technology Ventures, Primetals, Qiming Venture Partners, Softbank China and Suncor. B.V. acknowledges support from the SNSF Early Postdoc.Mobility fellowship (P2SKP3_184036).

Conflict of interest statement. A provisional patent has been filed on this vector system (US 62/943,036).

References

1. Clomburg, J.M., Crumbley, A.M. and Gonzalez, R. (2017) Industrial biomanufacturing: the future of chemical production. *Science*, 355, aag0804.
2. Liu, Z., Wang, K., Chen, Y., Tan, T. and Nielsen, J. (2020) Third-generation biorefineries as the means to produce fuels and chemicals from CO2. *Nat. Catal.*, 3, 274–288.
3. Köpke, M. and Simpson, S.D. (2020) Pollution to products: recycling of 'above ground' carbon by gas fermentation. *Curr. Opin. Biotechnol.*, 65, 180–189.
4. Yan, Q. and Fong, S.S. (2017) Challenges and advances for genetic engineering of non-model bacteria and uses in consolidated bioprocessing. *Front. Microbiol.*, 8, 2060.
5. Tracy, B.P., Jones, S.W., Fast, A.G., Indurthi, D.C. and Papoutsakis, E.T. (2012) Clostridia: the importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. *Curr. Opin. Biotechnol.*, 23, 364–381.
6. Lynd, L.R., Guss,A.M., Himmel,M.E., Beri,D., Herring,C., Holwerda,E.K., Murphy,S.J., Olson,D.G., Paye,J., Rydzak,T. et al. (2016) Advances in consolidated bioprocessing using clostridium thermocellum and thermoanaerobacter saccharolyticum. In: C Wittmann and JC Liao (eds). *Industrial Biotechnology: Microorganisms Vol. 1*. Wiley-VCH Verlag GmbH & Co. KGaA, USA, pp. 365–394.
7. Marcellin, E., Behrendorff, J.B., Nagaraju, S., DeTissera, S., Segovia, S., Palfreyman, R.W., Daniell, J., Licona-Cassani, C., Quek, L-e., Speight, R. et al. (2016) Low carbon fuels and commodity chemicals from waste gases—systematic approach to understand energy metabolism in a model acetogen. *Green Chem.*, 18, 3020–3028.
8. Yim, S.S., Johns, N.I., Park, J., Gomes, A.L., McBee, R.M., Richardson, M., Ronda, C., Chen, S.P., Garenne, D., Noireaux, V. et al. (2019) Multiplex transcriptional characterizations across diverse bacterial species using cell-free systems. *Mol. Syst. Biol.*, 15, e8875.
9. Moore, S.J., MacDonald, J.T., Wienecke, S., Ishwarbhai, A., Tsipa, A., Aw, R., Kylilis, N., Bell, D.J., McClymont, D.W., Jensen, K. et al. (2018) Rapid acquisition and model-based analysis of cell-free transcription-translation reactions from nonmodel bacteria. *Proc. Natl. Acad. Sci. U S A*, 115, E4340–E4349.
10. Krüger, A. et al. (2020) Development of a clostridia-based cell--free system for prototyping genetic parts and metabolic pathways. *Metab. Eng.*, 62, 95–105. doi: 10.1016/j.ymben.2020.06.004
11. Moore, S.J., MacDonald, J.T. and Freemont, P.S. (2017) Cell-free synthetic biology for in vitro prototype engineering. *Biochem. Soc. Trans.*, 45, 785–791.
12. Jiang, L., Zhao, J., Lian, J. and Xu, Z. (2018) Cell-free protein synthesis enabled rapid prototyping for metabolic engineering and synthetic biology. *Synth. Syst. Biotechnol.*, 3, 90–96.
13. Karim, A.S., Dudley, Q.M., Juminaga, A., Yuan, Y., Crowe, S.A., Heggestad, J.T., Garg, S., Abdalla, T., Grubbe, W.S., Rasor, B.J. et al. (2020) In vitro prototyping and rapid optimization of biosynthetic enzymes for cell design. *Nat. Chem. Biol.*, 16, 912–919.
14. Liu, Z., Zhang, Y., Jia, X., Hu, M., Deng, Z., Xu, Y. and Liu, T. (2017) In vitro reconstitution and optimization of the entire pathway to convert glucose into fatty acid. *ACS Synth. Biol.*, 6, 701–709.
15. Silverman, A.D., Karim, A.S. and Jewett, M.C. (2020) Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.*, 21, 151–170.
16. Grubbe, W.S., Rasor, B.J., Krüger, A., Jewett, M.C. and Karim, A.S. (2020) Cell-free styrene biosynthesis at high titers. *bioRxiv*, 2020.2003.2005.979302, doi: 10.1101/2020.03.05.979302
17. Dudley, Q.M., Karim, A.S., Nash, C.J. and Jewett, M.C. (2020) Cell-free prototyping of limonene biosynthesis using cell-free protein synthesis. *BioRxiv*, 2020.2004.2023.057737, doi: 10.1101/2020.04.23.057737
18. Kightlinger, W., Duncker, K.E., Ramesh, A., Thames, A.H., Natarajan, A., Stark, J.C., Yang, A., Lin, L., Mrksich, M., DeLisa, M.P. et al. (2019) A cell-free biosynthesis platform for modular construction of protein glycosylation pathways. *Nat. Commun.*, 10, 5404.
19. Karim, A.S. and Jewett, M.C. (2016) A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. *Metab. Eng.*, 36, 116–126.
20. Dudley, Q.M., Nash, C.J. and Jewett, M.C. (2019) Cell-free biosynthesis of limonene using enzyme-enriched *Escherichia coli* lysates. *Synth. Biol.*, 4, ysz003. doi:10.1093/synbio/ysz003
21. Karim, A.S. and Jewett, M.C. (2018) Cell-free synthetic biology for pathway prototyping. *Methods Enzymol.*, 608, 31–57.
22. Heap, J.T., Pennington, O.J., Cartman, S.T. and Minton, N.P. (2009) A modular system for Clostridium shuttle plasmids. *J. Microbiol. Methods*, 78, 79–85.
23. Wang, H., Li, J. and Jewett, M.C. (2018) Development of a *Pseudomonas putida* cell-free protein synthesis platform for rapid screening of gene regulatory elements. *Synth. Biol.*, 3, ysy003. doi:10.1093/synbio/ysy003
24. Li, J., Wang, H., Kwon, Y.C. and Jewett, M.C. (2017) Establishing a high yielding streptomyces-based cell-free protein synthesis system. *Biotechnol. Bioeng.*, 114, 1343–1353.
25. Engler, C., Kandzia, R. and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, 3, e3647.
26. Kwon, Y.C. and Jewett, M.C. (2015) High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.*, 5, 8663.
27. Jewett, M.C. and Swartz, J.R. (2004) Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.*, 86, 19–26.
28. Jewett, M.C. and Swartz, J.R. (2004) Substrate replenishment extends protein synthesis with an *in vitro* translation system designed to mimic the cytoplasm. *Biotechnol. Bioeng.*, 87, 465–472.
29. Hong, S.H., Ntai, I., Haimovich, A.D., Kelleher, N.L., Isaacs, F.J. and Jewett, M.C. (2014) Cell-free protein synthesis from a release factor 1 deficient *Escherichia coli* activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS Synth. Biol.*, 3, 398–409.
30. Jewett, M.C., Calhoun, K.A., Voloshin, A., Wuu, J.J. and Swartz, J.R. (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol.*, 4, 220.
31. Walsh, D.I. 3rd, Pavan, M., Ortiz, L., Wick, S., Bobrow, J., Guido, N.J., Leinicke, S., Fu, D., Pandit, S., Qin, L. et al. (2019) Standardizing automated DNA assembly: best practices, metrics, and protocols using robots. *SLAS Technol.*, 24, 282–290.
32. Shin, J. and Noireaux, V. (2010) Efficient cell-free expression with the endogenous *E. coli* RNA polymerase and sigma factor 70. *J. Biol. Eng.*, 4, 8.
33. Yeung, E., Dy, A.J., Martin, K.B., Ng, A.H., Del Vecchio, D., Beck, J.L., Collins, J.J. and Murray, R.M. (2017) Biophysical constraints arising from compositional context in synthetic gene networks. *Cell Syst.*, 5, 11–24.e12.
34. Des Soye, B.J., Gerbasi, V.R., Thomas, P.M., Kelleher, N.L. and Jewett, M.C. (2019) A highly productive, one-pot cell-free protein synthesis platform based on genomically recoded *Escherichia coli*. *Cell Chem. Biol.*, 26, 1743–1754.e9.

35. Caschera, F., Karim, A.S., Gazzola, G., d'Aquino, A.E., Packard, N.H. and Jewett, M.C. (2018) High-throughput optimization cycle of a cell-free ribosome assembly and protein synthesis system. *ACS Synth. Biol.*, 7, 2841–2853.
36. Karim, A.S., Heggestad, J.T., Crowe, S.A. and Jewett, M.C. (2018) Controlling cell-free metabolism through physiochemical perturbations. *Metab. Eng.*, 45, 86–94.
37. Marshall, R., Garamella, J., Noireaux, V. and Pierson, A. (2018) High-throughput microliter-sized cell-free transcription-translation reactions for synthetic biology applications using the Echo® 550 Liquid Handler. *Labcyte Appli. Note*, App-G124. San Jose, CA: Labcyte Inc.
38. Densmore, D.M. and Bhatia, S. (2014) Bio-design automation: software + biology + robots. *Trends Biotechnol.*, 32, 111–113.
39. Hillson, N.J., Rosengarten, R.D. and Keasling, J.D. (2012) j5 DNA assembly design automation software. *ACS Synth. Biol.*, 1, 14–21.
40. Hillson, N.J. (2014) j5 DNA assembly design automation. *Methods Mol. Biol.*, 1116, 245–269.
41. Xu, H., Liu, W.Q. and Li, J. (2020) Translation related factors improve the productivity of streptomyces-based cell-free protein synthesis system. *ACS Synth. Biol.*, doi: 10.1021/acssynbio.0c00140
42. Des Soye, B.J., Davidson, S.R., Weinstock, M.T., Gibson, D.G. and Jewett, M.C. (2018) Establishing a high-yielding cell-free protein synthesis platform derived from vibrio natriegens. *ACS Synth. Biol.*, 7, 2245–2255.
43. Wiegand, D.J., Lee, H.H., Ostrov, N. and Church, G.M. (2018) Establishing a cell-free vibrio natriegens expression system. *ACS Synth. Biol.*, 7, 2475–2479.
44. Minton, N.P., Ehsaan, M., Humphreys, C.M., Little, G.T., Baker, J., Henstra, A.M., Liew, F., Kelly, M.L., Sheng, L., Schwarz, K. et al. (2016) A roadmap for gene system development in Clostridium. *Anaerobe*, 41, 104–112.