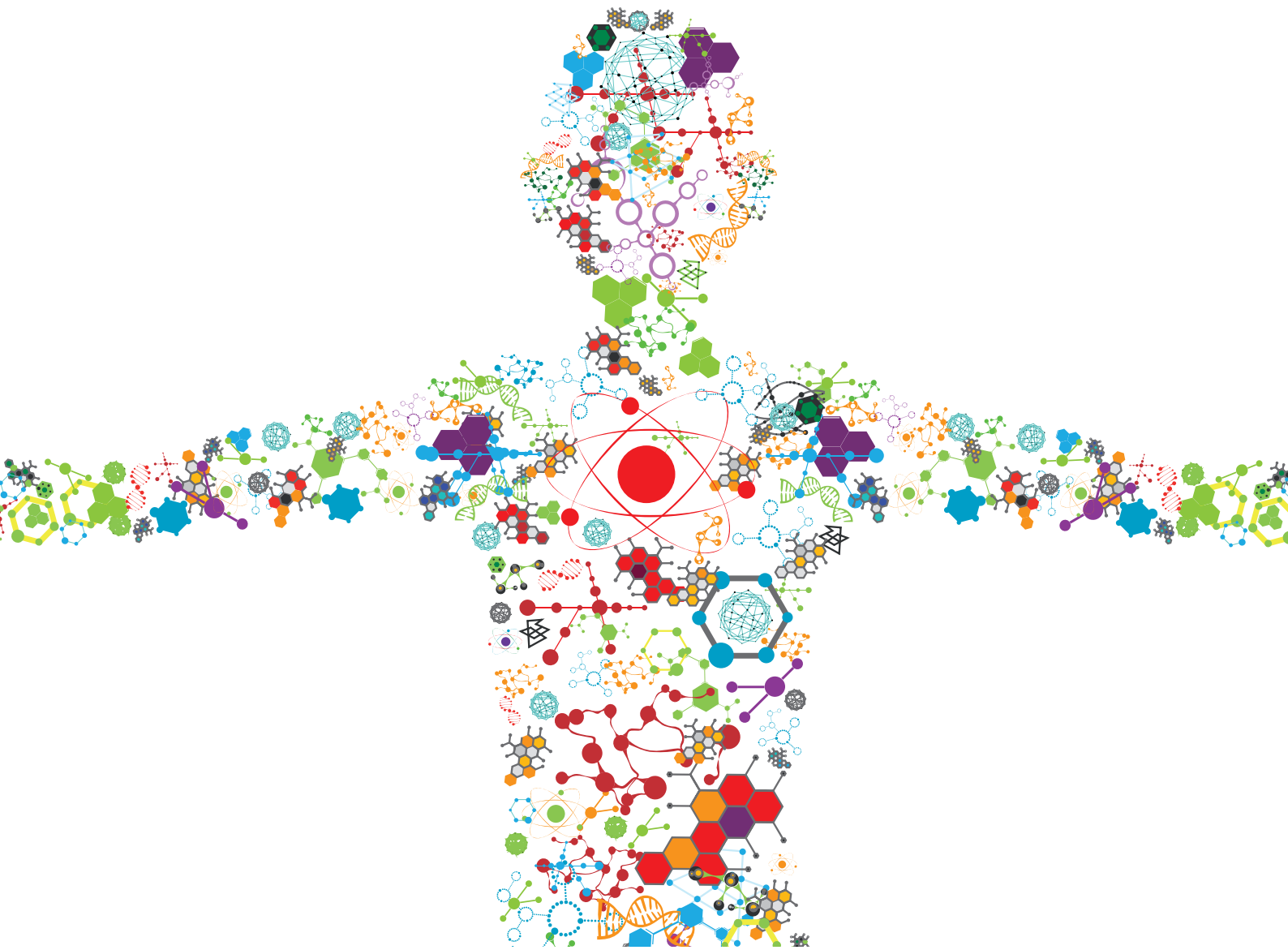
CELL-FREE SYNTHETIC BIOLOGY

EDITED BY: Jian Li, Yong-Chan Kwon, Yuan Lu and Simon J. Moore
PUBLISHED IN: Frontiers in Bioengineering and Biotechnology



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-080-2
DOI 10.3389/978-2-88974-080-2

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

CELL-FREE SYNTHETIC BIOLOGY

Topic Editors:
Jian Li, ShanghaiTech University, China
Yong-Chan Kwon, Louisiana State University, United States
Yuan Lu, Tsinghua University, China
Simon J. Moore, University of Kent, United Kingdom

Citation: Li, J., Kwon, Y.-C., Lu, Y., Moore, S. J., eds. (2022). Cell-Free Synthetic Biology. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-080-2

Table of Contents

05 ***Editorial: Cell-Free Synthetic Biology***
Jian Li, Yong-Chan Kwon, Yuan Lu and Simon J. Moore

08 ***Bottom-Up Construction of Complex Biomolecular Systems With Cell-Free Synthetic Biology***
Nadanai Laohakunakorn, Laura Grasemann, Barbora Lavickova, Grégoire Michielin, Amir Shahein, Zoe Swank and Sebastian J. Maerkl

34 ***O_2-Tuned Protein Synthesis Machinery in* Escherichia coli*-Based Cell-Free System***
Xiaomei Lin, Caijin Zhou, Songbiao Zhu, Haiteng Deng, Jisong Zhang and Yuan Lu

45 ***Biological Materials: The Next Frontier for Cell-Free Synthetic Biology***
Richard J. R. Kelwick, Alexander J. Webb and Paul S. Freemont

60 ***Establishing a Eukaryotic* Pichia pastoris *Cell-Free Protein Synthesis System***
Lingkai Zhang, Wan-Qiu Liu and Jian Li

70 ***Multi-Enzyme Assembly on T4 Phage Scaffold***
Jinny L. Liu, Daniel Zabetakis, Joyce C. Breger, George P. Anderson and Ellen R. Goldman

76 ***Emerging Methods for Efficient and Extensive Incorporation of Non-canonical Amino Acids Using Cell-Free Systems***
Yang Wu, Zhaoguan Wang, Xin Qiao, Jiaojiao Li, Xiangrong Shu and Hao Qi

90 ***Modular Enzymatic Cascade Synthesis of Nucleotides Using a (d)ATP Regeneration System***
Maryke Fehlau, Felix Kaspar, Katja F. Hellendahl, Julia Schollmeyer, Peter Neubauer and Anke Wagner

100 ***Corrigendum: Modular Enzymatic Cascade Synthesis of Nucleotides Using a (d)ATP Regeneration System***
Maryke Fehlau, Felix Kaspar, Katja F. Hellendahl, Julia Schollmeyer, Peter Neubauer and Anke Wagner

101 ***Functional Analysis of Aquaporin Water Permeability Using an* Escherichia coli*-Based Cell-Free Protein Synthesis System***
Ke Yue, Jihong Jiang, Peng Zhang and Lei Kai

112 ***The Genetic Code Kit: An Open-Source Cell-Free Platform for Biochemical and Biotechnology Education***
Layne C. Williams, Nicole E. Gregorio, Byungcheol So, Wesley Y. Kao, Alan L. Kiste, Pratish A. Patel, Katharine R. Watts and Javin P. Oza

125 ***Roadmap to Building a Cell: An Evolutionary Approach***
Zhanar Abil and Christophe Danelon

133 ***On the "Life-Likeness" of Synthetic Cells***
Luisa Damiano and Pasquale Stano

139 ***Modeling Cell-Free Protein Synthesis Systems—Approaches and Applications***
Jan Müller, Martin Siemann-Herzberg and Ralf Takors

146 ***Tuning the Cell-Free Protein Synthesis System for Biomanufacturing of Monomeric Human Filaggrin***
Jeehye Kim, Caroline E. Copeland, Kosuke Seki, Bastian Vögeli and Yong-Chan Kwon

160 ***Design, Development and Optimization of a Functional Mammalian Cell-Free Protein Synthesis Platform***
Chiara Heide, Gizem Buldum, Ignacio Moya-Ramirez, Oscar Ces, Cleo Kontoravdi and Karen M. Polizzi

170 ***Effective Use of Linear DNA in Cell-Free Expression Systems***
Megan A. McSweeney and Mark P. Styczynski

182 ***Designing Modular Cell-free Systems for Tunable Biotransformation of L-phenylalanine to Aromatic Compounds***
Chen Yang, Yushi Liu, Wan-Qiu Liu, Changzhu Wu and Jian Li

193 ***Decentralizing Cell-Free RNA Sensing With the Use of Low-Cost Cell Extracts***
Anibal Arce, Fernando Guzman Chavez, Chiara Gandini, Juan Puig, Tamara Matute, Jim Haseloff, Neil Dalchau, Jenny Molloy, Keith Pardee and Fernán Federici

frontiers in Bioengineering and Biotechnology

EDITORIAL
published: 29 November 2021
doi: 10.3389/fbioe.2021.799122

Editorial: Cell-Free Synthetic Biology

Jian Li[1]*, *Yong-Chan Kwon*[2]*, *Yuan Lu*[3]* and *Simon J. Moore*[4]*

[1]School of Physical Science and Technology, ShanghaiTech University, Shanghai, China, [2]Department of Biological and Agricultural Engineering, Louisiana State University, Baton Rouge, LA, United States, [3]Department of Chemical Engineering, Tsinghua University, Beijing, China, [4]School of Biosciences, University of Kent, Canterbury, United Kingdom

Keywords: cell-free systems, synthetic biology, cell-free protein synthesis, cell-free metabolic engineering, cell-free sensors, synthetic cells, biocatalysis and biotransformation

Editorial on the Research Topic

Cell-Free Synthetic Biology

Cell-free systems have historically been used to decipher the genetic code and elucidate key principles of biological systems (Silverman et al., 2020; Garenne et al., 2021). Recently, renewed scientific interest in cell-free systems has driven the rapid development of cell-free synthetic biology, including the establishment of new cell-free platforms, bottom-up design of genetic circuits and artificial cells, prototyping of metabolic pathways, and designing of medical diagnostics and biosensors, among others (Lu, 2017; Moore et al., 2017; Liu et al., 2019; Copeland et al., 2021). Today, with the renaissance of cell-free biotechnology, more and more scientists are entering this field to explore the frontiers of synthetic biology. To highlight the current progress of cell-free research, this Research Topic aims to bring together diverse and exciting achievements in cell-free synthetic biology.

Here, we collected a total of 17 papers. These papers present a broad range of cell-free research across the development of new cell-free protein synthesis (CFPS) systems and construction of an educational "Genetic Code Kit." The papers published in this topic are briefly introduced below.

Laohakunakorn et al. highlighted the current capabilities and challenges in the rapid growing field of cell-free synthetic biology. Due to the open nature of cell-free systems and breakthroughs in enabling technologies, bottom-up construction of increasingly complex biomolecular networks using cell-free systems will help address a wide range of fundamental biological questions such as the design of minimal life.

Kelwick et al. introduced key concepts and recent advancements in cell-free synthetic biology, with a focus on examples relevant to the materials sciences. In particular, the biological materials of industrial and societal importance manufactured by cell-free systems were summarized. While some challenges still remain, cell-free synthetic biology together with interdisciplinary collaborations and emerging technologies will lead to a new frontier for sustainable cell-free materials production and the growing bioeconomy.

Wu et al. summarized recent progresses on the efficient and extensive incorporation of non-canonical amino acids into proteins using CFPS. Different methods were used to increase the incorporation efficiency by elimination of endogenous competition factors, engineering of protein translation factors, and *in vitro* aminoacylation.

McSweeney and Styczynski reviewed the strategies of using linear DNA as expression templates in CFPS systems. Using linear DNA templates has some advantages, however, they can be quickly degraded by native exonucleases present in the cell extracts. This review described several approaches to stabilize linear DNA and compared their effectiveness and limitations, as well as future developments and applications.

Müller et al. introduced the modelling approaches and applications of CFPS systems, which can lay the foundation for understanding the production and degradation dynamics of macromolecules in CFPS. With the help of CFPS models, bottlenecks in the transcription and translation processes

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

***Correspondence:**
Jian Li
lijian@shanghaitech.edu.cn
Yong-Chan Kwon
yckwon@lsu.edu
Yuan Lu
yuanlu@tsinghua.edu.cn
Simon J. Moore
s.j.r.moore@kent.ac.uk

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

Received: *21 October 2021*
Accepted: *11 November 2021*
Published: *29 November 2021*

Citation:
Li J, Kwon Y-C, Lu Y and Moore SJ (2021) Editorial: Cell-Free Synthetic Biology. Front. Bioeng. Biotechnol. 9:799122. doi: 10.3389/fbioe.2021.799122

can be identified and key kinetic parameters can be determined from the modelled data, to aid the future development of enhanced CFPS systems.

Two papers discussed synthetic cells with different focuses. Damiano and Stano presented their opinion on the "life-likeness" of synthetic cells. Based on previous work, they promoted an "organizational approach" to the assessment of life-likeness and proposed the transition from behavioral assays (e.g., Turing test) to systemic strategies according to concepts like organization, complexity, networks, and emergence. Abil and Danelon pointed out in their perspective article that life is more than the sum of its constituted parts. These authors proposed a semi-rational, system's level evolutionary approach for bottom-up building a synthetic cell.

Lin et al. developed a flow platform by combining CFPS systems and a tube-in-tube reactor that can measure and analyze the effect of oxygen on protein synthesis. Using this platform, energy, transcription, translation, and pathway networks in the reaction system were analyzed. The results showed that a maximum protein yield could be obtained with 21% oxygen, which is the same as the natural atmosphere oxygen condition.

Zhang et al. established a eukaryotic *Pichia pastoris*-based CFPS system. After a systematic optimization, this cell-free system was able to synthesize 50 μg/ml of protein, which is comparable to other eukaryotic CFPS platforms. As a result, it is a valuable addition to the current CFPS toolbox and will provide an alternative for synthesizing complex proteins requiring post-translational modifications.

Williams et al. reported the transformation of the research-based CFPS biotechnology into a hands-on module called the "Genetic Code Kit" for implementation into teaching laboratories. This kit can help improve students' understanding of transcription and translation and prepare undergraduate students for careers in laboratory research.

Kim et al. reported the production of monomeric filaggrin, a human therapeutic protein, using an *E. coli*-based CFPS system. By relieving the limiting factors in transcription and translation, both protein yield and solubility were significantly improved with the enhanced CFPS system. This work demonstrates the potential of the *E. coli* CFPS to be adapted for studying therapeutic proteins.

Yang et al. designed modular cell-free systems for tunable biosynthesis of value-added chemicals. This approach was able to build long metabolic pathways (e.g., 6-7 enzymes) *in vitro* for defined biotransformation. Using two aromatic compounds (sharing the same upstream intermediate styrene) as examples, modular CFPS systems showed more than 80% conversion rates based on the same starting substrate of L-phenylalanine.

Arce et al. investigated cell-free RNA sensors using in-house made cell extracts instead of the commercial PURE CFPS system. CRISPRi-based technology was used to engineer the source strain to obtain nucleases silenced lysates, enabling CFPS from linear PCR-derived templates with high stability in the reaction. The home-made cell lysates were cost-effective and capable of virus detection using cell-free RNA toehold sensors.

Yue et al. described a systematic method for the expression and functional characterization of aquaporins based on an *E. coli* CFPS system, including template design, aquaporin expression and purification, proteoliposome preparation, and an aquaporin water permeability assay. This protocol can be adapted for the preparation of samples involving other channel proteins or transporters for their function assays.

Heide et al. provided a detailed protocol for the stepwise development and optimization of a CFPS platform derived from mammalian Chinese hamster ovary (CHO) cells. With the optimized CHO CFPS system, the protein yields were significantly improved, making it a useful research tool and a rapid production and screening platform for therapeutic protein development.

In addition to the above described cell lysates based cell-free systems, two research papers report the use of purified enzymes to implement *in vitro* biocatalysis. Fehlau et al. constructed a one-pot enzymatic cascade reaction for the synthesis of natural and modified nucleotides. Especially, a (d)ATP regeneration system was coupled in the reaction, leading to high-yielding production of nucleotides from nucleosides with high conversion rates. Liu et al. utilized T4 phage capsid as a protein scaffold for the immobilization of a three-enzyme cascade through SpyTag/SpyCatcher pairing. By doing this, the spatially organized enzymes exhibited higher enzymatic activity than the free enzymes, demonstrating the naturally occurring T4 phage scaffold is adaptable for multi-enzyme cascade assembly with enhanced biocatalytic activity.

In summary, the articles collated in this Research Topic demonstrate the potential of cell-free synthetic biology for a wide range of applications. Looking forward, cell-free synthetic biology research will continue to expand further to address current challenges in the field of synthetic biology when cell-based platforms are difficult or not amenable. With that, we now launch the second Research Topic on "Cell-Free Synthetic Biology, Volume II" and warmly welcome your significant contribution and submission.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work and approved it for publication.

REFERENCES

Copeland, C. E., Langlois, A., Kim, J., and Kwon, Y.-C. (2021). The Cell-free System: A New Apparatus for Affordable, Sensitive, and Portable Healthcare. *Biochem. Eng. J.* 175, 108124. doi:10.1016/j.bej.2021.108124

Garenne, D., Haines, M. C., Romantseva, E. F., Freemont, P., Strychalski, E. A., and Noireaux, V. (2021). Cell-free Gene Expression. *Nat. Rev. Methods Primers* 1, 49. doi:10.1038/s43586-021-00046-x

Liu, W.-Q., Zhang, L., Chen, M., and Li, J. (2019). Cell-free Protein Synthesis: Recent Advances in Bacterial Extract Sources and Expanded Applications. *Biochem. Eng. J.* 141, 182–189. doi:10.1016/j.bej.2018.10.023

Lu, Y. (2017). Cell-free Synthetic Biology: Engineering in an Open World. *Synth. Syst. Biotechnol.* 2, 23–27. doi:10.1016/j.synbio.2017.02.003

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. doi:10.1042/bst20170011

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. doi:10.1038/s41576-019-0186-3

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.

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

REVIEW
published: 24 March 2020
doi: 10.3389/fbioe.2020.00213

Bottom-Up Construction of Complex Biomolecular Systems With Cell-Free Synthetic Biology

*Nadanai Laohakunakorn[1], Laura Grasemann[2], Barbora Lavickova[2], Grégoire Michielin[2], Amir Shahein[2], Zoe Swank[2] and Sebastian J. Maerkl[2]**

[1] School of Biological Sciences, Institute of Quantitative Biology, Biochemistry, and Biotechnology, University of Edinburgh, Edinburgh, United Kingdom, [2] School of Engineering, Institute of Bioengineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Cell-free systems offer a promising approach to engineer biology since their open nature allows for well-controlled and characterized reaction conditions. In this review, we discuss the history and recent developments in engineering recombinant and crude extract systems, as well as breakthroughs in enabling technologies, that have facilitated increased throughput, compartmentalization, and spatial control of cell-free protein synthesis reactions. Combined with a deeper understanding of the cell-free systems themselves, these advances improve our ability to address a range of scientific questions. By mastering control of the cell-free platform, we will be in a position to construct increasingly complex biomolecular systems, and approach natural biological complexity in a bottom-up manner.

Keywords: cell-free synthetic biology, cell-free protein synthesis, *in vitro* reconstitution, microfluidics, compartmentalization, artificial cell, *in vitro* replication

OPEN ACCESS

Edited by:
Simon J. Moore,
University of Kent, United Kingdom

Reviewed by:
Pasquale Stano,
University of Salento, Italy
Guy-Bart Vincent Stan,
Imperial College London,
United Kingdom

***Correspondence:**
Sebastian J. Maerkl
sebastian.maerkl@epfl.ch

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

Received: *13 January 2020*
Accepted: *03 March 2020*
Published: *24 March 2020*

Citation:
Laohakunakorn N, Grasemann L, Lavickova B, Michielin G, Shahein A, Swank Z and Maerkl SJ (2020) Bottom-Up Construction of Complex Biomolecular Systems With Cell-Free Synthetic Biology. Front. Bioeng. Biotechnol. 8:213. doi: 10.3389/fbioe.2020.00213

1. INTRODUCTION

Synthetic biology promises to transform diverse domains including biomanufacturing, healthcare, food production, sustainable energy, and environmental remediation, by applying engineering principles to the design and construction of biological systems (Endy, 2005). Specifically, this was stipulated to involve abstracting away intricate biological complexity into simpler parts and modules whose behavior can be quantified (Heinemann and Panke, 2006; Arkin, 2008; Canton et al., 2008). The process of "building" thus involves assembling these subsystems together to obtain a required function, while quantitatively characterized components and their interactions ensure that the overall system may be predictively designed.

Practice currently diverges from the ideal framework set out above, due to the fact that we do not yet have a reliable approach to managing biological complexity (Kwok, 2010). While the idea of abstracting the behavior of a biological process, such as gene expression, into a simple mathematical model may indeed work well for single genes in isolation, as the gene circuit increases in size and complexity, the increased enzymatic and metabolic burden leads to reduced gene expression, changes in host cell state and growth rate, and increasing negative selection pressure. A seemingly modular component naturally loses its modularity as the system becomes more complex, and thus a major bottleneck preventing the current practice of synthetic biology from attaining the ideals outlined above lies in the transition from simple parts and circuits to larger systems (Purnick and Weiss, 2009).

There are several approaches to meet this challenge of reliable engineering of large biological systems, in the face of unknown complexity. One is to take advantage of increasing automation and experimental throughput to arrive at a functional design through screening large libraries of alternative constructs (Hillson et al., 2019). In order to effectively explore the parameter space, these screens may be guided by techniques, such as directed evolution (Agresti et al., 2010). A more rational approach is to discover designs which are robust to specific uncertainties, as exemplified by control theoretic approaches (Khammash, 2016; Vecchio et al., 2016; Hsiao et al., 2018). In this approach, it is not necessarily required to fully characterize the system, but merely to know which parts of the system are uncharacterized and varying, and therefore need to be buffered by an appropriate architecture.

Finally, a fully bottom-up approach attempts to rationally construct increasingly complex biomolecular systems from basic parts *in vitro* (Liu and Fletcher, 2009; Caschera and Noireaux, 2014a; Göpfrich et al., 2018; Schwille et al., 2018; Ganzinger and Schwille, 2019; Liu, 2019). In this approach, the major interactions within the system can in principle be fully quantified and understood. The payoffs from these efforts are well-informed models and understanding of increasingly complex biological systems (Elowitz and Lim, 2010), which may eventually guide fully predictive design in the future.

The rapidly growing field of cell-free synthetic biology (Garenne and Noireaux, 2019) brought forth numerous examples where such a constructivist approach has been adopted to elucidate basic principles associated with bottom-up construction of biomolecular complexity. The purpose of this review is to give a historical perspective and present an overview of the current capabilities and challenges facing this particular approach. We begin by giving an overview of the rich scientific history of cell-free gene expression systems and their use in deciphering fundamental biological processes by deconstructing them into their essential components. We then describe the current state of bottom-up cell-free synthetic biology, with a dual focus on both the cell-free systems themselves, as well as emerging technological platforms that enable increasingly complex and sophisticated manipulations of cell-free systems. Finally, we discuss how the construction of additional complexity on top of existing TX-TL systems stimulates the investigation of fundamental biological questions, which include context effects in gene expression, resource management, and possibilities for *in vitro* DNA replication.

Reliable engineering of synthetic biomolecular systems is an ambitious goal, whose success will depend on knowledge and insights gained from many different perspectives. We envision that the bottom-up approach, as exemplified in particular by cell-free synthetic biology, will play a key role in enabling the full potential of synthetic biology.

2. DECONSTRUCTING BIOLOGY USING CELL-FREE SYSTEMS

Cell-free systems are created by extracting cellular machinery, and combining them with energetic substrates and cofactors to recapitulate central biological processes, such as transcription and translation *in vitro*. While this approach has been in existence since Buchner's (1897) observation of cell-free fermentation in yeast extract (Buchner, 1897), it was only during the molecular biology revolution in the 1960s that cell-free systems began to be used in a rational and directed manner to elucidate biological mechanisms.

Early pioneers of cell-free investigations took advantage of two important properties of the system: its simplified biochemical nature, and its open reaction environment. Preparing a cell-free extract strips away much of the complexity of cellular regulation, homeostasis, and growth, revealing the isolated biochemical mechanisms underneath. By reconstituting the basic steps of protein synthesis, *E. coli* cell-free systems were used to demonstrate peptide synthesis from amino acids (Lamborg and Zamecnik, 1960), RNA (Nirenberg and Matthaei, 1961), and finally DNA, via coupled *in vitro* transcription and translation (Wood and Berg, 1962; DeVries and Zubay, 1967; Lederman and Zubay, 1967), thereby experimentally validating the central dogma of molecular biology. The first full protein synthesized *in vitro* was the coliphage F2 coat protein (Nathans et al., 1962).

The open nature of cell-free systems meant that factors which affected protein synthesis could be isolated and characterized, thus allowing direct study of transcriptional and translational regulation. Well-known examples of this work include the direct demonstration of the lac repressor's effect on peptide synthesis (Zubay et al., 1967), and the identification, isolation, and characterization of the catabolite activator protein (CAP) (Zubay et al., 1970). Cell-free systems were subsequently used to identify and elucidate genetic operons in *E. coli* (Zubay, 1973).

Another set of cell-free experiments of fundamental importance was the study of translation from synthetic polyribonucleotides by Nirenberg et al. They observed that cell-free extracts loaded with synthetic poly-uracil led to the production of only one type of polypeptide, poly-phenylalanine (Nirenberg and Leder, 1964). Thus, they hypothesized that poly-U must encode for phenylalanine. Over the next few years, the base composition, triplet nature, and eventually the genetic code mapping DNA sequence to amino acids was determined (Nirenberg et al., 1966).

Over the subsequent few decades, it became a standard approach to use *in vitro* systems to elucidate mechanisms in molecular biology [e.g., RNA replication (Mills et al., 1967), splicing (Kruger et al., 1982), Golgi trafficking (Balch et al., 1984), and chemiosmosis (Steinberg-Yfrach et al., 1998)]. In parallel, the growth of *in vitro* protein synthesis applications drove the development of increasingly efficient cell-free extracts, which achieved greater yields by incorporating more advanced metabolism to energize synthesis and recycle waste products (Jermutus et al., 1998). In the early 2000s, extract engineering merged with the nascent field of synthetic biology, giving rise to the field of cell-free synthetic biology (Noireaux et al., 2003), where instead of reconstituting existing biological processes, novel ones were constructed in the cell-free environment. This synthetic approach continues to characterize the field today.

3. TECHNOLOGIES

3.1. Lysates and Reconstituted Cell-Free Systems

In recent years the number of cell-free transcription-translation (TX-TL) systems from different organisms has grown rapidly (Zemella et al., 2015; Perez et al., 2016; Gregorio et al., 2019). The most common lysate systems include *E. coli*, insect, yeast, Chinese hamster ovary, rabbit reticulocyte, wheat germ, and human HeLa cells; and newly emerging systems include *B. subtilis* (Kelwick et al., 2016; Yim et al., 2019), *V. natriegens* (Failmezger, 2018; Yim et al., 2019), and *P. putida* (Wang et al., 2018; Yim et al., 2019), among others (Yim et al., 2019). Hybrid systems composed from multiple sources have also recently emerged (Anastasina et al., 2014; Panthu et al., 2018; Yim et al., 2019). Many of these lysate systems are currently commercially available. Concurrent with the expanding set of available lysate systems, there has also been a resurgence of interest in reconstituted recombinant systems, which are composed of mixtures of purified enzyme components. In this review, we will focus on *E. coli* lysate as well as recombinant systems, as they are commonly-used cell-free systems.

3.1.1. *E. coli* Lysates

The preparation and performance of *E. coli* lysate-based TX-TL systems vary tremendously and it is well-known that there can be large variability between different batch preparations (Takahashi et al., 2015b). For example, a recent study showed variability of more than 40% for TX-TL systems prepared in different laboratories, which resulted mainly from differences in personnel, and reagents used, and significantly, the laboratory in which the measurement was carried out (Cole et al., 2019). Fortunately, there is an increasing understanding of the role that each of the preparation steps plays in determining the final extract performance, as well as the factors responsible for reproducibility (Silverman et al., 2019b). Proteomics has been applied to elucidate the dependence of lysate composition and performance on batch variability, preparation methods (Failmezger et al., 2017; Foshag et al., 2018), as well as strain variability (Hurst et al., 2017; Garenne et al., 2019). The quest for a deeper understanding is also supported by the use of additional methods, such as metabolomics (Bujara et al., 2011), and other techniques such as polysome profiling (Liu et al., 2005), HPLC (Martin et al., 2018), and gel electrophoresis (Jaroentomeechai et al., 2018) (**Figure 1C**). These results raise the exciting prospect that lysates will become an engineerable substrate, where standardized and controlled preparation can result in extracts with a variety of defined behaviors. This approach has been particularly powerful in the context of cell-free metabolic engineering, and has been reviewed extensively by Karim et al. (2016) and Karim and Jewett (2018). Here we present an overview of different types of lysate preparation steps (**Figure 1A**), and their effects on lysate properties. The history of the field, recent advances, as well as the development, optimization, and applications of TX-TL systems are covered in recent reviews (Chiao et al., 2016; Silverman et al., 2019a).

E. coli extracts are prepared from a variety of different strains, whose choice strongly depends on the intended application. The most commonly used strains are BL21-derivatives (Sun et al., 2013; Kwon and Jewett, 2015; Didovyk et al., 2017; Cole et al.,

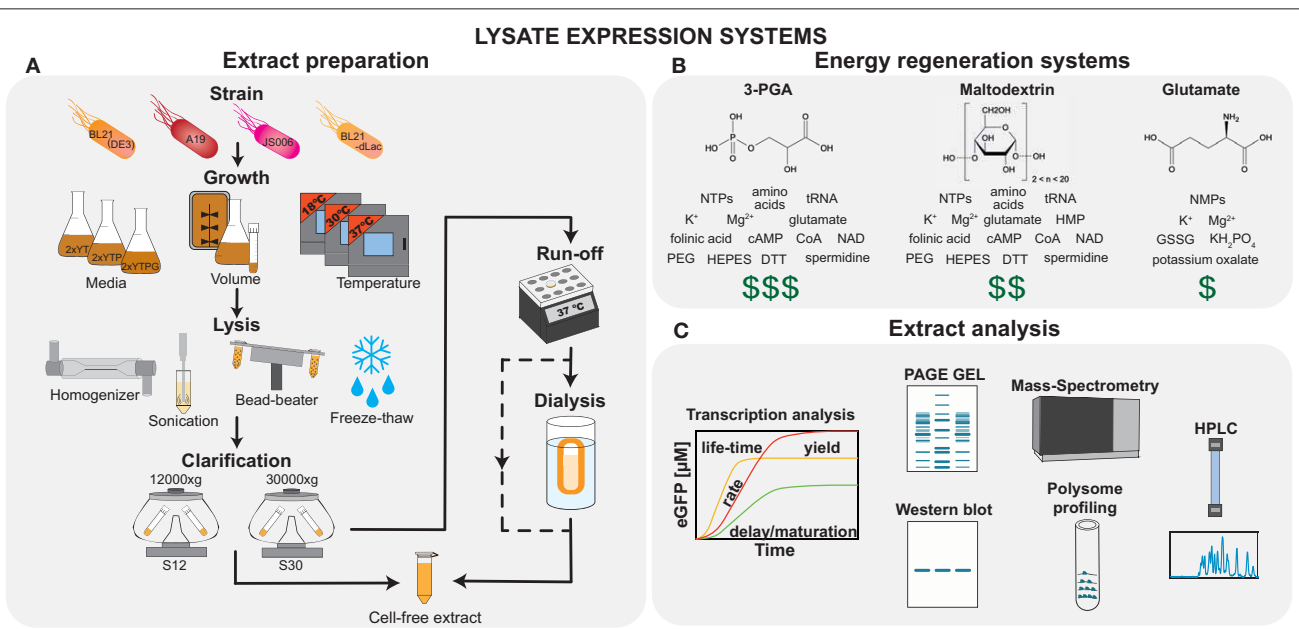


FIGURE 1 | Cell-free lysate systems. **(A)** The major steps in lysate preparation include growth, lysis, and clarification; however there exists a number of variables and options at each step, which can be adjusted to influence the final extract performance. **(B)** Examples of three energy regeneration systems are shown, which offer different cost-performance tradeoffs. **(C)** The final extract composition and performance may be analyzed using techniques, such as protein expression analysis, PAGE gel, Western blot, mass-spectrometry, polysome profiling, and HPLC.

2019), but the use of other strains can also be advantageous. For example, strains lacking DNAase, RNAase, and other *E. coli* enzymes can be used to enhance protein yield (Hong et al., 2015; Kwon and Jewett, 2015), for biosensing applications (Didovyk et al., 2017), or for circuit prototyping (Niederholtmeyer et al., 2015).

Different media, such as 2 × YT (Kim et al., 2006), 2 × YTP (Sun et al., 2013; Failmezger et al., 2017), or 2 × YTPG (Kwon and Jewett, 2015), as well as different temperatures and volumes can be used, which will influence the bacterial proteome and thus the composition of the lysate. For example, adding phosphate and glucose has suppressive effects on phosphatase activity (Kim and Choi, 2000). Bacteria can also be harvested at different time points during exponential or stationary phases. Surprisingly, this appears to have very little effect on lysate performance (Kwon and Jewett, 2015; Failmezger et al., 2017).

Cell lysis is a major and variable step of the overall lysate preparation, and different methods result in varying cost, scalability, and ease of use. Bacterial cells can be lysed by sonication (Kwon and Jewett, 2015), high-pressure homogenization (Hong et al., 2015), bead-beating (Sun et al., 2013), or enzymatic auto-lysis (Didovyk et al., 2017). Production yield between systems were shown to be comparable (Sun et al., 2013; Kwon and Jewett, 2015). However, other factors should also be considered. For example, the formation of inverted membrane vesicles is favored in lysates prepared with high-pressure homogenizers, and their preservation is essential for processes, such as oxidative phosphorylation (Jewett et al., 2008) and glycosylation (Jaroentomeechai et al., 2018). Subsequent lysate clarification usually involves centrifugation at 30,000 × g for S30 lysates or 12,000 × g for S12 lysates, which leads to different lysate clarity as distinct components sediment at different speeds, making the S30 lysate less viscous and opaque. For many applications no significant difference was observed between S30 and S12 lysates (Kim et al., 2006); however S12 lysates contain more inverted membrane vesicles which can support oxidative phosphorylation, and hence may be desirable for certain applications.

To reduce preparation time and simplify the process, some steps have been omitted in recent studies. Among these are run-off reaction and/or dialysis (Shrestha et al., 2012; Kwon and Jewett, 2015). Omitting these has minimal influence on final yield in T7 RNAP based systems (Kim et al., 2006; Kwon and Jewett, 2015) and might even be beneficial for retention of co-factors, amino acids, and tRNAs (Calhoun and Swartz, 2005a; Cai et al., 2015). However, the omission of both run-off reaction and dialysis has a profound effect when native transcriptional machinery is used (Kwon and Jewett, 2015; Silverman et al., 2019b).

Another important difference between systems is related to the energy regeneration approaches used (**Figure 1B**). The first systems based on substrates containing high-energy phosphate bonds (phosphoenolpyruvate, acetyl phosphate, creatine phosphate) were expensive and inefficient because of their fast degradation by non-specific phosphatases, and formation of inhibitory inorganic phosphate molecules. Over the last 20 years, a large amount of work has focused on yield improvement and price reduction. Most current energy regeneration systems are based on the native metabolic pathways of *E. coli*. These use either a part of—PANOx (Caschera and Noireaux, 2014b), 3-PGA (Sun et al., 2013)—or the entire *E. coli* glycolysis pathway—glucose (Calhoun and Swartz, 2005b), maltose (Caschera and Noireaux, 2014b), maltodextrin (Kim and Winfree, 2011; Caschera and Noireaux, 2015), and starch (Kim et al., 2011). These approaches have decreased the price per mg of synthesized protein to under one U.S. dollar. Nevertheless, we still lack systematic studies on the influence of these different energy regeneration methods on lysate properties other than simple protein yield. In particular, for prototyping and characterization of circuits, it is known that resource competition leading to improperly balanced energy usage (Siegal-Gaskins et al., 2014; Koch et al., 2018), efficiency of energy sources and small molecule replenishment (Siegal-Gaskins et al., 2014; Borkowski et al., 2018), changes in binding kinetics due to magnesium ion concentration changes (Kim et al., 2008), and pH variability (Calhoun and Swartz, 2005b) are all dependent on the energy system used and are expected to have profound influence on circuit behavior.

Finally, lysates can be directly supplemented with additives, such as liposomes, polymers, and detergents to facilitate folding of membrane proteins (Hein et al., 2014; Henrich et al., 2015). Enzymes, such as gamS (Sun et al., 2014) or short DNA decoy sequences (Marshall et al., 2017) can be added to prevent linear DNA degradation. The ease of adding functionality to lysates is a major advantage facilitated by the open nature of cell-free reactions.

3.1.2. Recombinant Systems

Lysate systems contain essentially all cytoplasmic components, which is advantageous for recapitulating cellular processes. However, this makes their composition ill-defined, leading to challenges in basic science and engineering. To address these difficulties, efforts were made to generate fully recombinant cell-free systems from a small number of purified enzyme components, whose composition can be defined exactly. Such defined systems are especially important for bottom-up synthetic biology for three main reasons. The first is that their use supports research into minimal cellular systems, as "minimality" of components and pathways can be directly tested. Secondly, the composition of the recombinant system is known much more precisely than for extract-based systems. This property is highly beneficial for modeling, optimization, troubleshooting, and mechanistic understanding of engineered pathways. Thirdly, the use of recombinant cell-free systems presents a viable approach toward the development of *de-novo* constructed synthetic cells.

Almost half a century ago, Weissbach's group developed the first such systems from recombinant *E. coli* proteins (Kung et al., 1977), but observed very low protein yield. About 25 years later, thanks to the advent of His-tag purification as well as the addition of a creatine-phosphate-based energy regeneration system, Shimizu et al. (2001) developed a very similar system called PURE (protein synthesis using recombinant elements) but with markedly higher protein synthesis yield (**Figures 2A,B**). Currently, there are three commercially available versions of this system: PUREfrex 2.0 (GeneFrontier), PURExpress (NEB)

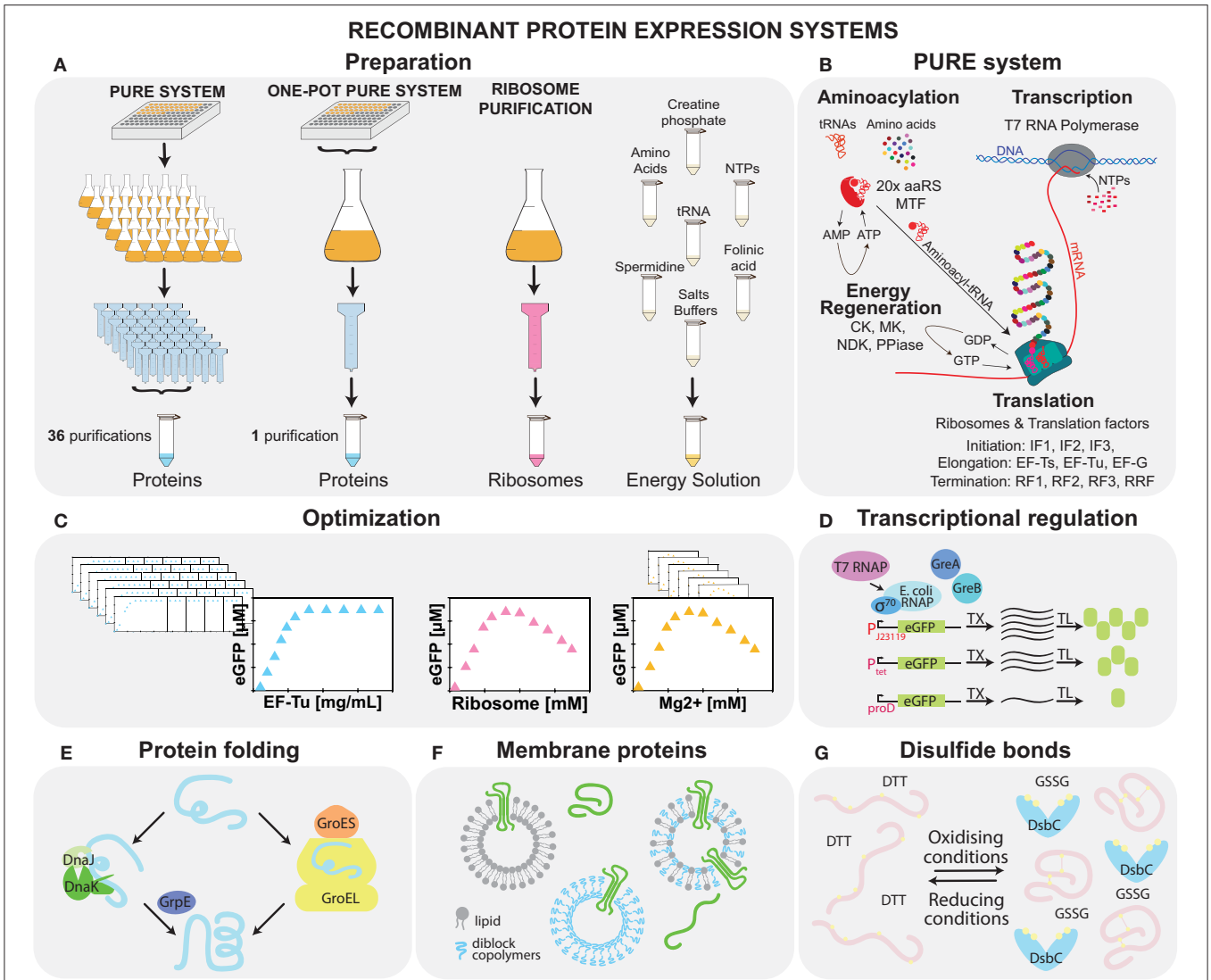


FIGURE 2 | Recombinant cell-free systems. **(A)** Schematic of the preparation of the three elements constituting the PURE system: proteins, ribosomes, and energy solution. **(B)** The four major reactions, aminoacylation, transcription, translation, and energy regeneration occurring during cell-free protein synthesis in the PURE system are shown along with a list of the components involved. **(C)** Optimization of the system can be carried out by adjusting both protein and energy solution components. Potential system modifications are shown: **(D)** supplementation with *E. coli* RNAP allows for more complex transcription regulation (Maddalena et al., 2016); **(E)** addition of chaperones aids protein folding (Niwa et al., 2012); **(F)** vesicles enable membrane protein folding and assembly (Kuruma and Ueda, 2015; Niwa et al., 2015a; Jacobs et al., 2019); and **(G)** oxidizing conditions allow for disulfide bond formation (Shimizu et al., 2005).

(Tuckey et al., 2014), and Magic PURE system (Creative Biolabs). Although highly popular, these systems are more expensive ($0.6–2/$\mu$L) than lysate systems ($0.3–0.5/μL). Moreover, despite the fact that the commercial systems are all based on the original PURE system, their exact composition is proprietary, and functional differences can be observed between them in terms of batch to batch variability, system yield, translation rate, lifespan of the reaction, and shelf-life (Doerr et al., 2019).

Cost-effective and modular PURE systems with user-defined compositions can be prepared in the laboratory (Shimizu and Ueda, 2010; Horiya et al., 2017), but the labor-intensive protocol requires ∼36 medium to large scale His-tag and ribosome purification steps (**Figure 2A**). Thus, different approaches to simplify the protocol have been developed, including His-tagging of *in vivo* enzyme pathways (Wang et al., 2012), microbial consortia (Villarreal et al., 2018), and bacterial artificial chromosomes (Shepherd et al., 2017). The first two systems achieved a 10–20% protein yield compared to the commercial PURExpress (NEB). Although the third approach reached protein synthesis levels comparable to PUREfrex, in all three of these approaches it is not possible to rapidly modify protein levels or omit proteins. We recently demonstrated that all proteins, except ribosomes, can be prepared from individual strains in a single co-culture and purification step

called the OnePot PURE system, which achieves a similar protein synthesis yield as commercial PURExpress (Lavickova and Maerkl, 2019) (**Figure 2A**).

Much work has been carried out to improve existing recombinant systems, particularly focusing on the protein expression yield: in addition to increasing the versatility of the system, this has also resulted in a better understanding of the system itself. Improved yield, lower cost, and the ability to adjust the system composition opens up many possibilities for applications, such as the development of defined artificial cells, gene network engineering, biosensors, and protein engineering. Here we separated the various approaches into two distinct types: the first includes experimental and theoretical approaches which aim to find an optimal composition of the system, while the second involves supplementing the existing system with factors that augment its behavior.

One direction for optimizing recombinant systems for protein synthesis yield is focused on finding optimal concentrations of the basic system components, such as proteins, energy sources, small molecules, and salts (Kazuta et al., 2014; Li et al., 2014, 2017; Doerr et al., 2019) (**Figure 2C**). Important work to improve our understanding of the system was done by Matsuura et al. (2009), who performed titrations of all protein components. These studies showed that although the system is composed of a relatively small number of components, its behavior is complex, and its analysis requires multivariate optimization. One of the most important parameters in the system is the magnesium ion concentration, which influences ribosome function. It is difficult to control the concentration of magnesium ions as they can be chelated by negatively charged molecules, such as NTPs, creatine phosphates, and pyrophosphates (Li et al., 2014, 2017). Studies focused on protein component concentrations showed that the performance of the system is mostly influenced by the concentration of ribosomes and translation factors. Increased yield depended strongly on high concentrations of EF-Tu, which often forms more than 50% of the non-ribosomal protein content *in vivo*. Moreover, finding optimal concentrations is essential for release factors and initiation factors, as an inhibitory effect was shown for these components when higher-than-optimal concentrations were used (Matsuura et al., 2009; Kazuta et al., 2014; Li et al., 2014). Finally, the optimal composition of the system will vary depending on the application. As an example, high concentrations of components, such as NTPs enhance transcription and translation, while inhibiting DNA replication (Sakatani et al., 2015).

To better understand the system behavior and to identify limiting factors, computational models of the PURE system have been developed. This includes coarse-grained ordinary differential equation (ODE) models containing effective lumped parameters and a small number of reactions (Mavelli et al., 2015; Carrara et al., 2018; Doerr et al., 2019), as well as more complex models based on modeling of a large number of elementary reactions, which can provide more detailed mechanistic insights but whose connection to experimental data as well as parameter inference is challenging (Matsuura et al., 2017, 2018). These models show that a number of steps involving ribosomes could potentially become rate-limiting: these include slow elongation rates, peptide release, and ribosome dissociation; qualitatively similar results were observed experimentally (Kempf et al., 2017; Li et al., 2017; Doerr et al., 2019).

As in the case of lysates, a second approach is based on augmenting the system with additional components, such as proteins (Kazuta et al., 2008), crowding agents, and liposomes. For example, yields can be slightly increased by adding proteins, such as EF-4 (Li et al., 2014), EF-P (Li et al., 2017), Pth (Kazuta et al., 2014), and HrpA (Kazuta et al., 2008). Recently, an energy regeneration system originally based on three kinases was replaced by one featuring a single polyphosphate kinase. This improvement lowers the price of the energy source and simplifies the energy regeneration process (Wang et al., 2019). While the original PURE system only contains T7 RNA polymerase, with its limited capability for transcriptional regulation, *E. coli* σ-factor based transcription has been successfully demonstrated, albeit with low efficiency with certain promoters, which can be enhanced by adding purified *E. coli* polymerase alone or in combination with transcription elongation factors (Maddalena et al., 2016) (**Figure 2D**).

Protein folding can be improved by incorporating chaperones, such as a trigger factor, DnaK/DnaJ/GrpE, and chaperonin GroEL/GroES (**Figure 2E**). Likewise, Niwa et al. (2012) showed that the solubility of 800 aggregation-prone *E. coli* cytoplasmic proteins can be enhanced if chaperones are added. Furthermore, an oxidizing environment and a disulfide bond isomerase are essential for the expression of proteins containing disulfide bonds (Shimizu et al., 2005) (**Figure 2G**). The addition of liposomes (Kuruma and Ueda, 2015; Niwa et al., 2015a) together with diblock copolymers (Jacobs et al., 2019) is important for membrane-protein synthesis (**Figure 2F**). Finally, the concentration of components in the cell-free system is up to 100 times lower than the native *E. coli* cytoplasm. Crowding agents, such as bovine serum albumin (BSA) (Li et al., 2014), Ficoll (Ge et al., 2011), polyethylene glycol (PEG) (Ge et al., 2011; Li et al., 2014), or osmolites (Moriizumi et al., 2019) can help mimic the *E. coli* cytosol (Ge et al., 2011), but they affect both transcription, translation (Norred et al., 2018), and the final synthesized proteins (Niwa et al., 2015b) in a complex way. Further studies will be needed to decipher the various physico-chemical effects of crowding on gene expression. Lastly, it was shown that temperature optimization is a key factor for chaperone-free assembly of protein complexes, such as DNA polymerase (Fujiwara et al., 2013).

3.2. Microfluidic Platforms

While cell-free reactions can be carried out successfully in a simple test tube, the complexity and sophistication of experiments can be dramatically augmented by coupling them to the appropriate technological platform. There have been numerous technological advancements with respect to cell-free gene expression over the past few decades, leveraging advances in microarraying, automation, and in particular, microfluidics. Offering reductions of orders of magnitude in sample volume, concomitant low cost, small device footprint, quantitative detection methods, and precise sample manipulation, microfluidic technology has offered tremendous

improvements in control and throughput of cell-free reactions (Damiati et al., 2018; Dubuc et al., 2019). We will focus on recent platforms enabling increased control over batch and, importantly, steady-state reactions, as well as describe recent work in the area of compartmentalization.

3.2.1. Increased Throughput and Spatial Control of Batch Reactions

Early high-throughput methods of spatially confined cell-free batch reactions were applied to the generation of protein arrays. In 2004, Ramachandran et al. showed that a plasmid array spotted on a glass slide could be transformed into a protein array by submersing the entire slide in a cell-free reaction. mRNA and proteins were locally transcribed and translated from the spotted plasmid DNA and proximally captured by surface bound antibodies (Ramachandran et al., 2004, 2008). The *in situ* generated protein array could then be interrogated with a protein of interest. A similar concept was later integrated into a microfluidic device for the automated mapping of protein-protein interactions (Gerber et al., 2009). Here linear expression DNA templates are spotted on a glass slide in pairs. The DNA array is then aligned to a MITOMI microfluidic device (Maerkl and Quake, 2007) so that each pair of linear templates is enclosed by a reaction chamber. Loading of the device with cell-free reaction solution synthesizes the bait and prey proteins, which are then assayed for interaction using the MITOMI method. A similar approach was used to generate large numbers of defined bHLH (basic helix-loop-helix) transcription factor mutants to assess the evolutionary accessible DNA binding specificity repertoire of these transcription factors (Maerkl and Quake, 2009). Martin et al. (2012) used the method to generate an RNA array for protein-RNA interaction studies. More recently, hundreds of full-length *Drosophila* transcription factors spanning a size range of 37–231 kDa were expressed on-chip using a wheat germ cell-free system (Rockel et al., 2013). Such approaches are becoming appealing for protein engineering, especially with the rapid decrease in synthetic DNA cost. In 2015, we demonstrated that over 400 synthetic zinc-finger transcription factors could be synthesized and characterized *in vitro* using this approach (Blackburn et al., 2015).

As synthetic gene networks began to emerge, the advantages of cell-free protein expression were adopted to rapidly screen large libraries of functional DNA parts, avoiding *in vivo* cloning steps, and speeding up the design-build-test cycle (Siegal-Gaskins et al., 2014; Takahashi et al., 2015a). The advent of acoustic liquid handling robots has enabled cell-free reactions to be carried out in standard microwell plate systems with increased throughput and precision, while simultaneously reducing reagent usage. This was recently demonstrated and coupled with a Bayesian modeling approach, which offered a fast route to characterizing regulatory elements from a non-model microbial host (Moore et al., 2018). With their rapid and automated method the authors were able to infer previously unknown transcription factor binding affinities as well as quantify resource competition in cell-free reactions (**Figure 3A**). Cell-free systems are particularly amenable to mechanistic modeling, and Bayesian inference of model parameters, which benefits from the possibility to perturb the composition of open cell-free reactions. Bayesian approaches uses probability distributions to quantify the degree of belief and uncertainty in the model, and can be deployed to quantitatively compare a number of models as well as determining parameter uncertainty. Automated acoustic liquid handling was also used to test serine integrase recombination dynamics (Swaminathan et al., 2017). A Python package built to model and simulate biological circuits was then applied to the cell-free prototyping data to carry out Bayesian parameter inference.

Microfluidic platforms applied to cell-free TX-TL have also enabled the exploration of larger design spaces at faster time scales. For example, droplet microfluidics was used to rapidly generate a library of distinct combinations of DNA templates, inducer molecules, and cell-free extract concentrations, with the possibility of generating millions of parameter combinations per hour (Hori et al., 2017). Together with a dye labeling scheme, it was possible to create a detailed map of biocircuit expression vs. parameter combination (**Figure 3A**). Sharing a common goal of characterizing gene network parameters, an alternative microfluidic platform was developed to carry out cell-free TX-TL in high-throughput, using different combinations of surface immobilized DNA as the reaction templates (Swank et al., 2019). Functional repression assays and quantitative affinity measurements (Maerkl and Quake, 2007) were used to characterize a library of synthetic transcription factors, enabling gene regulatory networks to be built from purely synthetic parts *de novo* (**Figure 3A**). Another quantitative and multi-dimensional study of genetic promoters was carried out using parallel piezoelectric cantilever beams that were able to generate an array of droplets containing cell-free TX-TL reaction mixtures with highly accurate concentration gradients (Fan et al., 2017) (**Figure 3A**).

Setting aside high-throughput techniques, there exist many other innovative technologies for cell-free gene expression, including methods that have sought to introduce spatial organization. In particular, a chip was developed to separate transcription and translation into different compartments (Georgi et al., 2016). Multi-compartment vesicles were used to predefine regions in which different proteins would be synthesized *in vitro* (Elani et al., 2014). Furthermore, Jiao et al. (2018) fabricated a microfluidic device for the encapsulation of plasmid integrated clay microgels. The incorporation of magnetic beads in the microgels permitted their recovery and re-use in subsequent cell-free TX-TL reactions. A bead-based approach was also used to express and capture recombinant proteins in a hydrogel matrix (Lee et al., 2012). Lastly, surface-bound DNA microarrays were aligned with a hydrogel matrix embedding protein synthesis machinery enabling localized protein synthesis (Byun et al., 2013). These studies will be discussed in more detail in section 3.3.

3.2.2. Steady-State Cell-Free Reactions

While cell-free batch reactions provide a means to characterize gene circuits, parts, and devices, the complexity of biological networks that can be implemented is constrained as the systems quickly reach chemical equilibrium. As discussed in section 3.1.1, batch cell-free reactions quickly equilibrate or reach a

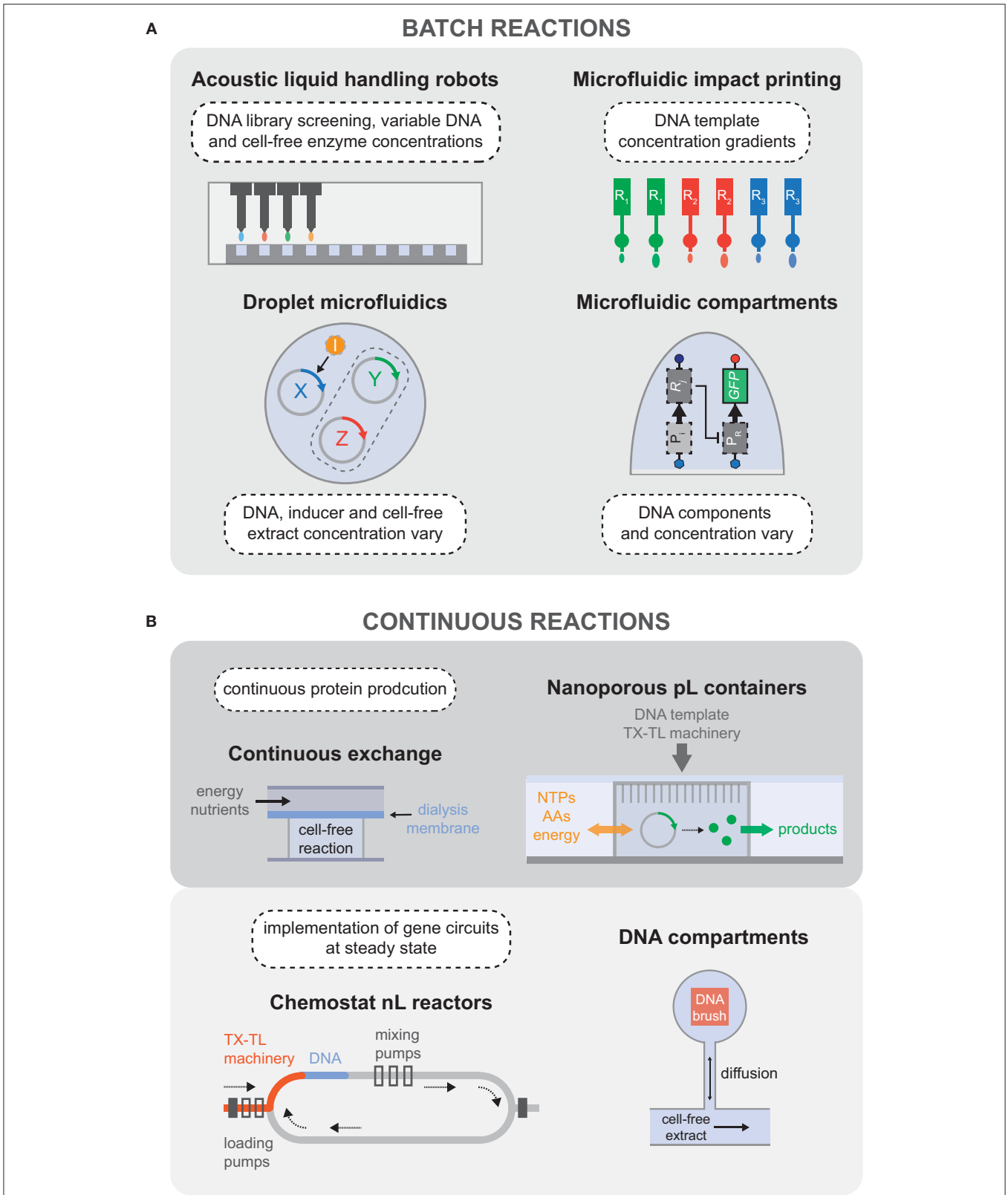


FIGURE 3 | Batch and continuous cell-free reaction platforms. **(A)** Overview of the technologies used to carry out high-throughput batch reactions, including the possibilities to vary the concentration of many reaction components in addition to exploring the sequence space of DNA templates. **(B)** Devices developed for continuous cell-free reactions, separated into two categories: continuous protein production, and steady-state reactors that enabled the implementation of genetic oscillatory circuits.

state of non-productivity for a number of reasons, such as byproduct or cofactor accumulation and subsequent drift from the initial reaction composition (e.g., inorganic phosphate, Mg^{2+}, H^{+}), denaturation or degradation of protein components, and simple exhaustion of substrate molecules. This has motivated the development of *in vitro* systems that can exchange reagents over time, maintaining the reaction in a non-equilibrium steady state, and mimicking the dilution and regeneration of cellular components during cell growth. Over 30 years ago there was interest in prolonging cell-free TX-TL reactions by providing a continuous flow of amino acids and energy sources to a reaction chamber from which synthesized proteins and by-products could be removed across an ultrafiltration membrane (Spirin et al., 1988). Successive work aimed to improve protein synthesis yield for cell-free TX-TL reactions by using a dialysis membrane to separate the reaction from the feeding solution of amino acids and energy sources, leading to a semi-continuous reaction (Kim and Choi, 1996; Madin et al., 2000). This idea was then extended to be compatible with standard micro-well plate systems that could be used for higher throughput applications (Mei et al., 2006, 2007; Khnouf et al., 2009, 2010). Following upon the same principles of continuous exchange cell-free reactions, a passive PDMS microreactor was built which separated the feeding and reaction chambers with a dialysis membrane, enabling protein synthesis for up to 15 h (Hahn et al., 2007) (**Figure 3B**).

Recent improvements in implementing continuous cell-free TX-TL reactions came in the form of novel microfluidic devices. For instance, continuous protein synthesis was demonstrated in an array of cell-sized nanoporous silicon containers that could exchange energy components and materials with the surrounding microfluidic environment (Siuti et al., 2011). In 2013, Niederholtmeyer et al. reported a two-layer PDMS device with eight independent nano-reactors that exchanged reagents at dilution rates similar to those of growing bacteria. Using this device, steady-state TX-TL reactions could be maintained for up to 30 h, enabling the first *in vitro* implementation of genetic oscillator circuits (Niederholtmeyer et al., 2013; van der Linden et al., 2019) (**Figure 3B**). Using the same device, Yelleswarapu et al. (2018) recently demonstrated the construction of synthetic oscillating networks using sigma-factor-based regulation of native RNAP in *E. coli* lysate. In 2014, Karzbrun et al. demonstrated two-dimensional DNA compartments capable of creating oscillating protein expression patterns and protein gradients. Each DNA compartment was linked to a supply channel by a small capillary channel for continuous diffusion of nutrients and products into and out of the compartment (Karzbrun et al., 2014) (**Figure 3B**). The geometry of the compartments determined the dilution rate of the reaction, giving rise to different observed reaction kinetics. Using high frequency localized electric field gradients, the same group was able to push the TX-TL machinery away from the DNA brush, thereby arresting transcription and translation. They showed that different biomolecules can be manipulated efficiently depending on the applied voltage and obtained sustained oscillation of gene expression from controlled ON/OFF switching of the TX-TL reaction (Efrat et al., 2018).

3.3. Compartmentalized Cell-Free Reactions

Compartmentalizing cell-free reactions spatially segregates a bulk reaction into smaller units. In addition to being a fundamental requirement in the construction of artificial cells, compartmentalized TX-TL opens up a number of scientific and practical opportunities, such as increased throughput for screening, *in vitro* directed evolution, distributed computation, and programmable communication. As discussed in sections 3.2.1 and 3.2.2, microwell plates with reaction volumes as low as 0.5 μL (Marshall et al., 2018), and microfluidic devices with volumes down to femtoliters (Karig et al., 2013), have been used to compartmentalize cell-free reactions.

Below, we will cover different types of compartmentalization including emulsions that allow for the rapid generation of multiple small volume compartments; liquid-liquid phase separation which can recapitulate naturally occurring crowded environments; hydrogels of natural or synthetic origin that immobilize DNA or proteinaceous factors and similarly provide a favorable crowded environment; liposomes which can provide a good starting point in the bottom-up assembly of synthetic cells by encapsulating a gene expression system; and other membrane-enclosed compartments with shells composed of polymers or protein-based materials that will expand the repertoire of physicochemical properties and functionalities.

3.3.1. Emulsion-Based Compartments

Emulsion-based compartmentalization allows for the rapid production of reaction vessels with volumes as low as femtoliters (Shojaeian et al., 2019). *In vitro* compartmentalization of TX-TL was first described in the context of *in vitro* evolution when Tawfik and Griffiths (1998) encapsulated a TX-TL system together with a DNA library of genes coding for an enzyme. Single copies of DNA templates were compartmentalized in $\sim$2 μm aqueous droplets dispersed in mineral oil, creating the crucial genotype-phenotype linkage (Contreras-Llano and Tan, 2018) which is required for selection and enrichment of improved enzymes. This eventually allowed a complete cycle of directed evolution of phosphotriesterases to be carried out (Griffiths and Tawfik, 2003).

One major drawback of emulsions produced by bulk methods is the size polydispersity of the obtained compartments (**Figure 4A**). This leads to enzymatic activity being convolved with noise resulting from variation in droplet size, making it difficult to select droplets containing improved enzymes. Dittrich et al. overcame this limitation using droplet microfluidics to generate monodispersed water-in-oil (W/O) droplets (**Figure 4A**) containing a TX-TL reaction expressing GFP. However, their setup did not allow for the production of droplets containing single DNA copies that gave rise to detectable signals, as would be required for *in vitro* evolution. Using a more efficient TX-TL system and stabilized W/O droplets, Courtois et al. (2008) were able to obtain efficient transcription and translation from a single DNA copy, opening the door for high throughput quantitative evolution experiments in droplets generated by microfluidics. Examples of these include multiple screening

rounds to enrich for active hydrogenase (Stapleton and Swartz, 2010) and beta-galactosidase enzymes (Fallah-Araghi et al., 2012).

The use of fluorogenic substrates in enzymatic assays can be problematic in surfactant stabilized emulsions as transport of fluorophores can occur between droplets both in single (Gruner et al., 2019) and double emulsions (Etienne et al., 2018). Woronoff et al. (2015) demonstrated an alternative methodology where a proteinogenic amino acid is released after enzymatic turnover and then incorporated in the translation of a reporter protein. Using this approach, they were able to screen for active penicillin acylase enzymes in single gene droplets. The literature contains fewer examples of compartmentalized *in vitro* assays to screen for protein binders. However, two-hybrid and three-hybrid systems have been developed in PURExpress supplemented with *E. coli* core RNAP enzyme (Zhou et al., 2014). Cui et al. (2016) used such an *in vitro* two-hybrid system encapsulated in single-emulsion droplets to screen a library of 105 peptide binders in a single day.

Recent work using droplets has diversified beyond the high-throughput screening studies discussed in the previous paragraphs to encompass physical effects, such as the influence of crowding (Hansen et al., 2015) or droplet size (Kato et al., 2012; Matsuura et al., 2012; Sakamoto et al., 2018) on protein expression. Schwarz-Schilling et al. (2018) used W/O droplets to compartmentalize streptavidin-coated magnetic beads which act as a scaffold on which complex RNA-protein nanostructures can be built using TX-TL. The high-throughput generation of such compartments is also attractive for the extensive parameter space mapping for genetic network prototyping, as exemplified by the work of Hori et al. (2017) discussed in section 3.2.1.

3.3.2. Liquid-Liquid Phase Separation

Liquid-liquid phase separation occurs when a water-soluble molecule, generally a polymer, is mixed with another aqueous solution containing either a high salt concentration or another water-soluble polymer. Under certain conditions, the first polymer cannot dissolve in the second solution, and a separation into two distinct phases occurs. The resulting "aqueous two-phase system" (ATPS) can form microscale, membrane-less compartments. The recent discovery that ATPS are ubiquitous in cells has attracted much attention to better understand their role in cell physiology (Alberti et al., 2019). Recreating cell-free transcription-translation reactions in these systems could help elucidate the properties of such condensates.

Torre et al. (2014) prepared ATPS of dextran/poly(ethylene glycol) or three-phase systems (A3PS) of dextran/poly(ethylene glycol)/ficoll containing TX-TL by vortexing in mineral oil (**Figure 4B**). In the ATPS, expression of the reporter protein indicated preferential partitioning of the TX-TL machinery to the dextran phase in the ATPS. The A3PS, on the other hand, exhibited lower expression, which was attributed to separation of TX-TL machinery into the different dextran and Ficoll phases, suggesting that different liquid phases could differentially partition TX-TL components.

When a liquid-liquid phase separated compartment consists of a condensate of biological polymers, it is most commonly referred to as a coacervate (**Figure 4B**). These coacervates are characterized by a high degree of macromolecular crowding, exhibiting protein concentrations of up to 272 g/L (Deng et al., 2018), similar to the *E. coli* cytosol. Such crowding can profoundly influence gene expression. Sokolova et al. (2013) used a microfluidic device to osmotically concentrate droplets containing lysate, and observed the formation of coacervates in lysate containing 2% PEG-8000. The resultant reporter gene expression was higher in coacervates than in single phase droplets. The work demonstrated that transcription rates were enhanced in the crowded environment of coacervates, offsetting the lower translation rate. Such observations are in agreement with previous studies in bulk cell-free reactions where macromolecular crowding enhances transcription and impairs translation (Ge et al., 2011). To generate monodisperse coacervates in high throughput, Tang et al. (2015) produced coacervates using a microfluidic device (van Swaay et al., 2015) starting from a mixture of carboxymethyl-dextran/polylysine and TX-TL. However, they observed lower gene expression in coacervates compared to the bulk reaction, with results suggesting charge-induced precipitation of the reporter protein after its production. This again indicates that protein expression is sensitive to the partitioning of the TX-TL machinery and that the charge of the coacervate and crowded environment can have opposite effects on yields.

3.3.3. Hydrogels

Similar environments to coacervates are found in hydrogels, where a highly porous hydrated network provides a crowded environment. Forming gel micropads by cross-linking X-shaped DNA entrapping plasmid DNA, or P-gel, Park et al. (2009a,b) obtained an up to 94-fold increase in protein production compared to a standard batch reaction (**Figure 4C**). They explained the increase in expression by an enhanced transcription rate due to the higher proximity of gene templates in the crowded DNA gel environment. The P-gel has also been prepared in a microdroplet format (Ruiz et al., 2012) and the microgel format was modified with Ni^{2+}-NTA to allow the immobilization of the expressed protein on the surface of the microgel (Kahn et al., 2016).

The same group showed that TX-TL was also increased in the presence of a clay hydrogel, which spontaneously forms when mixing hydrated clay in the presence of an ionic solution (Yang et al., 2013) (**Figure 4C**). DNA and RNA molecules localize to the clay hydrogel and are protected from enzymatic degradation by nucleases. The clay-DNA hydrogels were also formulated into microgels containing magnetic nanoparticles allowing for multiple successive TX-TL reactions after recovery of the magnetic microgel and refreshing of the TX-TL mixture (Jiao et al., 2018). Finally, clay-DNA microgels have been used as artificial nuclei inside W/O emulsions (Jiao et al., 2018) or inside permeable polymeric capsules (Niederholtmeyer et al., 2018).

Thiele et al. (2014) prepared hyaluronic acid functionalized with DNA template and produced porous hydrogel microparticles, which were further encapsulated in droplets containing TX-TL (**Figure 4C**). They observed efficient GFP

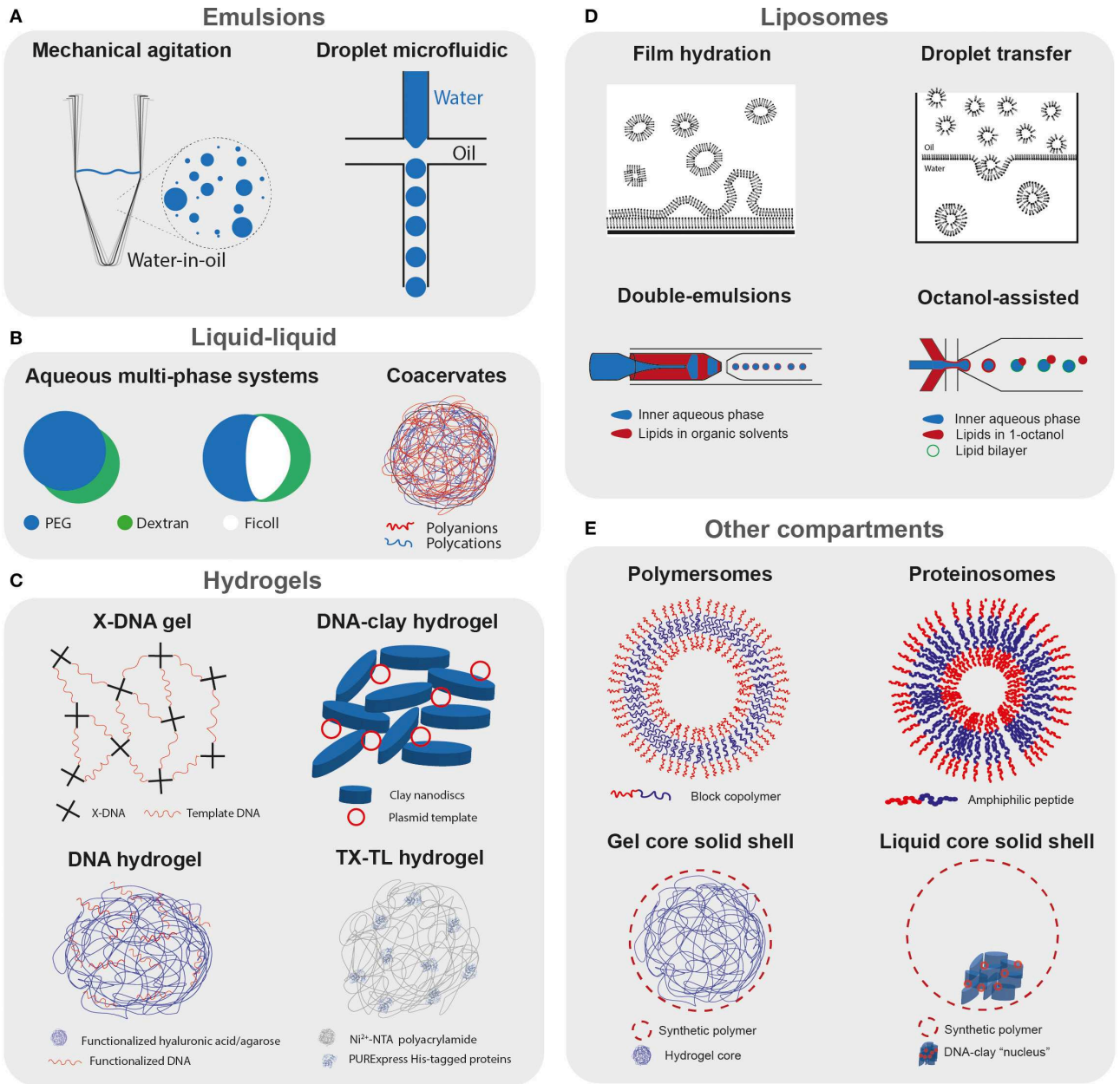


FIGURE 4 | Compartmentalized cell-free reactions. Schematic representation of the different strategies used to compartmentalize cell-free transcription translation reactions. **(A)** Emulsion-based compartments: polydisperse water-in-oil droplets obtained by mechanical agitation, and microfluidic production of monodisperse droplets. **(B)** Liquid-liquid phase separation: aqueous multiphase systems containing cell-free transcription translation machinery (Torre et al., 2014), and representation of a complex coacervate. **(C)** Hydrogels: X-DNA linking template DNA and forming a DNA hydrogel (Park et al., 2009a,b), a DNA-clay hydrogel (Yang et al., 2013), hyaluronic acid (Thiele et al., 2014), or agarose (Aufinger and Simmel, 2018) functionalized with DNA template, polyacrylamide hydrogel functionalized with Ni^{2+}-NTA binding PURExpress His-tagged proteins (Zhou et al., 2018). **(D)** Liposomes: rehydration of lipid films with an aqueous solution containing TX-TL, droplet transfer method where a lipid-stabilized W/O emulsion is layered on top of a feeding buffer and liposomes transferred to the bottom by centrifugation (Noireaux and Libchaber, 2004), double-emulsions with ultrathin shells containing lipids in organic solvent (Ho et al., 2015, 2017), and octanol-assisted assembly (Deshpande et al., 2016; Deshpande and Dekker, 2018). **(E)** Other compartments: polymersomes with membrane formed by amphiphilic polymers, proteinosomes with amphiphilic peptides (Vogele et al., 2018), alginate hydrogel coated with various polymers, artificial cells with polymeric shell and liquid core containing a DNA-clay "nucleus" (Niederholtmeyer et al., 2018).

protein expression proportional to the number of encapsulated DNA hydrogel beads, with the fluorescent protein diffusing inside the droplet. By using mRNA molecular beacons, they show that the transcribed mRNA remains trapped in the hyaluronic acid/DNA hydrogel, suggesting that transcription and translation both take place inside the hydrogel.

Aufinger and Simmel (2018) prepared agarose functionalized with alkynes and coupled to azide-modified DNA, and used it to prepare hydrogel-DNA "organelles" (**Figure 4C**). Transcription organelles contained template DNA coding for mVenus with a toehold switch on the 5′ end of the mRNA, whereas the translation organelles were functionalized with the corresponding toehold trigger. These organelles were re-encapsulated in W/O droplets containing TX-TL, and mVenus expression was observed only in droplets containing both the transcription and translation organelles. As these organelles can offer spatial organization of complex reactions while providing continuous exchange with the environment, they are useful for building more complex modular systems.

Whereas the previous studies focused on immobilizing the DNA template inside hydrogels, Zhou et al. (2018) immobilized the complete set of PURExpress His-tagged proteins on a polyacrylamide gel functionalized with Ni^{2+}-NTA or an anti-His-tag aptamer (Lai et al., 2020) (**Figure 4C**). The His-tagged proteins, ribosomes, and template plasmids are placed on pre-dried hydrogel particles, which effectively traps the ribosomes and plasmids in the hydrogel network by convection when rehydrated. Sustained gene expression is observed for as long as 11 days when the cell mimics are constantly supplied with fresh feeding buffer.

3.3.4. Liposomes

Liposomes are compartments encapsulated by a lipid bilayer similar to a cell membrane, making them attractive for the encapsulation of cell-free systems. Liposome technology has been recently reviewed by Stano (2019). Early studies used a film hydration method, where the reaction mix rehydrates a dried lipid film to produce liposomes encapsulating TX-TL (**Figure 4D**). This was deployed to translate peptides (Oberholzer et al., 1999), proteins (Yu et al., 2001; Oberholzer and Luisi, 2002; Nomura et al., 2003), and finally a more complex genetic cascade (Ishikawa et al., 2004). Noireaux and Libchaber (2004) presented a more convenient method of liposome production called droplet transfer, where a lipid stabilized emulsion of the reaction is first formed in oil and then layered on top of the feeding solution (**Figure 4D**). Liposomal vesicles are subsequently formed by centrifugation. By producing α-hemolysin *in situ*, which assembled to form pores in the liposome membrane, they were able to constantly supply feeding buffer to the encapsulated reaction and increase the duration of expression up to almost 100 h.

An interesting improvement in the lipid film rehydration method was presented by Nourian et al. (2012) where they dried the lipid films on 200 μm glass beads and rehydrated them with PURExpress. This allowed them to use low reaction volumes to produce liposomes in high yield and with high encapsulation efficiency. Moreover, they used phospholipids with shorter acyl chains to produce semi-permeable liposomes and incorporated biotinylated lipids for efficient immobilization of the vesicles on microscope slides.

Droplet microfluidics allows for the generation of double emulsions with ultrathin shells where the middle phase contains dissolved lipids and forms unilamellar vesicles after evaporation of the solvent (Arriaga et al., 2013) (**Figure 4D**). Ho et al. (2015) used this technology to encapsulate a mammalian cell-free system with very high encapsulation efficiency, and observe expression of GFP in the interior of the vesicles as well as expression and assembly of a trans-membrane protein. However, they observed in a consequent study that the surfactant necessary for double emulsion led to aggregation of the mammalian cell-free system (Ho et al., 2017).

By using triblock copolymer surfactants, Deng et al. (2016) could control the dewetting of the inner water drop from the middle organic phase thus forming perfectly unilamellar and uniform liposomes, in addition to solvent droplets that could be easily separated. A hierarchical assembly of liposomes inside other liposomes, or vesosomes, through multiple successive encapsulation and dewetting was also demonstrated (Deng et al., 2017). *In vitro* transcription of Spinach RNA was carried out in the interior "nucleus" liposome and translation of mRFP in the surrounding "cytoplasm" liposome, showing great potential toward bottom-up assembly of complex biomolecular structures, even though controlled transfer of mRNA from the interior to the surrounding liposome remains to be implemented. Finally, a similar method called octanol-assisted liposome assembly (OLA) was developed where the middle phase alkane solvents are replaced by octanol containing lipids and undergo rapid dewetting, which could further increase the efficiency and biocompatibility of the encapsulation method (Deshpande et al., 2016; Deshpande and Dekker, 2018) (**Figure 4D**).

3.3.5. Other Membrane Compartments

Other types of membrane compartments have also been used for cell-free protein expression, such as polymersomes, protein-based membranes, and polymeric shells (**Figure 4E**). Although there exist many different strategies and materials to make capsules (Cuomo et al., 2019), the conditions necessary for their production often prevent encapsulating cell-free systems. Martino et al. (2012) used a microfluidic capillary device to generate template double-emulsion for the direct encapsulation of a cell-free expression system inside polymersomes composed of PEG-*b*-PLA copolymer and PLA homopolymer to increase their stability. They successfully expressed an MreB protein which formed patches inside the aqueous core and also adhered to the membrane.

Vogele et al. (2018) used a film rehydration method similar to the one used for liposome production but with amphiphilic elastin-like peptides as building blocks, which formed vesicles upon rehydration with a TX-TL system (**Figure 4E**). They demonstrate that the expression of the elastin-like peptide led to its successful integration into the membrane and an increase in the size of the vesicles after a few hours of expression. Schreiber et al. (2019) also used amphiphilic peptides to form vesicles and encapsulate a cell-free expression system, and show the production and incorporation of amphiphilic peptide in the membrane. It will be interesting to see in future studies if pore-forming proteins can be incorporated in these "growing" protein-based membranes, which might allow for prolonged and higher protein expression, as was observed for cell-free protein expression in liposomes. By encapsulating a cell-free extract in

millimeter-sized alginate beads coated with polycationic chitosan (Kwon et al., 2008), silica (Lim et al., 2009), or polyethyleneimine (Saeki et al., 2014), researchers could show continuous expression of eGFP (**Figure 4E**). However, the core of the capsules presented in the previous studies is in a gel format and it is difficult to assess how well the capsules perform as no absolute quantification of the protein levels was provided.

To our knowledge, the only example to date where cell-free protein expression was demonstrated in liquid core-solid shell polymeric capsules was by Niederholtmeyer et al. (2018) where they produced porous polyacrylate capsules containing a DNA-clay hydrogel nucleus (**Figure 4E**). The capsules' pores are large enough to allow access by large macromolecules including ribosomes. Transcription-translation from the template DNA immobilized in the clay-DNA hydrogel "nucleus" can be achieved by immersing the capsules in a cell-free expression system. But, as the shell material leads to adsorption of proteins on the capsule surface and the pores are too large to retain the TX-TL machinery, the direct encapsulation of cell-free systems inside polymeric capsules remains to be demonstrated. Such direct encapsulation in synthetic polymeric capsules would be valuable as they could present attractive properties, such as high mechanical and chemical stability, as well as tunable porosity, based on the type of shell material and the fabrication method used.

3.3.6. Physical Effects of Compartmentalization

The effect of the compartment size and interface composition can have notable effects on gene expression. Initial work in Yomo's group showed that expression in sub-picoliter PDMS compartments severely hampered GFP synthesis, whereas quartz glass microcompartments passivated with amino acids showed expression as high as 41% of the test tube reaction with no dependence on compartment volume in a range from 40 fL to 7 pL (Okano et al., 2012). They later showed that synthesis of β-glucuronidase (GUS) with fourth-order reaction kinetics was favored in smaller compartments while GUS substrate depletion was rapidly occurring, pointing to an ideal compartment volume (Matsuura et al., 2012; Okano et al., 2014).

No size dependence on GFP synthesis was observed in a range from 1 to 100 μm in liposomes composed of a mixture of different phosphatidylcholine (PC) or phosphatidylglycerol (PG) lipids and cholesterol (Nishimura et al., 2012), in contradiction to previous reports where PG had inhibitory effect on protein synthesis (Sunami et al., 2010). In lipid stabilized droplets, the charge of the lipid used could also influence the synthesis rate, but in this case the relatively more negative PG lipid was favored over phosphatidylethanolamine (PE) or PC (Kato et al., 2012). Sakamoto et al. (2018) proposed a model with three regimes where there could be activation, no regulation, or repression at the surface. In droplets stabilized by PC lipids, they observed protein expression that did not scale with the droplet volume R^3, but with R^4 for droplets with radii below 17 μm, suggesting surface repression in their system. Other effects could explain variations in fluorescence intensity, such as the exchange of solutes between droplets which is influenced by the composition of the carrier oil, lipid or surfactant, as well as the radius of the droplets (Etienne et al., 2018).

The compartmentalization of biochemical reactions in smaller volumes increases the gene expression stochasticity as only a few molecules are present in each compartment. Hansen et al. (2015) suggest that such randomness can be explained by extrinsic noise, which results from the Poisson distribution of encapsulated reagents of the cell-free system, and intrinsic noise, which results from molecular crowding and other parameters, such as the stochasticity of the gene expression reactions or relative plasmid distributions. They co-encapsulated CFP and YFP plasmids in droplets with varying levels of crowding, and observed an increase in intrinsic noise with increased levels of crowding. Intrinsic noise in gene expression can also arise from the stochastic partitioning as was strikingly observed in liposomes prepared in dilute solutions of transcription-translation system (Stano et al., 2013). A small number of compartments ($<$ 0.5%) displayed detectable eGFP gene expression, whereas no expression occurred in free solution raising interesting questions about the mechanism of loading of the solute mixture.

High variability in gene expression was also observed in liposomes prepared in PURE solutions of normal concentration and interestingly gave rise to some compartments displaying particularly high or long lasting gene expression (Blanken et al., 2019). These large variations due to stochastic partitioning are interesting as a mechanism to generate diversity in the population, as recently discussed in a review by Altamura et al. (2018). Understanding and harnessing these physical effects of compartmentalization potentially offers yet another way of controlling cell-free gene expression.

3.3.7. Communication

Cellular communication is fundamental in biology and responsible for many processes ranging from development to tissue homeostasis. Following the successful developments in compartmentalizing cell-free systems, the next logical challenge consists of engineering inter-compartment communication. On-chip artificial cells consisting of DNA brushes (described in section 3.2.2) were interconnected in series by microfluidic channels, and communication is achieved by diffusion of molecules, which can be tuned by adjusting channel geometry (Tayar et al., 2015) (**Figure 5A**). Diffusion of a σ^{28} activator from one compartment to the next led to sequential switching of a bistable genetic circuit. In a follow-up study, Tayar et al. (2017) used a non-linear activator-repressor oscillator in compartments coupled by diffusion and observed that the oscillators could be synchronized and tuned by geometric control of diffusion. A key demonstration was that such reaction-diffusion systems could spontaneously form spatial patterns in good agreement with theory.

Moving away from microfluidic chips could potentially allow for the engineering of more complex, dynamic consortia of communicating compartments or even tissue-like assemblies. Schwarz-Schilling et al. (2016) used capillaries to align W/O droplets encapsulating cell-free extracts as well as *E. coli* cells (**Figure 5B**, top). The bacteria and cell-free systems contained either an AND gate circuit expressing GFP in

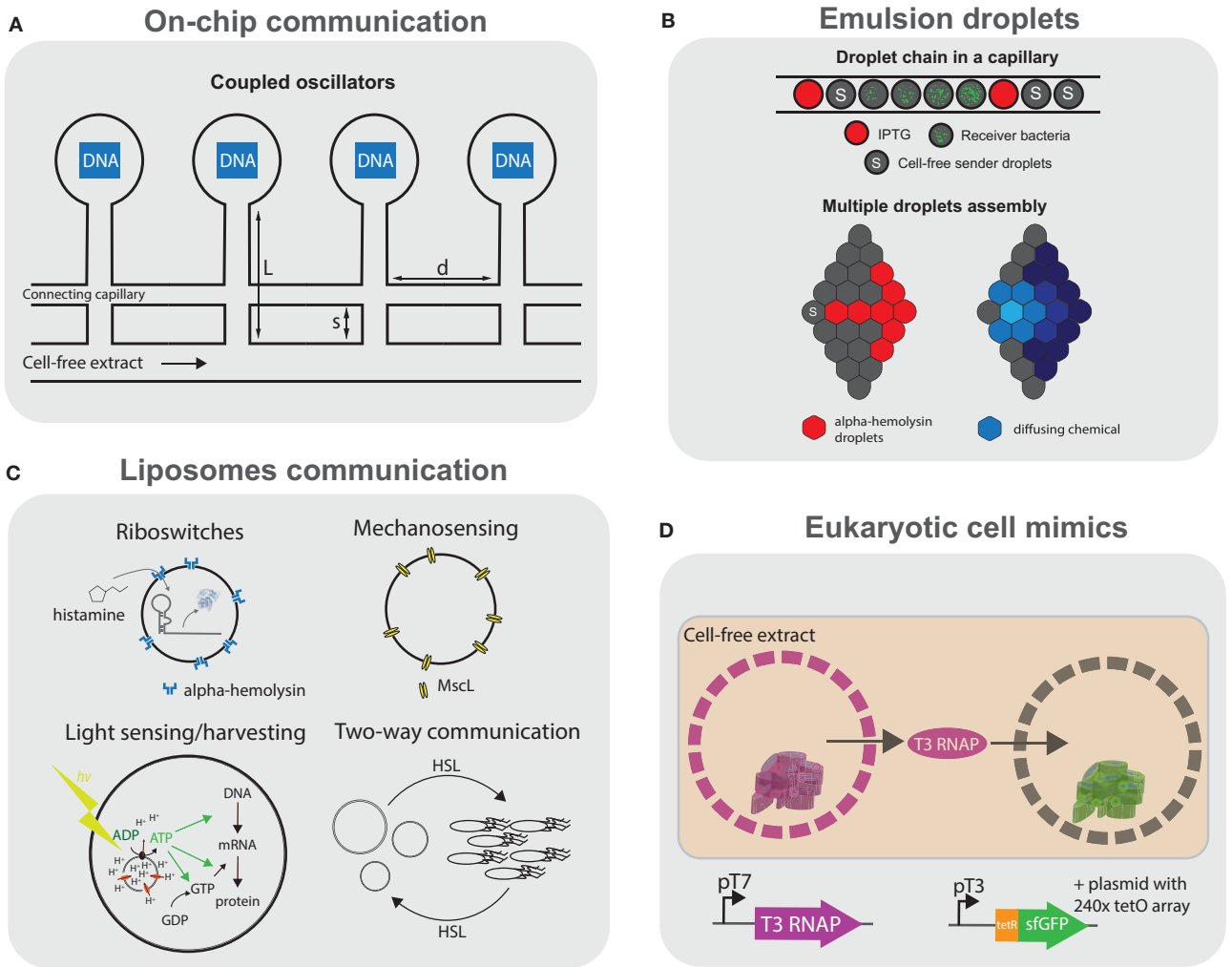


FIGURE 5 | Communication using cell-free transcription translation. Schematic representation of the different platforms using cell-free transcription translation reactions for communication. **(A)** Artificial cells on chip: DNA compartments are connected to a cell-free reaction feeding channel and interconnected by another capillary allowing the coupling of the compartments (Tayar et al., 2015). **(B)** Emulsion droplets: top, water-in-oil droplets containing small molecule activators, bacteria or cell-free genetic circuits arranged in a glass capillary (Schwarz-Schilling et al., 2016); bottom, multiple lipid-stabilized droplets assembled with a micromanipulator with some droplets containing pore forming α-hemolysin (Dupin and Simmel, 2019). **(C)** Sensing and communication with liposomes: liposomes encapsulating histamine-sensitive riboswitches (Dwidar et al., 2019), mechanosensing using MscL pores (Majumder et al., 2017; Garamella et al., 2019), light-driven ATP synthesis using bacteriorhodopsin and ATP synthase (Berhanu et al., 2019), and two-way communication between liposomes and bacteria using various AHLs (Lentini et al., 2017). **(D)** Eukaryotic cell mimics: microporous polymeric capsules containing a DNA-clay hydrogel "nucleus" are immersed in cell-free transcription translation. The expressed T3 polymerase can diffuse and activate transcription-translation in another compartment (Niederholtmeyer et al., 2018).

response to isopropyl β-d-1-thiogalactopyranoside (IPTG) and acyl homoserine lactone (AHL), or a sender circuit producing AHL in response to IPTG. Communication could be established between sender droplets and droplets containing the AND gate, in a cell-free-to-bacteria or bacteria-to-cell-free direction.

Dupin and Simmel (2019) used a micromanipulator to arrange multiple directly adjacent W/O droplets in a lipid-in-oil bath, forming a lipid bilayer interface between the compartments (**Figure 5B**, bottom). They show direct communication between sender droplets containing arabinose (ARA) or AHL and droplets containing a responder circuit. By using an incoherent feed-forward loop genelet circuit containing an RNA binding to 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), they observe the propagation of the DFHBI signal along multiple successive interconnected droplets. Finally, by encapsulating a positive feedback circuit expressing α-hemolysin in response to ARA, they observe an increased variability in protein expression levels among droplets, which they describe as "a primitive form of cellular differentiation."

Liposomes can more closely recapitulate cellular systems. Lentini et al. rehydrated liposomes containing a genetic circuit using a riboswitch responding to theophylline to express α-hemolysin and release co-encapsulated IPTG (**Figure 5C**). By incubating *E. coli* with these liposomes acting as signal

translators, the bacteria could effectively respond to theophylline in the medium (Lentini et al., 2014). They later demonstrated that two-way communication is possible between the artificial cells and bacteria by responding to and secreting different AHLs (Lentini et al., 2017) (**Figure 5C**). They even devised a "cellular Turing test" where they compare the expression of quorum sensing genes of *V. fischeri* in the presence of either artificial cells or in a consortium of bacteria. They measure that the artificial cells would be 39% "life-like," but warn that this estimation does not consider that the artificial cells are not fully genetically encoded. Rampioni et al. (2018) developed synthetic cells which could send quorum sensing molecule C4-HSL to the pathogenic *P. aeruginosa*. Such synthetic cells could have interesting theranostic applications once equipped with additional sensing capabilities, such as those discussed in this section.

Two-way communication has been implemented in various contexts, from buffer conditions ideal for artificial cells, to more simple environments, such as water or PBS (Ding et al., 2018). Other communication modalities have also been explored, such as osmoregulation using a mechanosensitive MscL channel incorporated into liposomes, which opens due to membrane stress in hypotonic environments (Majumder et al., 2017; Garamella et al., 2019). Impressively, Berhanu et al. (2019) encapsulated proteoliposomes containing ATP synthase and bacteriorhodopsin inside liposomes (**Figure 5C**). The artificial cells were able to convert photons to a proton gradient inside the proteoliposomes and drive the synthesis of ATP by ATP synthase, fueling the TX-TL system, effectively making these artificial cells capable of light sensing and even photosynthetic activity.

More complex communication between liposomes was presented by Adamala et al. (2016), where they use artificial cells containing either bacterial or mammalian TX-TL systems and use small molecules to communicate between the prokaryotic and eukaryotic artificial cells containing different genetic circuits and cascades. However, the sensing of small molecules is limited to known transcriptional regulators or the theophylline riboswitch. Dwidar et al. (2019) engineered a riboswitch for the biologically relevant small molecule histamine into liposome-based artificial cells, which could respond to the presence of histamine in a variety of programmed ways (**Figure 5C**). Finally, liposome-based artificial cells expressing *Pseudomonas* exotoxin A were injected *in vivo* inside mice tumors and an increase in caspase activity was shown (Krinsky et al., 2017), suggesting their potential use in therapeutic or diagnostic applications.

One major limitation of liposomes is the difficulty in implementing signaling mediated by protein factors, as only small signaling molecules can cross the lipid bilayer with the help of the α-hemolysin pore. The polymeric capsules presented by Niederholtmeyer et al. (2018) (as discussed in section 3.3) are permeabilized by 200–300 nm pores, allowing for the exchange of polymerases and even ribosomes (**Figure 5D**). The authors show a basic form of quorum sensing where the reporter expression increases sharply at a threshold of 400 cell-mimics per 4.5 μL droplet of TX-TL.

Models have been recently proposed to help understand and implement communication using cell-free systems. These include studies of quorum sensing (Shum and Balazs, 2017) and the design of spatially distributed compartments (Menon and Krishnan, 2019). More complex spatial assemblies of compartments capable of communication (Villar et al., 2013), combined with computation by cell-free TX-TL genetic circuits or other *in vitro* computation methods [such as DNA strand displacement reactions (Joesaar et al., 2019), the Polymerase-Exonuclease-Nickase (PEN) DNA toolbox (Genot et al., 2016), or transcriptional "genelet" circuits (Weitz et al., 2014)], and integration with orthogonal technologies, such as electronics (Selberg et al., 2018) may one day allow for the bottom-up engineering of programmable tissues with distributed functional capabilities.

4. SCIENTIFIC OPPORTUNITIES

The technical achievements described above have given rise to new research directions involving cell-free gene expression systems. While the pioneering scientific applications of cell-free systems have been the deconstruction and elucidation of molecular biological pathways, today the research landscape is much more varied. Of the numerous active research directions (including biosensing; biomanufacturing; diagnostics; screening; minimal, semi-synthetic, synthetic, and artificial cells; education; and genetic, metabolic, and protein engineering), here we highlight three topics which are particularly relevant in the context of bottom-up construction using cell free systems.

4.1. Gene Expression Regulation

We still lack a complete appreciation for how cells encode, execute, and regulate gene expression (Phillips et al., 2019), which restricts our ability to predictively design new gene regulatory networks or efficiently compose existing modules. Ever since cell-free systems were used to uncover the central dogma, they have contributed profoundly to our understanding of gene expression (Zubay, 1973). In this line of research, PURE and extract systems bring complementary advantages. The PURE system is based on the core components required by the central dogma, and accordingly, can serve as the foundation from which we can build-to-understand basic aspects of gene expression. Extract-based systems serve as environments more similar to their *in vivo* counterparts, but lacking endogenous mRNA and DNA, effectively decoupling them from host processes that can convolute design implementation and data interpretation (Siegal-Gaskins et al., 2014). This section will highlight recent work that has advanced our understanding of gene expression using cell-free systems to operate at the fertile interface between *in vitro* biochemistry and *in vivo* cell biology.

Biology employs promoters to process input logic and initiate informed transcriptional output (Bintu et al., 2005), an operation believed to lie at the heart of cellular decision-making, yet for which we still possess an incomplete understanding. In investigations of transcriptional regulation, cell-free biology has the benefit of combining complex functional assays with controlled and accessible environments. In contrast to purely *in vitro* research of promoter DNA and transcription factor interactions, cell-free systems have the potential to bridge the divide between promoter occupancy and mRNA production, and help to improve our understanding of the factors that

drive transcription. Research from our laboratory by Swank et al. (2019) used cell-free extract to study the interaction between promoters and the largest family of transcription factors, zinc-fingers. They leveraged the compatibility of cell-free systems with high-throughput assays to quantify the binding-energy landscapes of several synthetic zinc-finger regulators (Blackburn et al., 2015). The precise tuning of repression strength was demonstrated, by mutating the consensus sequence or flanking regions to create small changes in binding affinity. This control facilitated the engineering of gene circuits; adjusting individual binding-site affinities was crucial for optimizing logic gate function for example. By fusing interaction domains to repressors, cooperativity was engineered between different regulators binding to promoters possessing two binding sites. With the appropriate placement of binding sites, it was shown that cooperativity greatly increased fold-repression and response non-linearity. Notably, the optimal spacing between cooperative repressors was tied to the helical twist of DNA. The repression strength was greatest if the spacing was such that both repressors would bind to the same face of DNA, while repression decayed to match the non-cooperative level as the spacing changed to place the repressors on opposing sides of the DNA. The combination of predictable cooperative interactions and tunable binding affinity guided the engineering of NAND, AND, and OR gates.

Moving away from intragenic composition, intergenic compositional context effects (referring to the position and orientation of entire genes relative to each other on DNA) have also been shown to influence transcriptional regulation (Rhee et al., 1999; Shearwin et al., 2005; Chong et al., 2014; Yeung et al., 2017). Yeung et al. (2017) arranged genes in convergent, divergent, and tandem orientations, and modeled the relationships (based on torsional stress) between supercoiling and transcription, to support a picture of how supercoiling mediates transcriptional coupling between physically connected genes. Cell-free experimentation served as an important part of the toolkit used to validate their hypotheses and provide evidence for their model. Using cell-free systems, the authors were able to adjust gyrase expression freely, to relax supercoiling and observe the impact on reporter-gene transcription, while avoiding any interference by host-mediated effects. Running cell-free experiments also allowed the authors to control against possible effects coming from plasmid replication. Furthermore, by employing the common practice of expressing linear DNA in cell-free systems (Sun et al., 2014), Yeung et al. were able to investigate the outcome of dissipating peripheral torsional stress, since the ends of linear DNA can rotate freely in response to transcription. Using their insights, the authors leverage supercoiling to build a convergently-oriented toggle switch, which shows a sharper threshold for switching between stable states than the original toggle switch with divergent genes (Gardner et al., 2000).

4.2. Resource Constraints as a Design Feature

A current focal point in synthetic biology research is understanding the failure of synthetic biomolecular circuitry due to the coupling of individual circuit components through their competition for the same gene expression resource, and the added coupling with host processes seen in *in vivo* implementations (Cardinale and Arkin, 2012; Carbonell-Ballestero et al., 2016; Qian et al., 2017). This category of problems, along with other context dependencies, leads to a reduction in design composability, worsening in proportion to circuit size. In recent years, cell-free systems have served as an important research tool to deepen our understanding of resource constraints. Siegal-Gaskins et al. (2014) exploited the freedom with which DNA concentrations can be varied in cell-free systems to independently quantify the levels of transcriptional and translational cross-talk in cell-free extract (**Figure 6**). They show that increasing the concentration of a second load construct in their reaction results in a decrease in the transcription and translation of the original reporter construct (**Figure 6B**). Loading was largely abolished when the second construct lacked a ribosome binding site (**Figure 6C**), suggesting that the resource bottleneck was caused primarily through increased protein translation. This result was later found to generalize to *E. coli.* (Gyorgy et al., 2015). The effect of an increase in load DNA concentration on reporter protein translation is dependent on the total DNA concentration in the system. At higher total DNA concentrations, translational coupling between genes increases. This was observed experimentally by Siegal-Gaskins et al., where increasing the load DNA in the cell-free system has a greater impact on reporter protein expression when the system contains higher reporter DNA concentrations (**Figure 6A**). In contrast, the way an increase in load DNA concentration affects transcription was found to be independent of DNA for a larger range of concentration values. This result highlights a limiting translation (but not transcription) capacity, which above a certain level of load, causes a simple resource trade-off between proteins being produced.

A promising direction to improve predictability when composing synthetic parts, in light of resource problems, is to take the primary resources into account in mathematical models, thereby considering non-regulatory interactions between components through resource sequestration (Gyorgy et al., 2015; Gorochowski et al., 2016; Qian et al., 2017). Gyorgy and Murray (2016) developed a model that used the previous cell-free extract data obtained by Siegal-Gaskins et al. to account for resource competition between genes. They were able to successfully predict expression profiles of multiple co-expressed parts, from data where these parts were characterized individually.

Ceroni et al. (2015) developed a "resource capacity monitor" assay implemented in *E. coli*, designed to obtain a measure of load imposed on the host by synthetic circuits. They genomically integrated a GFP gene whose output was used to infer the load imposed by synthetic circuitry, from the relative decrease in GFP when the load is expressed in the host. In a subsequent paper, the same group established a similar approach but using cell-free extract (Borkowski et al., 2018), with the reasoning that this avoids growth-dependencies, which cause results to be difficult to interpret since the burden affects growth rate and promotes mutations. They feed the resource-impact data generated from cell-free experiments into a computational model to estimate

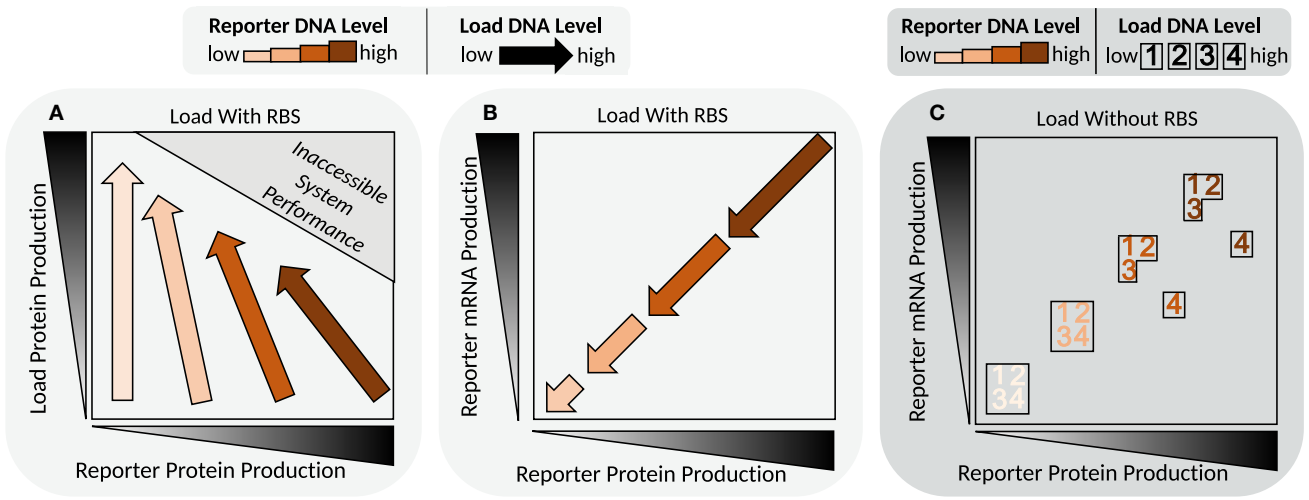


FIGURE 6 | Identifying resource constraints with cell-free gene expression. Schematic summary of results obtained by Siegal-Gaskins et al. (2014) **(A)** The authors observed that at greater reporter DNA concentrations, a given load imposed on the system will produce a larger decrease in reporter protein expression. **(B)** Loading decreased both transcriptional and translational output from cell-free extract. **(C)** When the load DNA lacked a ribosome binding site, loading had no effect, except for at the highest combined load and reporter DNA concentrations, suggesting that the bulk of the imposed load is realized through translational processes. In the figure, the relative positioning of numbers in a given box is arbitrary.

the resource cost that would be imposed on cells expressing synthetic circuitry employing the proteins they characterized. This strategy could be integrated with cell-free prototyping workflows, to improve the transfer of circuit design from cell-free to *in vivo*, by creating the opportunity to reject resource-demanding implementations. Furthermore, it is imaginable that cell-free extract systems could be adjusted to be resource-constrained in ways that better emulate a given host in order to improve predictive capacity.

Yelleswarapu et al. (2018) developed a clever oscillator design in cell-free extract that employs resource competition as a functional feature. Their delayed negative feedback topology leverages asymmetric competition between different sigma factors for core RNAP. Studies in this vein can help to improve our understanding of resource competition. By making resource sequestration a design element, circuit failure due to any "cross-talk" through this resource can be reframed as a problem of robust design. By learning design strategies that exhibit the desired behavior over large areas of parameter space, and by figuring out what models properly describe such circuits, we can learn to operate with, and perhaps around, the resource constraints in our biological systems. Even if such a circuit could be implemented successfully *in vivo* using an orthogonal RNAP and sigma-factor system, it would be difficult to untangle the signal of interest from the effects of the asymmetric load that would be imposed on the host. It would be interesting to investigate other resource-related phenomena, like modes of resource coupling or circuit failure following system overloading, using microfluidic chemostats (section 3.2.2), where reaction resources can be varied in a dynamic yet controllable manner.

One interesting strategy to alleviate the resource demands of translation is to implement transcriptional regulation with nucleic-acid hybridization interactions in cell-free systems (Chou and Shih, 2019). Chou et al. were able to do this by functionalizing T7 RNAP with single-stranded DNA, so that it can interact with cis-regulatory ssDNA domains on promoters, in a way that is dependent on nucleic-acid assemblies acting analogously to transcription factors. Although this may not directly advance our understanding of how biology encodes native promoters, making the link between gene regulatory networks and DNA strand-displacement reactions could reduce the cost of scaling up computation in genetic circuits, in order to fast-track the investigation of more sophisticated phenomena.

4.3. *In vitro* DNA Replication

Replication and propagation of genetic material is a key feature of life and is distributed among all living systems, and a robust *in vitro* implementation is crucial in particular for efforts in bottom-up construction of synthetic cells. While self-replicating systems including autocatalytic peptides, ribozyme replication, or RNA replicators have been established in the past (Ichihashi, 2019), it is crucial to develop a DNA replication system with regard to a transcription-translation based synthetic cell. Here we will focus on efforts to reconstitute DNA replication processes using cell-free TX-TL.

Organisms have evolved a great variety of mechanisms to replicate their DNA, with a broad range of complexity ranging from the eukaryotic replication machinery [consisting of at least five components some of which are further subdivided into complexes (Berg et al., 2012)], bacterial chromosome and plasmid replication, to simpler bacterial and viral replication strategies. Efforts to achieve *in vitro* reconstitution of DNA replication have focused mostly on the simpler systems.

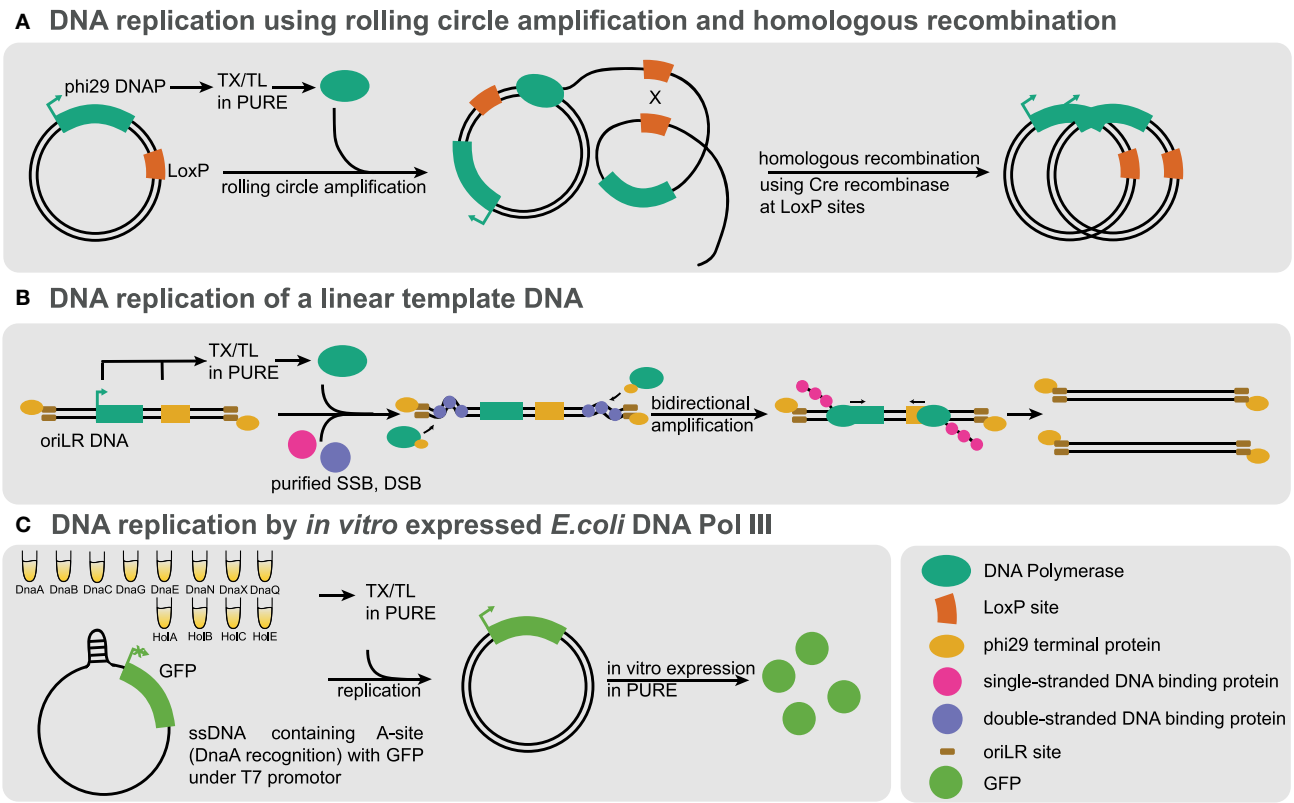


FIGURE 7 | Coupling DNA replication and cell-free gene expression. Schematic representation of methods to couple *in vitro* transcription-translation to DNA replication. **(A)** Sakatani et al. (2015, 2018) coupled *in vitro* expression of phi29 DNAP to rolling circle amplification of circular DNA and finally concluded their round of replication by re-circularizing the replicated DNA using homologous recombination by Cre recombinase at LoxP sites. **(B)** van Nies et al. (2018) reconstituted the native phi29 life cycle by replicating a linear DNA template flanked by oriLR sites expressing phi29 DNAP and TP *in vitro*, and adding recombinant SSB and DSB to the reaction. **(C)** Fujiwara et al. (2013) expressed the *E. coli* DNA Pol III holoenzyme *in vitro*. The enzyme was shown to replicate the second strand of a single stranded linear template containing an A-site; the resulting duplex DNA enables GFP expression.

In the 1980s, researchers reported *in vitro* DNA replication in crude cell extract of infected or transfected cells, including replication of plasmid RSF1010 in *P. aeruginosa* and *E. coli* (Diaz and Staudenbauer, 1982), and SV40 virus in monkey and human cell extract (Li and Kelly, 1984; Stillman and Gluzman, 1985; Wobbe et al., 1985). By the end of the decade, *in vitro* amplification of DNA became routine with the development of the polymerase chain reaction (PCR). Originally using the Klenow fragment of *E. coli* DNA Polymerase I, which was added anew after each hybridization step (Mullis and Faloona, 1987), the PCR method eventually adopted thermostable polymerases enabling continuous thermal cycling. However, repeated thermal cycling is not ideal for future applications involving synthetic cells, and so work on developing isothermal DNA replication methods remains of interest in this context.

Successful reconstitution of these isothermal machineries was eventually achieved *in vitro*, using partially or entirely recombinantly expressed and purified elements. Examples of these include the *E. coli* replication machinery (Kaguni and Kornberg, 1984; Su'etsugu et al., 2017), RSF1010 replication (Scherzinger et al., 1991), and viral replication systems including the phi29 (Blanco et al., 1994), T7 (Hürtgen et al., 2019), T4 (Schaerli et al., 2010), or SV40 (Waga et al., 1994) replication machineries.

The establishment of the PURE transcription-translation system has paved the way toward coupling *in vitro* protein expression with DNA replication, with the ultimate aim of reconstituting a self-sustaining system. Sakatani et al. (2015) expressed the phi29 DNA polymerase (DNAP) in PURE from a circular DNA template, which was then able to replicate the latter via a rolling circle amplification. The same group further developed their system based on a concept proposed by Forster and Church (2007), introducing recombinantly expressed Cre recombinase, that re-circularized an evolved form of the DNA template at the lox sites (Sakatani et al., 2018) (**Figure 7A**). They took advantage of the tunability of their home made PURE system by optimizing the NTP concentration, which is necessary for protein expression, yet was shown to inhibit DNA replication. van Nies et al. (2018) reported that PURE-expressed phi29 DNAP and terminal protein (TP) were able to amplify a linear DNA template encoding both proteins, in presence of recombinantly expressed single stranded and double stranded binding proteins (SSB, DSB) (**Figure 7B**). Those four proteins were shown to

be necessary and sufficient for DNA replication of the phi29 bacteriophage (Blanco et al., 1994; Salas et al., 2016).

Fujiwara et al. implemented an *in vitro* DNA replication machinery by mimicking *E. coli* DNA replication (**Figure 7C**). Using the PURE system, they expressed the machinery consisting of initiator (DnaA), helicase and helicase loader (DnaB and DnaC), DNA primase (DnaG), and the DNA polymerase III holoenzyme consisting of nine different proteins. By achieving the correct assembly of the holoenzyme in PURE, they furthermore showed the possibility to assemble a complex holoenzyme in the absence of chaperones by decreasing the cell-free expression temperature. The *in vitro*-expressed proteins were able to replicate an artificial gene circuit which expressed GFP in the PURE reaction system (Fujiwara et al., 2013).

Despite these advances, one major challenge on the way to implementing a self-sustaining DNA replication system remains to be addressed. Current approaches couple gene expression with DNA replication using only a couple of consecutive batch reactions. To ensure continuous replication in a future synthetic cell, it will be necessary to achieve continuous, multi-round replication, which could be explored for instance, in microfluidic chemostats as described in section 3.2.2. It has yet to be demonstrated that DNA replication can be achieved over many consecutive cycles, which may prove to be rather challenging as it appears that current DNA replication methods are rather inefficient and produce DNA in low-quantities (Sakatani et al., 2018; van Nies et al., 2018).

During long term replication, mutations will appear, among which some will enable the mutated DNA template to replicate faster than the original template, due to length or altered codon usage. This parasitic DNA may eventually out-compete the original DNA template, if no selection pressure is applied. Compartmentalization, as discussed above in section 3.3, may be a method to address this challenge, as discussed in Ichihashi (2019). Furthermore, implementation of a stable, continuous platform for *in vitro* DNA replication would enable the study of the evolutionary dynamics of molecular replicators, as the system is well-defined, simple, tunable, and does not rely on life-sustaining processes. This may additionally be linked with compartmentalization, where *in vitro* evolution of DNA polymerase using an error prone PCR approach has already been reported (Ghadessy et al., 2001).

In vitro coupling of transcription-translation with DNA replication is just at the beginning of its development, and it will be interesting to see what the limitations of the systems are. To our knowledge, only phi29 genomic DNA and plasmids have been replicated using coupled *in vitro* expression/replication systems to date. Successful determination of limits, such as size, accuracy, and energetic requirements to carry out *in vitro* replication may eventually enable the self-replication of all genes required to sustain a synthetic cell.

5. OUTLOOK

The bottom-up approach is but one way of addressing the formidable challenge of reliably building complex synthetic biological systems, and it will necessarily be combined with other complementary methods. However, the key principle of building to understand is undoubtedly a powerful motivation, and cell-free systems represent perhaps one of the best examples where this is currently being put into practice. While cell-free systems have historically been used to deconstruct biology, allowing its core processes to be elucidated, recent advances have led to its increasing application to construct biological systems.

Today, basic cell-free lysate systems are less of a black-box, and better characterization of their properties and preparation methods has made them an increasingly engineerable, and maybe more importantly, accessible tool. Recombinant systems have been the focus of increasing investigation as users demand more modularity and cost-effectiveness. Technological innovation in automation, microfluidics, and materials science have enabled increased throughput, dynamic control of steady-state reactions, and sophisticated compartmentalization strategies, while at the same time becoming accessible to more labs around the world.

However, there are also clear challenges ahead. Compartmentalizing cell-free reactions has exposed important physical effects, such as crowding and differential partitioning, which, while complex, may one day be harnessed to control the microscale spatial organization of gene expression. This level of fine control, exhibited by all cells, currently eludes us. Cell-free gene expression studies have unveiled a number of effects, such as physical properties of promoters, supercoiling and compositional context dependencies, and the ever-present resource burden of heterologous gene circuits. Replication studies have pointed out to the difficulty of achieving efficient DNA replication and protein synthesis in a cell-free reaction. And while increasingly complex communication systems have been implemented, the field is still in a nascent stage.

A common theme in constructing complex systems is emergence: as the system grows in size, effects appear which cannot be predicted by assessing the parts independently. In synthetic biology, these confounding effects currently stymie many efforts. But it is exactly because cell-free studies allow us to work at the interface between simple and complex systems that they are well-poised to address these issues. Ultimately, a thorough understanding of these effects will allow us to turn what are currently viewed as design constraints into design features, thereby expanding the scope and potential of synthetic biology.

AUTHOR CONTRIBUTIONS

BL wrote section 3.1, ZS wrote section 3.2, GM wrote section 3.3, AS wrote sections 4.1 and 4.2, LG wrote section 4.3. NL wrote sections 1, 2, and 5. SM and NL contributed to all sections. All authors edited the manuscript and approved it for publication.

FUNDING

This work was supported by Human Frontier Science Program Grant RGP0032/2015, the European Research Council under the European Union's Horizon 2020 research and innovation program Grant 723106, and a Swiss National Science Foundation grant (182019). NL was supported by a Chancellor's Fellowship from the University of Edinburgh. GM was supported by the M.D.-Ph.D. Program of the Swiss National Science Foundation (SNF, 323530_171144).

REFERENCES

Adamala, K. P., Martin-Alarcon, D. A., Guthrie-Honea, K. R., and Boyden, E. S. (2016). Engineering genetic circuit interactions within and between synthetic minimal cells. *Nat. Chem.* 9, 431–439. doi: 10.1038/nchem.2644

Agresti, J. J., Antipov, E., Abate, A. R., Ahn, K., Rowat, A. C., Baret, J.-C., et al. (2010). Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4004–4009. doi: 10.1073/pnas.0910781107

Alberti, S., Gladfelter, A., and Mittag, T. (2019). Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* 176, 419–434. doi: 10.1016/j.cell.2018.12.035

Altamura, E., Carrara, P., D'Angelo, F., Mavelli, F., and Stano, P. (2018). Extrinsic stochastic factors (solute partition) in gene expression inside lipid vesicles and lipid-stabilized water-in-oil droplets: a review. *Synth. Biol.* 3, 1–16. doi: 10.1093/synbio/ysy011

Anastasina, M., Terenin, I., Butcher, S. J., and Kainov, D. E. (2014). A technique to increase protein yield in a rabbit reticulocyte lysate translation system. *Biotechniques* 56, 36–39. doi: 10.2144/000114125

Arkin, A. (2008). Setting the standard in synthetic biology. *Nat. Biotechnol.* 26, 771–774. doi: 10.1038/nbt0708-771

Arriaga, L. R., Datta, S. S., Kim, S.-H., Amstad, E., Kodger, T. E., Monroy, F., et al. (2013). Ultrathin shell double emulsion templated giant unilamellar lipid vesicles with controlled microdomain formation. *Small* 10, 950–956. doi: 10.1002/smll.201301904

Aufinger, L., and Simmel, F. C. (2018). Artificial gel-based organelles for spatial organization of cell-free gene expression reactions. *Angew. Chem. Int. Ed.* 57, 17245–17248. doi: 10.1002/anie.201809374

Balch, W. E., Dunphy, W. G., Braell, W. A., and Rothman, J. E. (1984). Reconstitution of the transport of protein between successive compartments of the golgi measured by the coupled incorporation of N-acetylglucosamine. *Cell* 39, 405–416. doi: 10.1016/0092-8674(84)90019-9

Berg, J. M. J. M., Tymoczko, J. L., and Stryer, L. (2012). *Biochemistry*. New York, NY: W. H. Freeman.

Berhanu, S., Ueda, T., and Kuruma, Y. (2019). Artificial photosynthetic cell producing energy for protein synthesis. *Nat. Commun.* 10, 1325–1310. doi: 10.1038/s41467-019-09147-4

Bintu, L., Buchler, N. E., Garcia, H. G., Gerland, U., Hwa, T., Kondev, J., et al. (2005). Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* 15, 116–124. doi: 10.1016/j.gde.2005.02.007

Blackburn, M. C., Petrova, E., Correia, B. E., and Maerkl, S. J. (2015). Integrating gene synthesis and microfluidic protein analysis for rapid protein engineering. *Nucleic Acids Res.* 44:e68. doi: 10.1093/nar/gkv1497

Blanco, L., Lázaro, J. M., de Vega, M., Bonnin, A., and Salas, M. (1994). Terminal protein-primed DNA amplification. *Proc. Natl. Acad. Sci. U.S.A.* 91, 12198–12202. doi: 10.1073/pnas.91.25.12198

Blanken, D., van Nies, P., and Danelon, C. (2019). Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes. *Phys. Biol.* 16, 045002–045015. doi: 10.1088/1478-3975/ab0c62

Borkowski, O., Bricio, C., Murgiano, M., Rothschild-Mancinelli, B., Stan, G.-B., and Ellis, T. (2018). Cell-free prediction of protein expression costs for growing cells. *Nat. Commun.* 9:1457. doi: 10.1038/s41467-018-03970-x

Buchner, E. (1897). Alkoholische Gärung ohne Hefezellen. *Berich. Deutsch. Chem. Gesellsch.* 30, 1110–1113. doi: 10.1002/cber.189703001215

Bujara, M., Schümperli, M., Pellaux, R., Heinemann, M., and Panke, S. (2011). Optimization of a blueprint for *in vitro* glycolysis by metabolic real-time analysis. *Nat. Chem. Biol.* 7, 271–277. doi: 10.1038/nchembio.541

Byun, J.-Y., Lee, K.-H., Lee, K.-Y., Kim, M.-G., and Kim, D.-M. (2013). In-gel expression and *in situ* immobilization of proteins for generation of three dimensional protein arrays in a hydrogel matrix. *Lab Chip* 13, 886–891. doi: 10.1039/c2lc41137g

Cai, Q., Hanson, J. A., Steiner, A. R., Tran, C., Masikat, M. R., Chen, R., et al. (2015). A simplified and robust protocol for immunoglobulin expression in Escherichia coli cell-free protein synthesis systems. *Biotechnol. Prog.* 31, 823–831. doi: 10.1002/btpr.2082

Calhoun, K. A., and Swartz, J. R. (2005a). An economical method for cell-free protein synthesis using glucose and nucleoside monophosphates. *Biotechnol. Prog.* 21, 1146–1153. doi: 10.1021/bp050052y

Calhoun, K. A., and Swartz, J. R. (2005b). Energizing cell-free protein synthesis with glucose metabolism. *Biotechnol. Bioeng.* 90, 606–613. doi: 10.1002/bit.20449

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

Carbonell-Ballestero, M., Garcia-Ramallo, E., Montañez, R., Rodriguez-Caso, C., and Macía, J. (2016). Dealing with the genetic load in bacterial synthetic biology circuits: convergences with the Ohm's law. *Nucleic Acids Res.* 44, 496–507. doi: 10.1093/nar/gkv1280

Cardinale, S., and Arkin, A. P. (2012). Contextualizing context for synthetic biology–identifying causes of failure of synthetic biological systems. *Biotechnol. J.* 7, 856–866. doi: 10.1002/biot.201200085

Carrara, P., Altamura, E., D'Angelo, F., Mavelli, F., and Stano, P. (2018). Measurement and numerical modeling of cell-free protein synthesis: combinatorial block-variants of the PURE system. *Data* 3:41. doi: 10.3390/data3040041

Caschera, F., and Noireaux, V. (2014a). Integration of biological parts toward the synthesis of a minimal cell. *Curr. Opin. Chem. Biol.* 22, 85–91. doi: 10.1016/j.cbpa.2014.09.028

Caschera, F., and Noireaux, V. (2014b). Synthesis of 2.3 mg/ml of protein with an all *Escherichia coli* cell-free transcription–translation system. *Biochimie* 99, 162–168. doi: 10.1016/j.biochi.2013.11.025

Caschera, F., and Noireaux, V. (2015). A cost-effective polyphosphate-based metabolism fuels an all *E. coli* cell-free expression system. *Metab. Eng.* 27, 29–37. doi: 10.1016/j.ymben.2014.10.007

Ceroni, F., Algar, R., Stan, G. B., and Ellis, T. (2015). Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* 12, 415–418. doi: 10.1038/nmeth.3339

Chiao, A. C., Murray, R. M., and Sun, Z. Z. (2016). Development of prokaryotic cell-free systems for synthetic biology. *bioRxiv* 1–38. doi: 10.1101/048710

Chong, S., Chen, C., Ge, H., and Xie, X. S. (2014). Mechanism of transcriptional bursting in bacteria. *Cell* 158, 314–326. doi: 10.1016/j.cell.2014.05.038

Chou, L., and Shih, W. (2019). Cell-free transcriptional regulation via nucleic-acid-based transcription factors. *biorxiv*. doi: 10.1101/644021

Cole, S. D., Beabout, K., Turner, K. B., Smith, Z. K., Funk, V. L., Harbaugh, S. V., et al. (2019). Quantification of interlaboratory cell-free protein synthesis variability. *ACS Synth. Biol.* 8, 2080–2091. doi: 10.1021/acssynbio.9b00178

Contreras-Llano, L. E., and Tan, C. (2018). High-throughput screening of biomolecules using cell-free gene expression systems. *Synth. Biol.* 3:ysy012. doi: 10.1093/synbio/ysy012

Courtois, F., Olguin, L. F., Whyte, G., Bratton, D., Huck, W. T. S., Abell, C., et al. (2008). An integrated device for monitoring time-dependent *in vitro* expression from single genes in picolitre droplets. *ChemBioChem* 9, 439–446. doi: 10.1002/cbic.200700536

Cui, N., Zhang, H., Schneider, N., Tao, Y., Asahara, H., Sun, Z., et al. (2016). A mix-and-read drop-based *in vitro* two-hybrid method for screening high-affinity peptide binders. *Sci. Rep.* 6, 1–10. doi: 10.1038/srep22575

Cuomo, F., Ceglie, A., De Leonardis, A., and Lopez, F. (2019). Polymer capsules for enzymatic catalysis in confined environments. *Catalysts* 9, 1–18. doi: 10.3390/catal9010001

Damiati, S., Mhanna, R., Kodzius, R., Ehmoser, E.-K., Damiati, S., Mhanna, R., et al. (2018). Cell-free approaches in synthetic biology utilizing microfluidics. *Genes* 9:144. doi: 10.3390/genes9030144

Deng, N.-N., Vibhute, M. A., Zheng, L., Zhao, H., Yelleswarapu, M., and Huck, W. T. S. (2018). Macromolecularly crowded protocells from reversibly shrinking monodisperse liposomes. *J. Am. Chem. Soc.* 140, 7399–7402. doi: 10.1021/jacs.8b03123

Deng, N.-N., Yelleswarapu, M., and Huck, W. T. S. (2016). Monodisperse uni- and multicompartment liposomes. *J. Am. Chem. Soc.* 138, 7584–7591. doi: 10.1021/jacs.6b02107

Deng, N.-N., Yelleswarapu, M., Zheng, L., and Huck, W. T. S. (2017). Microfluidic assembly of monodisperse vesosomes as artificial cell models. *J. Am. Chem. Soc.* 139, 587–590. doi: 10.1021/jacs.6b10977

Deshpande, S., Caspi, Y., Meijering, A. E., and Dekker, C. (2016). Octanol-assisted liposome assembly on chip. *Nat. Commun.* 7, 1–9. doi: 10.1038/ncomms10447

Deshpande, S., and Dekker, C. (2018). On-chip microfluidic production of cell-sized liposomes. *Nat. Protoc.* 13, 856–874. doi: 10.1038/nprot.2017.160

DeVries, J. K., and Zubay, G. (1967). DNA-directed peptide synthesis II. The synthesis of the alpha-fragment of the enzyme beta-galactosidase. *Proc. Natl. Acad. Sci. U.S.A.* 57, 1010–1012. doi: 10.1073/pnas.57.4.1010

Diaz, R., and Staudenbauer, W. L. (1982). Replication of the broad host range plasmid RSF1010 in cell-free extracts of *Escherichia coli* and *Pseudomonas aeruginosa*. *Nucleic Acids Res.* 10, 4687–4702. doi: 10.1093/nar/10.15.4687

Didovyk, A., Tonooka, T., Tsimring, L., and Hasty, J. (2017). Rapid and scalable preparation of bacterial lysates for cell-free gene expression. *ACS Synth. Biol.* 6, 2198–2208. doi: 10.1021/acssynbio.7b00253

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. Mater. Interfaces* 10, 30137–30146. doi: 10.1021/acsami.8b10029

Doerr, A., de Reus, E., van Nies, P., van der Haar, M., Wei, K., Kattan, J., et al. (2019). Modelling cell-free RNA and protein synthesis with minimal systems. *Phys. Biol.* 16:025001. doi: 10.1088/1478-3975/aaf33d

Dubuc, E., Pieters, P. A., van der Linden, A. J., van Hest, J. C., Huck, W. T., and de Greef, T. F. (2019). Cell-free microcompartmentalised transcription–translation for the prototyping of synthetic communication networks. *Curr. Opin. Biotechnol.* 58, 72–80. doi: 10.1016/j.copbio.2018.10.006

Dupin, A., and Simmel, F. C. (2019). Signalling and differentiation in emulsion-based multi-compartmentalized *in vitro* gene circuits. *Nat. Chem.* 11, 32–39. doi: 10.1038/s41557-018-0174-9

Dwidar, M., Seike, Y., Kobori, S., Whitaker, C., Matsuura, T., and Yokobayashi, Y. (2019). Programmable artificial cells using histamine-responsive synthetic riboswitch. *J. Am. Chem. Soc.* 141, 11103–11114. doi: 10.1021/jacs.9b03300

Efrat, Y., Tayar, A. M., Daube, S. S., Levy, M., and Bar-Ziv, R. H. (2018). Electric-field manipulation of a compartmentalized cell-free gene expression reaction. *ACS Synth. Biol.* 7, 1829–1833. doi: 10.1021/acssynbio.8b00160

Elani, Y., Law, R. V., and Ces, O. (2014). Vesicle-based artificial cells as chemical microreactors with spatially segregated reaction pathways. *Nat. Commun.* 5, 1–5. doi: 10.1038/ncomms6305

Elowitz, M., and Lim, W. A. (2010). Build life to understand it. *Nature* 468, 889–890. doi: 10.1038/468889a

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

Etienne, G., Vian, A., Biočanin, M., Deplancke, B., and Amstad, E. (2018). Cross-talk between emulsion drops: how are hydrophilic reagents transported across oil phases? *Lab Chip* 18, 3903–3912. doi: 10.1039/C8LC01000E

Failmezger, J. (2018). *Understanding Limitations to Increased Potential of Cell-Free Protein Synthesis*. Stuttgart: University of Stuttgart.

Failmezger, J., Rauter, M., Nitschel, R., Kraml, M., and Siemann-Herzberg, M. (2017). Cell-free protein synthesis from non-growing, stressed *Escherichia coli*. *Sci. Rep.* 7:16524. doi: 10.1038/s41598-017-16767-7

Fallah-Araghi, A., Baret, J.-C., Ryckelynck, M., and Griffiths, A. D. (2012). A completely *in vitro* ultrahigh-throughput droplet-based microfluidic screening system for protein engineering and directed evolution. *Lab Chip* 12, 882–811. doi: 10.1039/c2lc21035e

Fan, J., Villarreal, F., Weyers, B., Ding, Y., Tseng, K. H., Li, J., et al. (2017). Multi-dimensional studies of synthetic genetic promoters enabled by microfluidic impact printing. *Lab Chip* 17, 2198–2207. doi: 10.1039/C7LC00382J

Forster, A. C., and Church, G. M. (2007). Synthetic biology projects *in vitro*. *Genome Res.* 17, 1–6. doi: 10.1101/gr.5776007

Foshag, D., Henrich, E., Hiller, E., Schäfer, M., Kerger, C., Burger-Kentischer, A., et al. (2018). The *E. coli* S30 lysate proteome: a prototype for cell-free protein production. *New Biotechnol.* 40, 245–260. doi: 10.1016/j.nbt.2017.09.005

Fujiwara, K., Katayama, T., and Nomura, S. I. M. (2013). Cooperative working of bacterial chromosome replication proteins generated by a reconstituted protein expression system. *Nucleic Acids Res.* 41, 7176–7183. doi: 10.1093/nar/gkt489

Ganzinger, K. A., and Schwille, P. (2019). More from less–bottom-up reconstitution of cell biology. *J Cell Sci.* 132:jcs227488. doi: 10.1242/jcs.227488

Garamella, J., Majumder, S., Liu, A. P., and Noireaux, V. (2019). An adaptive synthetic cell based on mechanosensing, biosensing, and inducible gene circuits. *ACS Synth. Biol.* 8, 1913–1920. doi: 10.1021/acssynbio.9b00204

Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403, 339–342. doi: 10.1038/35002131

Garenne, D., Beisel, C. L., and Noireaux, V. (2019). Characterization of the all-*E. coli* transcription-translation system myTXTL by mass spectrometry. *Rapid Commun. Mass Spectr.* 33, 1036–1048. doi: 10.1002/rcm.8438

Garenne, D., and Noireaux, V. (2019). Cell-free transcription–translation: engineering biology from the nanometer to the millimeter scale. *Curr. Opin. Biotechnol.* 58, 19–27. doi: 10.1016/j.copbio.2018.10.007

Ge, X., Luo, D., and Xu, J. (2011). Cell-free protein expression under macromolecular crowding conditions. *PLoS ONE* 6:e28707. doi: 10.1371/journal.pone.0028707

Genot, A. J., Baccouche, A., Sieskind, R., Aubert-Kato, N., Bredeche, N., Bartolo, J. F., et al. (2016). High-resolution mapping of bifurcations in nonlinear biochemical circuits. *Nat. Chem.* 8, 760–767. doi: 10.1038/nchem.2544

Georgi, V., Georgi, L., Blechert, M., Bergmeister, M., Zwanzig, M., Wüstenhagen, D. A., et al. (2016). On-chip automation of cell-free protein synthesis: new opportunities due to a novel reaction mode. *Lab Chip* 16, 269–281. doi: 10.1039/C5LC00700C

Gerber, D., Maerkl, S. J., and Quake, S. R. (2009). An *in vitro* microfluidic approach to generating protein-interaction networks. *Nat. Methods* 6, 71–74. doi: 10.1038/nmeth.1289

Ghadessy, F. J., Ong, J. L., and Holliger, P. (2001). Directed evolution of polymerase function by compartmentalized self-replication. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4552–4557. doi: 10.1073/pnas.071052198

Göpfrich, K., Platzman, I., and Spatz, J. P. (2018). Mastering complexity: towards bottom-up construction of multifunctional eukaryotic synthetic cells. *Trends Biotechnol.* 36, 938–951. doi: 10.1016/j.tibtech.2018.03.008

Gorochowski, T. E., Avcilar-Kucukgoze, I., Bovenberg, R. A., Roubos, J. A., and Ignatova, Z. (2016). A minimal model of ribosome allocation dynamics captures trade-offs in expression between endogenous and synthetic genes. *ACS Synth. Biol.* 5, 710–720. doi: 10.1021/acssynbio.6b00040

Gregorio, N. E., Levine, M. Z., and Oza, J. P. (2019). A user's guide to cell-free protein synthesis. *Methods Protoc.* 2:24. doi: 10.3390/mps2010024

Griffiths, A. D., and Tawfik, D. S. (2003). Directed evolution of an extremely fast phosphotriesterase by *in vitro* compartmentalization. *EMBO J.* 22, 24–35. doi: 10.1093/emboj/cdg014

Gruner, P., Riechers, B., Semin, B., Lim, J., Johnston, A., Short, K., et al. (2019). Controlling molecular transport in minimal emulsions. *Nat. Commun.* 7, 1–9. doi: 10.1038/ncomms10392

Gyorgy, A., Jiménez, J. I., Yazbek, J., Huang, H. H., Chung, H., Weiss, R., et al. (2015). Isocost lines describe the cellular economy of genetic circuits. *Biophys. J.* 109, 639–646. doi: 10.1016/j.bpj.2015.06.034

Gyorgy, A., and Murray, R. M. (2016). "Quantifying resource competition and its effects in the TX-TL system," in *2016 IEEE 55th Conference on Decision and Control, CDC 2016* (Las Vegas, NV). doi: 10.1109/CDC.2016.7798775

Hahn, G.-H., Asthana, A., Kim, D.-M., and Kim, D.-P. (2007). A continuous-exchange cell-free protein synthesis system fabricated on a chip. *Anal. Biochem.* 365, 280–282. doi: 10.1016/j.ab.2007.03.025

Hansen, M. M. K., Meijer, L. H. H., Spruijt, E., Maas, R. J. M., Rosquelles, M. V., Groen, J., et al. (2015). Macromolecular crowding creates heterogeneous environments of gene expression in picolitre droplets. *Nat. Nanotechnol.* 11, 191–197. doi: 10.1038/nnano.2015.243

Hein, C., Henrich, E., Orbán, E., Dötsch, V., and Bernhard, F. (2014). Hydrophobic supplements in cell-free systems: designing artificial environments for membrane proteins. *Eng. Life Sci.* 14, 365–379. doi: 10.1002/elsc.201300050

Heinemann, M., and Panke, S. (2006). Synthetic biology–putting engineering into biology. *Bioinformatics* 22, 2790–2799. doi: 10.1093/bioinformatics/btl469

Henrich, E., Hein, C., Dötsch, V., and Bernhard, F. (2015). Membrane protein production in *Escherichia coli* cell-free lysates. *FEBS Lett.* 589, 1713–1722. doi: 10.1016/j.febslet.2015.04.045

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, 1–4. doi: 10.1038/s41467-019-10079-2

Ho, K. K. Y., Lee, J. W., Durand, G., Majumder, S., and Liu, A. P. (2017). Protein aggregation with poly(vinyl) alcohol surfactant reduces double emulsion-encapsulated mammalian cell-free expression. *PLoS ONE* 12:e0174689. doi: 10.1371/journal.pone.0174689

Ho, K. K. Y., Murray, V. L., and Liu, A. P. (2015). Engineering artificial cells by combining HeLa-based cell-free expression and ultrathin double emulsion template. *Methods Cell Biol.* 128, 303–318. doi: 10.1016/bs.mcb.2015.01.014

Hong, S. H., Kwon, Y.-C., Martin, R. W., Des Soye, B. J., de Paz, A. M., Swonger, K. N., et al. (2015). Improving cell-free protein synthesis through genome engineering of *Escherichia coli* lacking release factor 1. *Chembiochem* 16, 844–853. doi: 10.1002/cbic.201402708

Hori, Y., Kantak, C., Murray, R. M., and Abate, A. R. (2017). Cell-free extract based optimization of biomolecular circuits with droplet microfluidics. *Lab Chip* 17, 3037–3042. doi: 10.1039/C7LC00552K

Horiya, S., Bailey, J. K., and Krauss, I. J. (2017). "Chapter 4–directed evolution of glycopeptides using mRNA display," in *Methods in Enzymology, Volume 597 of Chemical Glycobiology Part A. Synthesis, Manipulation and Applications of Glycans*, ed B. Imperiali (Cambridge, MA: Academic Press), 83–141. doi: 10.1016/bs.mie.2017.06.029

Hsiao, V., Swaminathan, A., and Murray, R. M. (2018). Control theory for synthetic biology: recent advances in system characterization, control design, and controller implementation for synthetic biology. *IEEE Control Syst.* 38, 32–62. doi: 10.1109/MCS.2018.2810459

Hurst, G. B., Asano, K. G., Doktycz, C. J., Consoli, E. J., Doktycz, W. L., Foster, C. M., et al. (2017). Proteomics-based tools for evaluation of cell-free protein synthesis. *Anal. Chem.* 89, 11443–11451. doi: 10.1021/acs.analchem.7b02555

Hürtgen, D., Mascarenhas, J., Heymann, M., Murray, S. M., Schwille, P., and Sourjik, V. (2019). Reconstitution and coupling of DNA replication and segregation in a biomimetic system. *ChemBioChem* 20, 2633–2642. doi: 10.1002/cbic.201900299

Ichihashi, N. (2019). What can we learn from the construction of *in vitro* replication systems? *Ann. N. Y. Acad. Sci.* 1447, 144–156. doi: 10.1111/nyas.14042

Ishikawa, K., Sato, K., Shima, Y., Urabe, I., and Yomo, T. (2004). Expression of a cascading genetic network within liposomes. *FEBS Lett.* 576, 387–390. doi: 10.1016/j.febslet.2004.09.046

Jacobs, M. L., Boyd, M. A., and Kamat, N. P. (2019). Diblock copolymers enhance folding of a mechanosensitive membrane protein during cell-free expression. *Proc. Natl. Acad. Sci. U.S.A.* 116, 4031–4036. doi: 10.1073/pnas.1814775116

Jaroentomeechai, T., Stark, J. C., Natarajan, A., Glasscock, C. J., Yates, L. E., Hsu, K. J., et al. (2018). Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nat. Commun.* 9:2686. doi: 10.1038/s41467-018-05620-8

Jermutus, L., Ryabova, L. A., and Plückthun, A. (1998). Recent advances in producing and selecting functional proteins by using cell-free translation. *Curr. Opin. Biotechnol.* 9, 534–548. doi: 10.1016/S0958-1669(98)80042-6

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. doi: 10.1038/msb.2008.57

Jiao, Y., Liu, Y., Luo, D., Huck, W. T., and Yang, D. (2018). Microfluidic-assisted fabrication of clay microgels for cell-free protein synthesis. *ACS Appl. Mater. Interfaces* 10, 29308–29313. doi: 10.1021/acsami.8b09324

Joesaar, A., Yang, S., Gels, B. B. X., Linden, A., Pieters, P., Kumar, B. V. V. S. P., et al. (2019). DNA-based communication in populations of synthetic protocells. *Nat. Nanotechnol.* 21, 1–12. doi: 10.1038/s41565-019-0399-9

Kaguni, J. M., and Kornberg, A. (1984). Replication initiated at the origin (oriC) of the *E. coli* chromosome reconstituted with purified enzymes. *Cell* 38, 183–190. doi: 10.1016/0092-8674(84)90539-7

Kahn, J. S., Ruiz, R. C. H., Sureka, S., Peng, S., Derrien, T. L., An, D., et al. (2016). DNA microgels as a platform for cell-free protein expression and display. *Biomacromolecules* 17, 2019–2026. doi: 10.1021/acs.biomac.6b00183

Karig, D. K., Jung, S.-Y., Srijanto, B., Collier, C. P., and Simpson, M. L. (2013). Probing cell-free gene expression noise in femtoliter volumes. *ACS Synth. Biol.* 2, 497–505. doi: 10.1021/sb400028c

Karim, A. S., Dudley, Q. M., and Jewett, M. C. (2016). Cell-free synthetic systems for metabolic engineering and biosynthetic pathway prototyping. *Ind. Biotechnol.* 1, 125–148. doi: 10.1002/9783527807796.ch4

Karim, A. S., and Jewett, M. C. (2018). Cell-free synthetic biology for pathway prototyping. *Methods Enzymol.* 608, 31–57. doi: 10.1016/bs.mie.2018.04.029

Karzbrun, E., Tayar, A. M., Noireaux, V., and Bar-Ziv, R. H. (2014). Programmable on-chip DNA compartments as artificial cells. *Science* 345, 829–832. doi: 10.1126/science.1255550

Kato, A., Yanagisawa, M., Sato, Y. T., Fujiwara, K., and Yoshikawa, K. (2012). Cell-sized confinement in microspheres accelerates the reaction of gene expression. *Sci. Rep.* 2, 1172–1175. doi: 10.1038/srep00283

Kazuta, Y., Adachi, J., Matsuura, T., Ono, N., Mori, H., and Yomo, T. (2008). Comprehensive analysis of the effects of *Escherichia coli* ORFs on protein translation reaction. *Mol. Cell. Proteomics* 7, 1530–1540. doi: 10.1074/mcp.M800051-MCP200

Kazuta, Y., Matsuura, T., Ichihashi, N., and Yomo, T. (2014). Synthesis of milligram quantities of proteins using a reconstituted *in vitro* protein synthesis system. *J. Biosci. Bioeng.* 118, 554–557. doi: 10.1016/j.jbiosc.2014.04.019

Kelwick, R., Webb, A. J., MacDonald, J. T., and Freemont, P. S. (2016). Development of a *Bacillus subtilis* cell-free transcription-translation system for prototyping regulatory elements. *Metab. Eng.* 38, 370–381. doi: 10.1016/j.ymben.2016.09.008

Kempf, N., Remes, C., Ledesch, R., Züchner, T., Höfig, H., Ritter, I., et al. (2017). A novel method to evaluate ribosomal performance in cell-free protein synthesis systems. *Sci. Rep.* 7:46753. doi: 10.1038/srep46753

Khammash, M. (2016). An engineering viewpoint on biological robustness. *BMC Biol.* 14:22. doi: 10.1186/s12915-016-0241-x

Khnouf, R., Beebe, D. J., and Fan, Z. H. (2009). Cell-free protein expression in a microchannel array with passive pumping. *Lab Chip* 9, 56–61. doi: 10.1039/B808034H

Khnouf, R., Olivero, D., Jin, S., and Fan, Z. H. (2010). Miniaturized fluid array for high-throughput protein expression. *Biotechnol. Prog.* 26, 1590–1596. doi: 10.1002/btpr.474

Kim, D.-M., and Choi, C.-Y. (1996). A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol. Prog.* 12, 645–649. doi: 10.1021/bp960052l

Kim, H.-C., Kim, T.-W., and Kim, D.-M. (2011). Prolonged production of proteins in a cell-free protein synthesis system using polymeric carbohydrates as an energy source. *Process Biochem.* 46, 1366–1369. doi: 10.1016/j.procbio.2011.03.008

Kim, J., and Winfree, E. (2011). Synthetic *in vitro* transcriptional oscillators. *Mol. Syst. Biol.* 7, 1–15. doi: 10.1038/msb.2010.119

Kim, R. G., and Choi, C. Y. (2000). Expression-independent consumption of substrates in cell-free expression system from *Escherichia coli*. *J. Biotechnol.* 84, 27–32. doi: 10.1016/S0168-1656(00)00326-6

Kim, T.-W., Keum, J.-W., Oh, I.-S., Choi, C.-Y., Park, C.-G., and Kim, D.-M. (2006). Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system. *J. Biotechnol.* 126, 554–561. doi: 10.1016/j.jbiotec.2006.05.014

Kim, T.-W., Kim, H.-C., Oh, I.-S., and Kim, D.-M. (2008). A highly efficient and economical cell-free protein synthesis system using the S12 extract of *Escherichia coli*. *Biotechnol. Bioprocess Eng.* 13, 464–469. doi: 10.1007/s12257-008-0139-8

Koch, M., Faulon, J.-L., and Borkowski, O. (2018). Models for cell-free synthetic biology: make prototyping easier, better, and faster. *Front. Bioeng. Biotechnol.* 6:182. doi: 10.3389/fbioe.2018.00182

Krinsky, N., Kaduri, M., Zinger, A., Shainsky-Roitman, J., Goldfeder, M., Benhar, I., et al. (2017). Synthetic cells synthesize therapeutic proteins inside tumors. *Adv. Healthc. Mater.* 7:e1701163. doi: 10.1002/adhm.201701163

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* 31, 147–157. doi: 10.1016/0092-8674(82)90414-7

Kung, H. F., Redfield, B., Treadwell, B. V., Eskin, B., Spears, C., and Weissbach, H. (1977). DNA-directed *in vitro* synthesis of beta-galactosidase. Studies with purified factors. *J. Biol. Chem.* 252, 6889–6894.

Kuruma, Y., and Ueda, T. (2015). The PURE system for the cell-free synthesis of membrane proteins. *Nat. Protoc.* 10, 1328–1344. doi: 10.1038/nprot.2015.082

Kwok, R. (2010). Five hard truths for synthetic biology. *Nature* 463, 288–290. doi: 10.1038/463288a

Kwon, Y.-C., Hahn, G.-H., Huh, K. M., and Kim, D.-M. (2008). Synthesis of functional proteins using *Escherichia coli* extract entrapped in calcium alginate microbeads. *Anal. Biochem.* 373, 192–196. doi: 10.1016/j.ab.2007.10.045

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. doi: 10.1038/srep08663

Lai, S. N., Zhou, X., Ouyang, X., Zhou, H., Liang, Y., Xia, J., et al. (2020). Artificial cells capable of long-lived protein synthesis by using aptamer grafted polymer hydrogel. *ACS Synth. Biol.* 9, 76–83. doi: 10.1021/acssynbio.9b00338

Lamborg, M. R., and Zamecnik, P. C. (1960). Amino acid incorporation into protein by extracts of *E. coli*. *Biochim. Biophys. Acta* 42, 206–211. doi: 10.1016/0006-3002(60)90782-4

Lavickova, B., and Maerkl, S. J. (2019). A simple, robust, and low-cost method to produce the PURE cell-free system. *ACS Synth. Biol.* 8, 455–462. doi: 10.1021/acssynbio.8b00427

Lederman, M., and Zubay, G. (1967). DNA-directed peptide synthesis I. A comparison of T2 and *Escherichia coli* DNA-directed peptide synthesis in two cell-free systems. *Biochim. Biophys. Acta* 149, 253–258. doi: 10.1016/0005-2787(67)90706-X

Lee, K.-H., Lee, K.-Y., Byun, J.-Y., Kim, B.-G., and Kim, D.-M. (2012). On-bead expression of recombinant proteins in an agarose gel matrix coated on a glass slide. *Lab Chip* 12, 1605–1610. doi: 10.1039/c2lc21239k

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 Central Sci.* 3, 117–123. doi: 10.1021/acscentsci.6b00330

Lentini, R., Santero, S. P., Chizzolini, F., Cecchi, D., Fontana, J., Marchioretto, M., et al. (2014). Integrating artificial with natural cells to translate chemical messages that direct *E. coli* behaviour. *Nat. Commun.* 5:4012. doi: 10.1038/ncomms5012

Li, J., Gu, L., Aach, J., and Church, G. M. (2014). Improved cell-free RNA and protein synthesis system. *PLoS ONE* 9:e106232. doi: 10.1371/journal.pone.0106232

Li, J., Zhang, C., Huang, P., Kuru, E., Forster-Benson, E. T. C., Li, T., et al. (2017). Dissecting limiting factors of the protein synthesis using recombinant elements (PURE) system. *Translation* 5:e1327006. doi: 10.1080/21690731.2017.1327006

Li, J. J., and Kelly, T. J. (1984). Simian virus 40 DNA replication *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* 81, 6973–6977. doi: 10.1073/pnas.81.22.6973

Lim, S. Y., Kim, K.-O., Kim, D.-M., and Park, C. B. (2009). Silica-coated alginate beads for *in vitro* protein synthesis via transcription/translation machinery encapsulation. *J. Biotechnol.* 143, 183–189. doi: 10.1016/j.jbiotec.2009.07.006

Liu, A. P. (2019). The rise of bottom-up synthetic biology and cell-free biology. *Phys. Biol.* 16:040201. doi: 10.1088/1478-3975/ab1bed

Liu, A. P., and Fletcher, D. A. (2009). Biology under construction: *in vitro* reconstitution of cellular function. *Nat. Rev. Mol. Cell Biol.* 10, 644–650. doi: 10.1038/nrm2746

Liu, D. V., Zawada, J. F., and Swartz, J. R. (2005). Streamlining *Escherichia coli* S30 extract preparation for economical cell-free protein synthesis. *Biotechnol. Prog.* 21, 460–465. doi: 10.1021/bp049789y

Maddalena, L. L., Niederholtmeyer, H., Turtola, M., Swank, Z. N., Belogurov, G. A., and Maerkl, S. J. (2016). GreA and GreB enhance expression of *Escherichia coli* RNA polymerase promoters in a reconstituted transcription–translation system. *ACS Synth. Biol.* 5, 929–935. doi: 10.1021/acssynbio.6b00017

Madin, K., Sawasaki, T., Ogasawara, T., and Endo, Y. (2000). A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 559–564. doi: 10.1073/pnas.97.2.559

Maerkl, S. J., and Quake, S. R. (2007). A systems approach to measuring the binding energy landscapes of transcription factors. *Science* 315, 233–237. doi: 10.1126/science.1131007

Maerkl, S. J., and Quake, S. R. (2009). Experimental determination of the evolvability of a transcription factor. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18650–18655. doi: 10.1073/pnas.0907688106

Majumder, S., Garamella, J., Wang, Y.-L., DeNies, M., Noireaux, V., and Liu, A. P. (2017). Cell-sized mechanosensitive and biosensing compartment programmed with DNA. *Chem. Commun.* 53, 7349–7352. doi: 10.1039/C7CC03455E

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 Application Note, App-G124. San Jose, CA: Labcyte Inc.

Marshall, R., Maxwell, C. S., Collins, S. P., Beisel, C. L., and Noireaux, V. (2017). Short DNA containing χ sites enhances DNA stability and gene expression in *E. coli* cell-free transcription–translation systems. *Biotechnol. Bioeng.* 114, 2137–2141. doi: 10.1002/bit.26333

Martin, L. L., Meier, M. M., Lyons, S. M. S., Sit, R. V. R., Marzluff, W. F. W., Quake, S. R. S., et al. (2012). Systematic reconstruction of RNA functional motifs with high-throughput microfluidics. *Nat. Methods* 9, 1192–1194. doi: 10.1038/nmeth.2225

Martin, R. W., Soye, B. J. D., Kwon, Y.-C., Kay, J., Davis, R. G., Thomas, P. M., et al. (2018). Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat. Commun.* 9, 1–9. doi: 10.1038/s41467-018-03469-5

Martino, C., Kim, S.-H., Horsfall, L., Abbaspourrad, A., Rosser, S. J., Cooper, J., et al. (2012). Protein expression, aggregation, and triggered release from polymersomes as artificial cell-like structures. *Angew. Chem. Int. Ed.* 51, 6416–6420. doi: 10.1002/anie.201201443

Matsuura, T., Hosoda, K., Kazuta, Y., Ichihashi, N., Suzuki, H., and Yomo, T. (2012). Effects of compartment size on the kinetics of intracompartmental multimeric protein synthesis. *ACS Synth. Biol.* 1, 431–437. doi: 10.1021/sb300041z

Matsuura, T., Hosoda, K., and Shimizu, Y. (2018). Robustness of a reconstituted *Escherichia coli* protein translation system analyzed by computational modeling. *ACS Synth. Biol.* 7, 1964–1972. doi: 10.1021/acssynbio.8b00228

Matsuura, T., Kazuta, Y., Aita, T., Adachi, J., and Yomo, T. (2009). Quantifying epistatic interactions among the components constituting the protein translation system. *Mol. Syst. Biol.* 5:297. doi: 10.1038/msb.2009.50

Matsuura, T., Tanimura, N., Hosoda, K., Yomo, T., and Shimizu, Y. (2017). Reaction dynamics analysis of a reconstituted *Escherichia coli* protein translation system by computational modeling. *Proc. Natl. Acad. Sci. U.S.A.* 114, E1336–E1344. doi: 10.1073/pnas.1615351114

Mavelli, F., Marangoni, R., and Stano, P. (2015). A simple protein synthesis model for the PURE system operation. *Bull. Math. Biol.* 77, 1185–1212. doi: 10.1007/s11538-015-0082-8

Mei, Q., Fredrickson, C. K., Lian, W., Jin, S., and Fan, Z. H. (2006). Ricin detection by biological signal amplification in a well-in-a-well device. *Anal. Chem.* 78, 7659–7664. doi: 10.1021/ac0610006

Mei, Q., Fredrickson, C. K., Simon, A., Khnouf, R., and Fan, Z. H. (2007). Cell-free protein synthesis in microfluidic array devices. *Biotechnol. Prog.* 23, 1305–1311. doi: 10.1021/bp070133p

Menon, G., and Krishnan, J. (2019). Design principles for compartmentalization and spatial organization of synthetic genetic circuits. *ACS Synth. Biol.* 8, 1601–1619. doi: 10.1021/acssynbio.8b00522

Mills, D. R., Peterson, R. L., and Spiegelman, S. (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl. Acad. Sci. U.S.A.* 58, 217–224. doi: 10.1073/pnas.58.1.217

Moore, S. J., MacDonald, J. T., Wienecke, S., Ishwarbhai, A., Tsipa, A., Aw, R., 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. doi: 10.1073/pnas.1715806115

Moriizumi, Y., Tabata, K. V., Miyoshi, D., and Noji, H. (2019). Osmolyte-enhanced protein synthesis activity of a reconstituted translation system. *ACS Synth Biol.* 8, 557–567. doi: 10.1021/acssynbio.8b00513

Mullis, K. B., and Faloona, F. A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155, 335–350. doi: 10.1016/0076-6879(87)55023-6

Nathans, D., Notani, G., Schwartz, J. H., and Zinder, N. D. (1962). Biosynthesis of the coat protein of coliphage F2 by *E. coli* extracts. *Proc. Natl. Acad. Sci. U.S.A.* 48, 1424–1431. doi: 10.1073/pnas.48.8.1424

Niederholtmeyer, H., Chaggan, C., and Devaraj, N. K. (2018). Communication and quorum sensing in non-living mimics of eukaryotic cells. *Nat. Commun.* 9, 1–8. doi: 10.1038/s41467-018-07473-7

Niederholtmeyer, H., Sun, Z., Hori, Y., Yeung, E., Verpoorte, A., Murray, R. M., et al. (2015). Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife* 4:e09771. doi: 10.7554/eLife.09771

Niederholtmeyer, H., Xu, L., and Maerkl, S. J. (2013). Real-time mRNA measurement during an *in vitro* transcription and translation reaction using binary probes. *ACS Synth. Biol.* 2, 411–417. doi: 10.1021/sb300104f

Nirenberg, M., and Leder, P. (1964). RNA codewords and protein synthesis. *Science* 145, 1399–1407. doi: 10.1126/science.145.3639.1399

Nirenberg, M. W., Caskey, T., Marshall, R., Brimacombe, R., Kellog, D., Doctor, B., et al. (1966). The RNA code and protein synthesis. *Cold Spring Harbor Symp. Quant. Biol.* 31, 11–24. doi: 10.1101/SQB.1966.031.01.008

Nirenberg, M. W., and Matthaei, J. H. (1961). The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 47, 1588–1602. doi: 10.1073/pnas.47.10.1588

Nishimura, K., Matsuura, T., Nishimura, K., Sunami, T., Suzuki, H., and Yomo, T. (2012). Cell-free protein synthesis inside giant unilamellar vesicles analyzed by flow cytometry. *Langmuir* 28, 8426–8432. doi: 10.1021/la3001703

Niwa, T., Kanamori, T., Ueda, T., and Taguchi, H. (2012). Global analysis of chaperone effects using a reconstituted cell-free translation system. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8937–8942. doi: 10.1073/pnas.1201380109

Niwa, T., Sasaki, Y., Uemura, E., Nakamura, S., Akiyama, M., Ando, M., et al. (2015a). Comprehensive study of liposome-assisted synthesis of membrane proteins using a reconstituted cell-free translation system. *Sci. Rep.* 5:18025. doi: 10.1038/srep18025

Niwa, T., Sugimoto, R., Watanabe, L., Nakamura, S., Ueda, T., and Taguchi, H. (2015b). Large-scale analysis of macromolecular crowding effects on protein aggregation using a reconstituted cell-free translation system. *Front. Microbiol.* 6:1113. doi: 10.3389/fmicb.2015.01113

Noireaux, V., Bar-Ziv, R., and Libchaber, A. (2003). Principles of cell-free genetic circuit assembly. *Proc. Natl. Acad. Sci. U.S.A.* 100, 12672–12677. doi: 10.1073/pnas.2135496100

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

Nomura, S.-I. M., Tsumoto, K., Hamada, T., Akiyoshi, K., Nakatani, Y., and Yoshikawa, K. (2003). Gene expression within cell-sized lipid vesicles. *ChemBioChem* 4, 1172–1175. doi: 10.1002/cbic.200300630

Norred, S. E., Caveney, P. M., Chauhan, G., Collier, L. K., Collier, C. P., Abel, S. M., et al. (2018). Macromolecular crowding induces spatial correlations that control gene expression bursting patterns. *ACS Synth. Biol.* 7, 1251–1258. doi: 10.1021/acssynbio.8b00139

Nourian, Z., Roelofsen, W., and Danelon, C. (2012). Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angew. Chem. Int. Ed.* 51, 3114–3118. doi: 10.1002/anie.201107123

Oberholzer, T., and Luisi, P. L. (2002). The use of liposomes for constructing cell models. *J. Biol. Phys.* 28, 733–744. doi: 10.1023/A:1021267512805

Oberholzer, T., Nierhaus, K. H., and Luisi, P. L. (1999). Protein expression in liposomes. *Biochem. Biophys. Res. Commun.* 261, 238–241. doi: 10.1006/bbrc.1999.0404

Okano, T., Matsuura, T., Kazuta, Y., Suzuki, H., and Yomo, T. (2012). Cell-free protein synthesis from a single copy of DNA in a glass microchamber. *Lab Chip* 12, 2704–2708. doi: 10.1039/c2lc40098g

Okano, T., Matsuura, T., Suzuki, H., and Yomo, T. (2014). Cell-free protein synthesis in a microchamber revealed the presence of an optimum compartment volume for high-order reactions. *ACS Synth. Biol.* 3, 347–352. doi: 10.1021/sb400087e

Panthu, B., Ohlmann, T., Perrier, J., Schlattner, U., Jalinot, P., Elena-Herrmann, B., et al. (2018). Cell-free protein synthesis enhancement from real-time nmr metabolite kinetics: redirecting energy fluxes in hybrid RRL systems. *ACS Synth. Biol.* 7, 218–226. doi: 10.1021/acssynbio.7b00280

Park, N., Kahn, J. S., Rice, E. J., Hartman, M. R., Funabashi, H., Xu, J., et al. (2009a). High-yield cell-free protein production from P-gel. *Nat. Protoc.* 4, 1759–1770. doi: 10.1038/nprot.2009.174

Park, N., Um, S. H., Funabashi, H., Xu, J., and Luo, D. (2009b). A cell-free protein-producing gel. *Nat. Mater.* 8, 432–437. doi: 10.1038/nmat2419

Perez, J. G., Stark, J. C., and Jewett, M. C. (2016). Cell-free synthetic biology: engineering beyond the cell. *Cold Spring Harbor Perspect. Biol.* 8:a023853. doi: 10.1101/cshperspect.a023853

Phillips, R., Belliveau, N. M., Chure, G., Garcia, H. G., Razo-Mejia, M., and Scholes, C. (2019). Figure 1 theory meets Figure 2 experiments in the study of gene expression. *Annu. Rev. Biophys.* 48, 121–163. doi: 10.1146/annurev-biophys-052118-115525

Purnick, P. E., and Weiss, R. (2009). The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* 10, 410–422. doi: 10.1038/nrm2698

Qian, Y., Huang, H. H., Jiménez, J. I., and Del Vecchio, D. (2017). Resource competition shapes the response of genetic circuits. *ACS Synth. Biol.* 6, 1263–1272. doi: 10.1021/acssynbio.6b00361

Ramachandran, N., Hainsworth, E., Bhullar, B., Eisenstein, S., Rosen, B., Lau, A. Y., et al. (2004). Self-assembling protein microarrays. *Science* 305, 86–90. doi: 10.1126/science.1097639

Ramachandran, N., Raphael, J. V., Hainsworth, E., Demirkan, G., Fuentes, M. G., Rolfs, A., et al. (2008). Next-generation high-density self-assembling functional protein arrays. *Nat. Methods* 5, 535–538. doi: 10.1038/nmeth.1210

Rampioni, G., D'Angelo, F., Messina, M., Zennaro, A., Kuruma, Y., Tofani, D., et al. (2018). Synthetic cells produce a quorum sensing chemical signal perceived by: *Pseudomonas aeruginosa*. *Chem. Commun.* 54, 2090–2093. doi: 10.1039/C7CC09678J

Rhee, K. Y., Opel, M., Ito, E., Hung, S. P., Arfin, S. M., and Hatfield, G. W. (1999). Transcriptional coupling between the divergent promoters of a prototypic LysR-type regulatory system, the ilvYC operon of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14294–14299. doi: 10.1073/pnas.96.25.14294

Rockel, S., Geertz, M., Hens, K., Deplancke, B., and Maerkl, S. J. (2013). iSLIM: a comprehensive approach to mapping and characterizing gene regulatory networks. *Nucleic Acids Res.* 41:e52. doi: 10.1093/nar/gks1323

Ruiz, R. C. H., Kiatwuthinon, P., Kahn, J. S., Roh, Y. H., and Luo, D. (2012). Cell-free protein expression from DNA-based hydrogel (P-gel) droplets for scale-up production. *Ind. Biotechnol.* 8, 372–377. doi: 10.1089/ind.2012.0024

Saeki, D., Sugiura, S., Kanamori, T., Sato, S., and Ichikawa, S. (2014). Microcompartmentalized cell-free protein synthesis in semipermeable microcapsules composed of polyethylenimine-coated alginate. *J. Biosci. Bioeng.* 118, 199–204. doi: 10.1016/j.jbiosc.2014.01.014

Sakamoto, R., Noireaux, V., and Maeda, Y. T. (2018). Anomalous scaling of gene expression in confined cell-free reactions. *Sci. Rep.* 8:7364. doi: 10.1038/s41598-018-25532-3

Sakatani, Y., Ichihashi, N., Kazuta, Y., and Yomo, T. (2015). A transcription and translation-coupled DNA replication system using rolling-circle replication. *Sci. Rep.* 5:10404. doi: 10.1038/srep10404

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

Salas, M., Holguera, I., Redrejo-Rodríguez, M., and de Vega, M. (2016). DNA-binding proteins essential for protein-primed bacteriophage Φ29 DNA replication. *Front. Mol. Biosci.* 3:37. doi: 10.3389/fmolb.2016.00037

Schaerli, Y., Stein, V., Spiering, M. M., Benkovic, S. J., Abell, C., and Hollfelder, F. (2010). Isothermal DNA amplification using the T4 replisome: circular nicking endonuclease-dependent amplification and primase-based whole-genome amplification. *Nucleic Acids Res.* 38:e201. doi: 10.1093/nar/gkq795

Scherzinger, E., Haring, V., Lurz, R., and Otto, S. (1991). Plasmid RSF1010 DNA replication *in vitro* promoted by purified RSF1010 RepA, RepB and RepC proteins. *Nucleic Acids Res.* 19, 1203–1211. doi: 10.1093/nar/19.6.1203

Schreiber, A., Huber, M. C., and Schiller, S. M. (2019). Prebiotic protocell model based on dynamic protein membranes accommodating anabolic reactions. *Langmuir* 35, 9593–9610. doi: 10.1021/acs.langmuir.9b00445

Schwarz-Schilling, M., Aufinger, L., Mückl, A., and Simmel, F. C. (2016). Chemical communication between bacteria and cell-free gene expression systems within linear chains of emulsion droplets. *Integr. Biol.* 8, 564–570. doi: 10.1039/C5IB00301F

Schwarz-Schilling, M., Dupin, A., Chizzolini, F., Krishnan, S., Mansy, S. S., and Simmel, F. C. (2018). Optimized assembly of a multifunctional RNA-protein nanostructure in a cell-free gene expression system. *Nano Lett.* 18, 2650–2657. doi: 10.1021/acs.nanolett.8b00526

Schwille, P., Spatz, J., Landfester, K., Herminghaus, S., Sourjik, V., Erb, T. J., et al. (2018). MaxSynBio: avenues towards creating cells from the bottom up. *Angew. Chem. Int. Ed.* 57, 13382–13392. doi: 10.1002/anie.201802288

Selberg, J., Gomez, M., and Rolandi, M. (2018). The potential for convergence between synthetic biology and bioelectronics. *Cell Syst.* 7, 231–244. doi: 10.1016/j.cels.2018.08.007

Shearwin, K. E., Callen, B. P., and Egan, J. B. (2005). Transcriptional interference–a crash course. *Trends Genet.* 21, 339–345. doi: 10.1016/j.tig.2005.04.009

Shepherd, T. R., Du, L., Liljeruhm, J., Samudyata, Wang, J., Sjödin, M. O., et al. (2017). *De novo* design and synthesis of a 30-cistron translation-factor module. *Nucleic Acids Res.* 45, 10895–10905. doi: 10.1093/nar/gkx753

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., et al. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755. doi: 10.1038/90802

Shimizu, Y., Kanamori, T., and Ueda, T. (2005). Protein synthesis by pure translation systems. *Methods* 36, 299–304. doi: 10.1016/j.ymeth.2005.04.006

Shimizu, Y., and Ueda, T. (2010). PURE technology. *Methods Mol. Biol.* 607, 247–248. doi: 10.1007/978-1-60327-331-2_2

Shojaeian, M., Lehr, F.-X., Göringer, H. U., and Hardt, S. (2019). On-demand production of femtoliter drops in microchannels and their use as biological reaction compartments. *Anal. Chem.* 91, 3484–3491. doi: 10.1021/acs.analchem.8b05063

Shrestha, P., Holland, T. M., and Bundy, B. C. (2012). Streamlined extract preparation for *Escherichia coli*-based cell-free protein synthesis by sonication or bead vortex mixing. *Biotechniques* 53, 163–174. doi: 10.2144/0000113924

Shum, H., and Balazs, A. C. (2017). Synthetic quorum sensing in model microcapsule colonies. *Proc. Natl. Acad. Sci. U.S.A.* 114, 8475–8480. doi: 10.1073/pnas.1702288114

Siegal-Gaskins, D., Tuza, Z. A., Kim, J., Noireaux, V., and Murray, R. M. (2014). Gene circuit performance characterization in a cell-free 'breadboard'. *ACS Synth. Biol.* 3, 416–425. doi: 10.1021/sb400203p

Silverman, A. D., Karim, A. S., and Jewett, M. C. (2019a). Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* 21, 151–170. doi: 10.1038/s41576-019-0186-3

Silverman, A. D., Kelley-Loughnane, N., Lucks, J. B., and Jewett, M. C. (2019b). Deconstructing cell-free extract preparation for *in vitro* activation of transcriptional genetic circuitry. *ACS Synth. Biol.* 8, 403–414. doi: 10.1021/acssynbio.8b00430

Siuti, P., Retterer, S. T., and Doktycz, M. J. (2011). Continuous protein production in nanoporous, picolitre volume containers. *Lab Chip* 11, 3523–3529. doi: 10.1039/c1lc20462a

Sokolova, E., Spruijt, E., Hansen, M. M. K., Dubuc, E., Groen, J., Chokkalingam, V., et al. (2013). Enhanced transcription rates in membrane-free protocells formed by coacervation of cell lysate. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11692–11697. doi: 10.1073/pnas.1222321110

Spirin, A. S., Baranov, V. I., Ryabova, L. A., Ovodov, S. Y., and Alakhov, Y. B. (1988). A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* 242, 1162–1164. doi: 10.1126/science.3055301

Stano, P. (2019). Gene expression inside liposomes: from early studies to current protocols. *Chem. A Eur. J.* 25, 7798–7814. doi: 10.1002/chem.201806445

Stano, P., D'Aguanno, E., Bolz, J., Fahr, A., and Luisi, P. L. (2013). A remarkable self-organization process as the origin of primitive functional cells. *Angew. Chem. Int. Ed.* 52, 13397–13400. doi: 10.1002/anie.201306613

Stapleton, J. A., and Swartz, J. R. (2010). Development of an *in vitro* compartmentalization screen for high-throughput directed evolution of [FeFe] hydrogenases. *PLoS ONE* 5:e15275. doi: 10.1371/journal.pone.0015275

Steinberg-Yfrach, G., Rigaud, J. L., Durantini, E. N., Moore, A. L., Gust, D., and Moore, T. A. (1998). Light-driven production of ATP catalysed by F0F1-ATP synthase in an artificial photosynthetic membrane. *Nature* 392, 479–482. doi: 10.1038/33116

Stillman, B. W., and Gluzman, Y. (1985). Replication and supercoiling of simian virus 40 DNA in cell extracts from human cells. *Mol. Cell. Biol.* 5, 2051–2060. doi: 10.1128/MCB.5.8.2051

Su'etsugu, M., Takada, H., Katayama, T., and Tsujimoto, H. (2017). Exponential propagation of large circular DNA by reconstitution of a chromosome-replication cycle. *Nucleic Acids Res.* 45, 11525–11534. doi: 10.1093/nar/gkx822

Sun, Z. Z., Hayes, C. A., Shin, J., Caschera, F., Murray, R. M., and Noireaux, V. (2013). Protocols for implementing an *Escherichia coli* based TX-TL cell-free expression system for synthetic biology. *J. Vis. Exp.* 79:e50762. doi: 10.3791/50762

Sun, Z. Z., Yeung, E., Hayes, C. A., Noireaux, V., and Murray, R. M. (2014). Linear DNA for rapid prototyping of synthetic biological circuits in an *Escherichia coli* based TX-TL cell-free system. *ACS Synth. Biol.* 3, 387–397. doi: 10.1021/sb400131a

Sunami, T., Hosoda, K., Suzuki, H., Matsuura, T., and Yomo, T. (2010). Cellular compartment model for exploring the effect of the lipidic membrane on the kinetics of encapsulated biochemical reactions. *Langmuir* 26, 8544–8551. doi: 10.1021/la904569m

Swaminathan, A., Hsiao, V., and Murray, R. M. (2017). Quantitative modeling of integrase dynamics using a novel python toolbox for parameter inference in synthetic biology. *bioRxiv* 121152. doi: 10.1101/121152

Swank, Z., Laohakunakorn, N., and Maerkl, S. J. (2019). Cell-free gene-regulatory network engineering with synthetic transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* 116, 5892–5901. doi: 10.1073/pnas.1816591116

Takahashi, M. K., Chappell, J., Hayes, C. A., Sun, Z. Z., Kim, J., Singhal, V., et al. (2015a). Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription–translation (TX-TL) systems. *ACS Synth. Biol.* 4, 503–515. doi: 10.1021/sb400206c

Takahashi, M. K., Hayes, C. A., Chappell, J., Sun, Z. Z., Murray, R. M., Noireaux, V., et al. (2015b). Characterizing and prototyping genetic networks with cell-free transcription-translation reactions. *Methods* 86, 60–72. doi: 10.1016/j.ymeth.2015.05.020

Tang, T. Y. D., van Swaay, D., deMello, A., Anderson, J. L. R., and Mann, S. (2015). *In vitro* gene expression within membrane-free coacervate protocells. *Chem. Commun.* 51, 11429–11432. doi: 10.1039/C5CC04220H

Tawfik, D. S., and Griffiths, A. D. (1998). Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* 16, 652–656. doi: 10.1038/nbt0798-652

Tayar, A. M., Karzbrun, E., Noireaux, V., and Bar-Ziv, R. H. (2015). Propagating gene expression fronts in a one-dimensional coupled system of artificial cells. *Nat. Phys.* 11, 1037–1041. doi: 10.1038/nphys3469

Tayar, A. M., Karzbrun, E., Noireaux, V., and Bar-Ziv, R. H. (2017). Synchrony and pattern formation of coupled genetic oscillators on a chip of artificial cells. *Proc. Natl. Acad. Sci. U.S.A.* 114, 11609–11614. doi: 10.1073/pnas.1710620114

Thiele, J., Ma, Y., Foschepoth, D., Hansen, M. M. K., Steffen, C., Heus, H. A., et al. (2014). DNA-functionalized hydrogels for confined membrane-free *in vitro* transcription/translation. *Lab Chip* 14, 2651–2656. doi: 10.1039/c3lc51427g

Torre, P., Keating, C. D., and Mansy, S. S. (2014). Multiphase water-in-oil emulsion droplets for cell-free transcription–translation. *Langmuir* 30, 5695–5699. doi: 10.1021/la404146g

Tuckey, C., Asahara, H., Zhou, Y., and Chong, S. (2014). Protein synthesis using a reconstituted cell-free system. *Curr. Protoc. Mol. Biol.* 108, 16.31.1–16.31.22. doi: 10.1002/0471142727.mb1631s108

van der Linden, A. J., Yelleswarapu, M., Pieters, P. A., Swank, Z., Huck, W. T. S., Maerkl, S. J., et al. (2019). A multilayer microfluidic platform for the conduction of prolonged cell-free gene expression. *J. Vis. Exp.* 152, 1–14. doi: 10.3791/59655

van Nies, P., Westerlaken, I., Blanken, D., Salas, M., Mencía, M., and Danelon, C. (2018). Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun.* 9:1583. doi: 10.1038/s41467-018-03926-1

van Swaay, D., Tang, T. Y. D., Mann, S., and de Mello, A. (2015). Microfluidic formation of membrane-free aqueous coacervate droplets in water. *Angew. Chem. Int. Ed.* 127, 8518–8521. doi: 10.1002/ange.201502886

Vecchio, D. D., Dy, A. J., and Qian, Y. (2016). Control theory meets system biology. *J. R. Soc. Interface* 13:20160380. doi: 10.1098/rsif.2016.0380

Villar, G., Graham, A. D., and Bayley, H. (2013). A tissue-like printed material. *Science* 340, 48–52. doi: 10.1126/science.1229495

Villarreal, F., Contreras-Llano, L. E., Chavez, M., Ding, Y., Fan, J., Pan, T., et al. (2018). Synthetic microbial consortia enable rapid assembly of pure translation machinery. *Nat. Chem. Biol.* 14, 29–35. doi: 10.1038/nchembio.2514

Vogele, K., Frank, T., Gasser, L., Goetzfried, M. A., Hackl, M. W., Sieber, S. A., et al. (2018). Towards synthetic cells using peptide-based reaction compartments. *Nat. Commun.* 9, 1–7. doi: 10.1038/s41467-018-06379-8

Waga, S., Bauer, G., and Stillman, B. (1994). Reconstitution of complete SV40 DNA replication with purified replication factor. *J. Biol. Chem.* 3, 1907–1941.

Wang, H. H., Huang, P.-Y., Xu, G., Haas, W., Marblestone, A., Li, J., et al. (2012). Multiplexed *in vivo* his-tagging of enzyme pathways for *in vitro* single-pot multienzyme catalysis. *ACS Synth. Biol.* 1, 43–52. doi: 10.1021/sb3000029

Wang, P., Fujishima, K., Berhanu, S., Kuruma, Y., Jia, T. Z., Khusnutdinova, A. N., et al. (2019). A single polyphosphate kinase-based NTP regeneration system driving cell-free protein synthesis. *chemRxiv [preprint]*. doi: 10.26434/chemrxiv.8874410.v1

Wang, S., Majumder, S., Emery, N. J., and Liu, A. P. (2018). Simultaneous monitoring of transcription and translation in mammalian cell-free expression in bulk and in cell-sized droplets. *Synth. Biol.* 3:ysy005. doi: 10.1093/synbio/ysy005

Weitz, M., Kim, J., Kapsner, K., Winfree, E., Franco, E., and Simmel, F. C. (2014). Diversity in the dynamical behaviour of a compartmentalized programmable biochemical oscillator. *Nat. Chem.* 6, 295–302. doi: 10.1038/nchem.1869

Wobbe, C. R., Dean, F., Weissbach, L., and Hurwitz, J. (1985). *In vitro* replication of duplex circular DNA containing the simian virus 40 DNA origin site. *Proc. Natl. Acad. Sci. U.S.A.* 82, 5710–5714. doi: 10.1073/pnas.82.17.5710

Wood, W. B., and Berg, P. (1962). The effect of enzymatically synthesized ribonucleic acid on amino acid incorporation by a soluble protein-ribosome system from *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 48, 94–104. doi: 10.1073/pnas.48.1.94

Woronoff, G., Ryckelynck, M., Wessel, J., Schicke, O., Griffiths, A. D., and Soumillion, P. (2015). Activity-fed translation (AFT) assay: a new high-throughput screening strategy for enzymes in droplets. *ChemBioChem* 16, 1343–1349. doi: 10.1002/cbic.201500087

Yang, D., Peng, S., Hartman, M. R., Gupton-Campolongo, T., Rice, E. J., Chang, A. K., et al. (2013). Enhanced transcription and translation in clay hydrogel and implications for early life evolution. *Sci. Rep.* 3, 27–26. doi: 10.1038/srep03165

Yelleswarapu, M., van der Linden, A. J., van Sluijs, B., Pieters, P. A., Dubuc, E., de Greef, T. F. A., et al. (2018). Sigma factor-mediated tuning of bacterial cell-free synthetic genetic oscillators, supplementary information. *ACS Synth. Biol.* 7:2879. doi: 10.1021/acssynbio.8b00300

Yeung, E., Dy, A. J., Martin, K. B., Ng, A. H., Del Vecchio, D., Beck, J. L., et al. (2017). Biophysical constraints arising from compositional context in synthetic gene networks. *Cell Syst.* 5, 11–24.e12. doi: 10.1016/j.cels.2017.06.001

Yim, S. S., Johns, N. I., Park, J., Gomes, A. L., McBee, R. M., Richardson, M., et al. (2019). Multiplex transcriptional characterizations across diverse bacterial species using cell-free systems. *Mol. Syst. Biol.* 15:e8875. doi: 10.15252/msb.20198875

Yu, W., Sato, K., Wakabayashi, M., Nakaishi, T., Ko-Mitamura, E. P., Shima, Y., et al. (2001). Synthesis of functional protein in liposome. *J. Biosci. Bioeng.* 92, 590–593. doi: 10.1016/S1389-1723(01)80322-4

Zemella, A., Thoring, L., Hoffmeister, C., and Kubick, S. (2015). Cell-free protein synthesis: pros and cons of prokaryotic and eukaryotic systems. *ChemBioChem* 16, 2420–2431. doi: 10.1002/cbic.201500340

Zhou, X., Wu, H., Cui, M., Lai, S. N., and Zheng, B. (2018). Long-lived protein expression in hydrogel particles: towards artificial cells. *Chem. Sci.* 9, 4275–4279. doi: 10.1039/C8SC00383A

Zhou, Y., Asahara, H., Schneider, N., Dranchak, P., Inglese, J., and Chong, S. (2014). Engineering bacterial transcription regulation to create a synthetic *in vitro* two-hybrid system for protein interaction assays. *J. Am. Chem. Soc.* 136, 14031–14038. doi: 10.1021/ja502512g

Zubay, G. (1973). *In vitro* synthesis of protein in microbial systems. *Annu. Rev. Genet.* 7, 267–287. doi: 10.1146/annurev.ge.07.120173.001411

Zubay, G., Lederman, M., and DeVries, J. K. (1967). DNA-directed peptide synthesis, III. Repression of beta-galactosidase synthesis and inhibition of repressor by inducer in a cell-free system. *Proc. Natl. Acad. Sci. U.S.A.* 58, 1669–1675. doi: 10.1073/pnas.58.4.1669

Zubay, G., Schwartz, D., and Beckwith, J. (1970). Mechanism of activation of catabolite-sensitive genes: a positive control system. *Proc. Natl. Acad. Sci. U.S.A.* 66, 104–110. doi: 10.1073/pnas.66.1.104

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

ORIGINAL RESEARCH
published: 09 April 2020
doi: 10.3389/fbioe.2020.00312

O_2-Tuned Protein Synthesis Machinery in *Escherichia coli*-Based Cell-Free System

Xiaomei Lin[1†], Caijin Zhou[2†], Songbiao Zhu[3], Haiteng Deng[3], Jisong Zhang[2] and Yuan Lu[1*]*

[1] Key Laboratory of Industrial Biocatalysis, Ministry of Education, Department of Chemical Engineering, Tsinghua University, Beijing, China, [2] The State Key Lab of Chemical Engineering, Department of Chemical Engineering, Tsinghua University, Beijing, China, [3] Ministry of Education Key Laboratory of Bioinformatics, Center for Synthetic and Systematic Biology, School of Life Sciences, Tsinghua University, Beijing, China

OPEN ACCESS

Edited by:
Shihui Yang,
Hubei University, China

Reviewed by:
Zhiguang Zhu,
Tianjin Institute of Industrial Biotechnology (CAS), China
Chun You,
Tianjin Institute of Industrial Biotechnology (CAS), China

***Correspondence:**
Jisong Zhang
jiszhang@tsinghua.edu.cn
Yuan Lu
yuanlu@tsinghua.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: 10 February 2020
Accepted: 23 March 2020
Published: 09 April 2020

Citation:
Lin X, Zhou C, Zhu S, Deng H, Zhang J and Lu Y (2020) O_2-Tuned Protein Synthesis Machinery in Escherichia coli-Based Cell-Free System. Front. Bioeng. Biotechnol. 8:312. doi: 10.3389/fbioe.2020.00312

Involved in most aerobic biochemical processes, oxygen affects cellular functions, and organism behaviors. Protein synthesis, as the underlying biological process, is unavoidably affected by the regulation of oxygen delivery and utilization. Bypassing the cell wall, cell-free protein synthesis (CFPS) systems are well adopted for the precise oxygen regulation analysis of bioprocesses. Here a reliable flow platform was developed for measuring and analyzing the oxygen regulation on the protein synthesis processes by combining *Escherichia coli*-based CFPS systems and a tube-in-tube reactor. This platform allows protein synthesis reactions conducted in precisely controlled oxygen concentrations. For analysis of the intrinsic role of oxygen in protein synthesis, O_2-tuned CFPS systems were explored with transcription-translation related parameters (transcripts, energy, reactive oxygen species, and proteomic pathway analysis). It was found that 2% of oxygen was the minimum requirement for protein synthesis. There was translation-related protein degradation in the high oxygen condition leading to a reduction. By combining the precise gas level controlling and open biosystems, this platform is also potential for fundamental understanding and clinical applications by diverse gas regulation in biological processes.

Keywords: cell-free system, tube-in-tube reactor, oxygen-tuned protein synthesis, ATP analysis, reactive oxygen species (ROS), proteomic analysis

INTRODUCTION

Oxygen, serving as essential nutrition of living life, is the critical substrate of many physiological processes in cells, such as ATP generation in aerobic metabolism, which is considered as the energy supply of thousands of biochemical reactions as well as many oxidative reactions (Lindahl, 2008; Heffner et al., 2013; Palmer and Clegg, 2014). Each life is well adopted to a particular oxygen concentration so that changes in oxygen concentration will cause a series of complicated adaptive efforts or misfunction on cells, tissues, and organisms (Ivan et al., 2001; Jaakkola et al., 2001). Adaptions are including but not limited to enzyme activity, membrane transport (Moore et al., 2018), proliferation (Carmeliet and Jain, 2000; Ke and Costa, 2006; Mohyeldin et al., 2010; Heffner et al., 2013), and systemic inflammatory response. After a long period of evolution, the atmospheric conditions were transferred into the present stable element's composition,

the 21% oxygen concentration (Heffner et al., 2013). However, the natural atmospheric condition does not actually reflect the oxygen microenvironments typically in cellular pathways, and many physiological reactions are conducted in low-level oxygen tension, generally in 2–9% (Brahimi-Horn and Pouysségur, 2007; Mohyeldin et al., 2010).

Although oxygen is vital for life, the mechanisms of oxygen sensing and oxygen utilization in cells are still not completely understood, for the limitation of precise control on gradient oxygen concentrations and effective methods to quantify oxygen concentration in its physiological microenvironment (Roussakis et al., 2015). For gas controlling, pioneering efforts in microfluidic devices have demonstrated that it is possible to conduct oxygen concentrations in tiny gradients by physical diffusion equilibrium, biological cellular uptake, and chemical generation for the study of bacterium growth, cell differentiation, migration and disease mechanism (Leclerc et al., 2004; Lam et al., 2009; Adler et al., 2010; Abaci et al., 2012). Moreover, the ineffective and destructive oxygen concentration quantifying methods still limit the fundamental physiological oxygen-derived research (Roussakis et al., 2015). Clark-style electrodes (Clark et al., 1953; Nozue et al., 1997; Buerk, 2004) and luminescent optical sensors (Sud et al., 2006; Wang and Hu, 2012) are commonly used for oxygen measurement in biological process research, which is either focusing on the extracellular measurement or intracellular dye toxicity (Brennan et al., 2014). Therefore, constructing a more flexible gas controlling and measuring systems are critical challenges for the research of essential oxygen affecting cellular processes.

Cell-free synthetic biology, an emerging biological tool for the study of biological reactions without using the intact cells, has been widely used for the fundamental research in biomolecules, pathways, and whole-cell interactions (Perez et al., 2016). With the advantage of bypassing cell membranes, cell-free systems provide unprecedented freedom for applications of precise physiological analysis researches (Shin and Noireaux, 2012; Lu, 2017; Marshall et al., 2018). Therefore, cell-free systems are well-suitable for oxygen-tuned biological change analysis that could directly reflect the intracellular response to the microenvironment changes. Recently, the tube-in-tube reactor, as a powerful tool for studying the gas-liquid system, has shown significant advantages, which was widely used in gas-liquid reactions in flow mode (Koos et al., 2011; O'brien et al., 2011; Polyzos et al., 2011; Brzozowski et al., 2015). As the device's name suggests, the tube-in-tube reactor consists of a semipermeable inner tubing and a gas non-permeable outer tubing (Yang and Jensen, 2013; Zhang et al., 2017). Due to the high gas permeability of the inner tubing, the tube-in-tube reactor has a rapid oxygen diffusion rate, which allows oxygen to penetrate rapidly into the liquid in about 30 s (Zhang et al., 2017, 2019). In particular, by using the tube-in-tube reactor, it can easily realize the accurate control and alteration of oxygen concentration in liquid by changing the volume fraction of oxygen at the oxygen-mediated reaction process (Petersen et al., 2012; Ringborg et al., 2017; Burkholder et al., 2019).

Here, a new robust method for oxygen-tuned biological regulation was presented by combining the tube-in-tube flow reactor technology and cell-free systems, in which a physiological process in a precise oxygen concentration was conducted. By exposing cell-free protein synthesis (CFPS) systems to 0–100% oxygen gradients, the features of oxygen regulation in fluorescent protein sfGFP (super fold green fluorescent protein) synthesis were analyzed among energy, transcription, translation, and pathway networks.

MATERIALS AND METHODS

Tube-in-Tube Device

The tube-in-tube reactor was composed of an inner semipermeable membrane tube and an outer impermeable tube. The inner tube was fabricated by the materials of Teflon AF-2400 with an inner diameter of 0.6 mm, and the outer tube was made by polytetrafluoroethylene (PTFE) with an inner diameter of 3.175 mm. Due to the high permeability and short mass transfer distance (0.6 mm) of the inner tube, oxygen could penetrate rapidly into the reaction substrates and saturated the reaction substrates in 30 s (Zhang et al., 2017). A total of 300 μL cell-free reactants were injected into the inner tube through a Harvard pump (Harvard, United States). Twenty microliters of reactants were pushed out and gathered for further analysis at every 20 min in 2 h (**Supplementary Figures S1, S2**). More details were shown in the **Supplementary Material**.

Cell Extract Preparation

For standard cell-free expression reactions, *E. coli* Rosetta (DE3) were fermented in 4 L of 2xYTP (1.6% tryptone; 1% yeast extraction; 0.5% NaCl; 40 mM K_2HPO_4; 22 mM KH_2PO_4) containing chloramphenicol (34 μg/mL) at 37°C, 300 rpm. Cells were harvested in the late logarithmic growth phase (∼4 h, OD600 = 2). Cell pellets were washed with ice-cold S30A buffer (14 mM magnesium glutamate; 60 mM potassium glutamate; 50 mM Tris, pH 7.7) three times. Cells were resuspended in S30A buffer (1 mL buffer for 1 g of wet cells) for disruption in a high-pressure homogenizer (1,000 bar). Then lysate was centrifuged at 12,000 × g for 10 min at 4°C. The supernatant was incubated at 37°C for 80 min after 3 mM DTT added and centrifuged at 12,000 × g for 10 min at 4°C again. Then it was dialyzed in molecular porous membrane tubing (6-8 KD MWCO) for 3h at 4°C with magnetic stirring. The dialysate was then centrifuged at 12,000 × g for 10 min at 4°C, flash frozen and stored at −80°C (Sun et al., 2013; **Supplementary Figure S3**).

Cell-Free Reaction

Cell-free reagents for tube-in-tube reactions were assembled on ice as generally described and immediately injected into the inner tube for incubation at 37°C. The general cell-free reaction mixture consisted of the following components: 30% S30 cell extract (v/v%); 15–20 ng/μL DNA templates; 175 mM potassium glutamate; 10 mM ammonium glutamate; 2.7 mM potassium oxalate monohydrate; 10 mM magnesium glutamate; 50 mM each of 19 amino acids without glutamic acid; 3 mM phosphenol pyruvate (PEP); 1 mM putrescine; 1.5 mM spermidine; 0.33 mM nicotinamide adenine dinucleotide

(NAD); 1.2 mM ATP; 0.86 mM each of CTP, GTP and UTP; 0.27 mM coenzyme A; 170 μg/mL tRNA; 34 μg/mL folinic; T7 RNA polymerase prepared from *E. coli* BL21 (DE3) cell extract; reactions also contained 2% PEG8000 (Sun et al., 2013; **Supplementary Figure S4**).

Fluorescence Measurement and Data Processing

For fluorescent protein sfGFP measurement, the cell-free samples were diluted and measured in the plate reader (TECAN infinite M200 Pro) with the wavelength of excitation at 485 nm and emission at 520 nm. To minimize the error caused by different batches of cell extracts or other reagents, the fluorescence values were normalized.

Since the long-term storage of cell-free reagents would lead to the fluctuation in fluorescence production, one of the time points (120 min) of 10% treatment was used for the normalization of each cell-free reactions from the same batch of cell-free reagents. The fluorescence of 120 min time point of 10% (F_{106}) was considered as the same value at each experiment, that the other fluorescence (F_X) was calculated as follow: $F_{XN} = F_X/F_{106}{}^{*}0.825$ (0.825, the relative fluorescence to one of 21% treatment in the time point of 150 min). Experiments were run more than twice.

The ROS Level Analysis

The tube-in-tube cell-free reaction samples were determined by Reactive Oxygen Species Assay Kit (Beyotime, S0033) through the DCFH-DA ROS probe. Cell-free reactions without sfGFP expression were injected into the tube-in-tube reactor along with the DCFH-DA probe for incubation directly. At predetermined time intervals, samples were removed onto the ice. Measure the fluorescence intensity in the plate reader (TECAN infinite M200 Pro) with the wavelength of excitation at 485 nm and emission at 520 nm.

The ATP Level Analysis

The ATP concentrations were determined by ATP Assay Kit (Beyotime, S0026) through firefly luciferase. Cell-free samples were diluted into lysis buffer and added to the diluted working solution. Luminescence was determined by the plate reader (TECAN infinite M200 Pro). Series concentrations of ATP were used for a standard curve to calculate the ATP concentration of samples.

RNA Extraction and Target mRNA Quantification

The total mRNA was extracted from cell-free samples by Eastep Super Total RNA Extraction Kit. Then it was stored at −80°C, or the reverse transcription was performed immediately by FastKing RT Kit (With gDNase) (TIANGEN, KR116). The cDNAs were quantified by TranStart Green qPCR SuperMix kit (TranStart, AQ101). The PCR product of sfGFP amplified by the same primers as qPCR was used for the standard curve to relate the sfGFP cDNA concentrations with CT values read by ABI 7300 Real-Time PCR system.

HPLC-MS/MS Analysis

The samples were separated by SDS-PAGE (12% separating gel) for further MS analysis. The proteins were analyzed by Center of Biomedical Analysis, Tsinghua University. The peptide samples were injected to Q-Exactive mass spectrometor for analysis after In-gel digestion process. The raw files were processed by using the MaxQuant software version 1.6.0.16 (Max Planck Institute, Munich, Germany). The *Escherichia coli* database (strain K12, UniProt release 17/03/2019; containing 29,810 entries) was used by the search engine andromeda implemented in MaxQuant. The criteria for the searching process were listed below: full tryptic specificity required, two missed cleavage tolerance, 20 ppm and 0.02 Da for precursor and fragment ion mass tolerance, respectively. The variable modifications were oxidation on methionine and acetylation on protein N-terminal, and the fixed modifications were carbamidomethylation on cysteine. Both peptide spectrum matches (PSMs) and protein identifications were filtered based on 1% false discovery rate (FDR) based on decoy search for revert mode. The relative label-free quantification (LFQ) of proteomics data based on precursor (MS1) signal intensities were used for quantifying proteins across biological samples. The algorithm MaxLFQ implemented in MaxQuant software was used for LFQ calculation.

One-way ANOVA method and Turkey *post-hoc* analysis were used to determine the significant difference among groups with different oxygen concentrations. *P*-value below 0.05 was considered to be significant. Gene ontology (GO) enrichment analysis was performed using DAVID tools (version 6.8, The Database for Annotation, Visualization and Integrated Discovery) with the complete *Escherichia coli* genome information as the background (Jiao et al., 2012). Visualization of results, including the GO annotations and volcano, were conducted with R packages GOplot, ggplot2, in RStudio (version 1.1, base R 3.5).

RESULTS AND DISCUSSION

Design of Oxygen Gradient Generator

The fabrication of the tube-in-tube reactor was constructed with a semipermeable fluoropolymer inner tube (Teflon AF-2400) encased by a gas-non-permeable outer tube (**Figure 1A**). The Teflon AF-2400 inner tubing is highly permeable for gas but non-permeable for liquid with an inner diameter of 0.6 mm and an outer diameter of 0.8 mm. The structure of the reactor has been introduced in detail previously (Yang and Jensen, 2013; Zhang et al., 2017, 2019; Zhou et al., 2020). The cell-free reactions were injected into the inner tube with a Harvard pump (Harvard, United States). The impermeable outer tube was made of PTFE with an inner diameter of 1.6 mm and an outer diameter of 3.175 mm. Different tensions of oxygen were created by an oxygen stream and a nitrogen stream whose flow rates were controlled by two thermal mass flow controllers (Sevenstar, Beijing, China), respectively (**Supplementary Figure S1**).

Because of the rapid oxygen diffusion rate (30 s) from the gas phase into the reaction substances in the tube-in-tube reactor, the

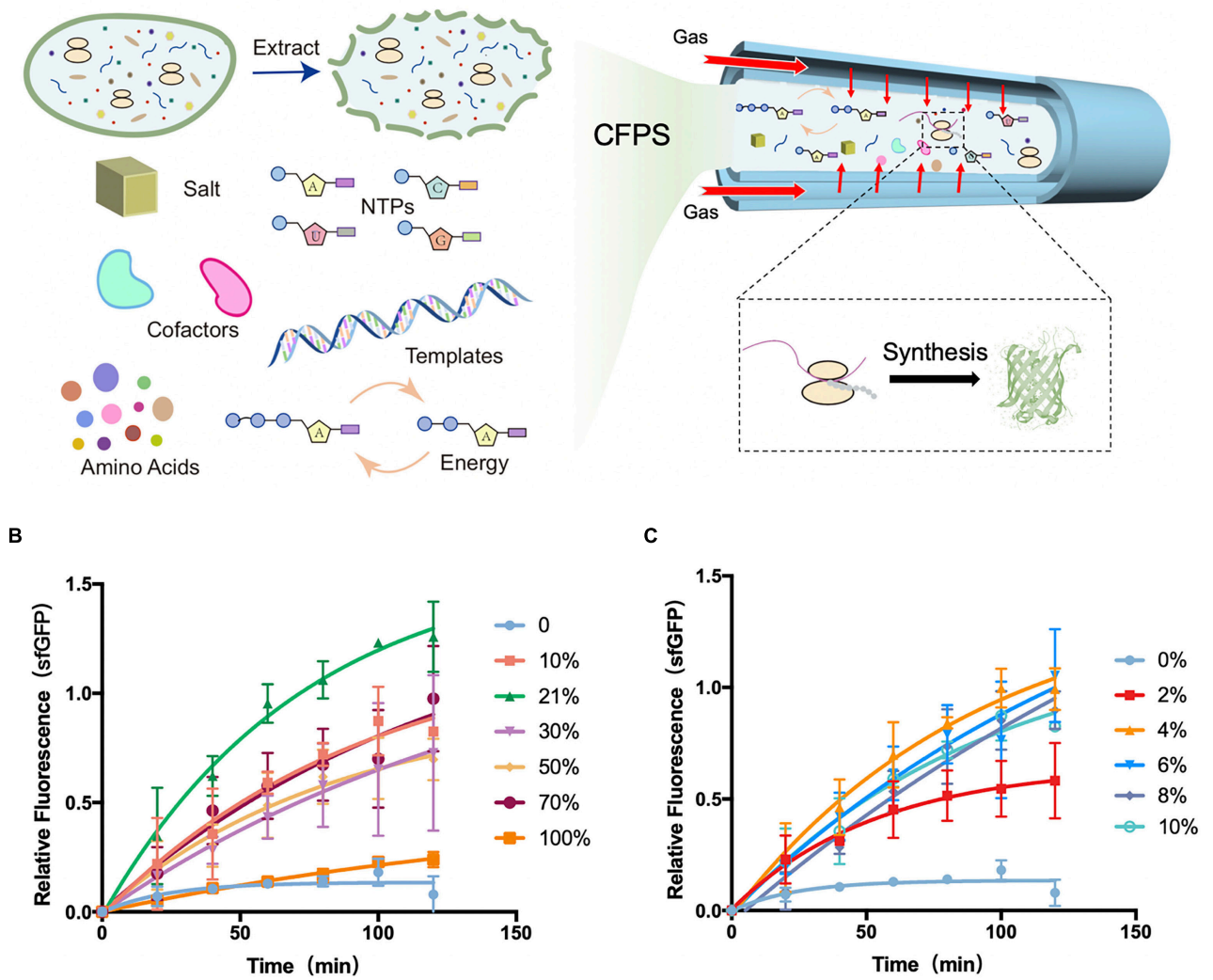


FIGURE 1 | Schematic diagram of tube-in-tube devices for cell-free protein synthesis and fluorescence gained from devices. **(A)** Cell-free protein synthesis (CFPS) system was injected into the inner tube of the tube-in-tube reactor, surrounded by the controllable gas. **(B)** Normalized fluorescence of sfGFP synthesized in cell-free systems in the tube-in-tube reactor oxygen tensions of 0, 10, 21, 30, 50, 70, and 100% in 2 h and 37°C. **(C)** Normalized fluorescence of sfGFP synthesized in cell-free systems in the tube-in-tube reactor with oxygen tensions of 0, 2, 4, 6, 8, and 10% in 2 h and 37°C. Values were calculated as the mean of biological replicates.

reaction substrates in the inner tube were always fully saturated with oxygen throughout the protein synthesis process (several hours). For calculating the saturated oxygen concentration of reaction substrates, Henry's law was used as following:

$$C^* = Ky$$

Where C^* (mg/L) is the saturated oxygen concentration of reaction substrates, y is the gas volume fraction of oxygen in mixture gas, and K [mg/(L)] is Henry's constant with a constant value of 34.4505 mg/(L) at 37°C. With this formula, the oxygen concentrations in cell-free liquid could be calculated, and the saturation concentrations of oxygen at different oxygen tensions were shown in **Supplementary Figure S5**.

Cell-Free Protein Synthesis in Precise Gradients of Oxygen

For the global analysis of the influence of oxygen concentration in CFPS, CFPS reactions for sfGFP (**Supplementary Figure S6**) in different levels of oxygen were firstly conducted in the tube-in-tube reactor. In this analysis platform, cell-free reaction systems were injected into the inner tube of the tube-in-tube reactor for protein synthesis. The gas tensions varied from 0 to 100% were continuously delivered by changing the flow rates of nitrogen and oxygen streams and maintaining the overall flow rate at 30 mL/min. Before performing the protein synthesis in the tube-in-tube reactor, the cell-free reactants were placed on the ice to stop the synthesis,

ensuring that the reactions were performed at a certain oxygen condition.

The protein synthesis features were first explored at a range of O_2 tension varied from 0 to 100% to investigate the effect of oxygen concentration in CFPS. The model *E. coli*-based cell-free system was employed for this investigation. All of the cell-free reactions for oxygen concentration investigation were performed at 37°C as the natural growth environment of *E. coli*. Experiments with oxygen tensions of 0, 10, 21, 30, 50, 70, and 100% were carried out to separately analyze the protein synthesis characteristics (**Figure 1B** and **Supplementary Figure S7A**). To minimize the error from the reagents, the fluorescence was normalized among each batch experiment. Partial fluorescence data were shown in the **Supplementary Material** (**Supplementary Figure S8**). As shown in **Figure 1B**, the natural atmosphere condition 21% oxygen showed the highest yields and protein synthesis rate among the conditions. With the increasing of the oxygen concentration, the protein synthesis rate increased obviously between the tension of 10 and 70%, but a drastic decline in the 100%. As expected, the synthesis was greatly limited when there was no oxygen (0% oxygen and 100% nitrogen) delivery into the synthesis reactions, suggesting that the oxygen was an essential substrate or intermediate in the protein synthesis. As the fluorescence of green fluorescent protein depended on the oxygen oxidation, the expression of 0% oxygen treatment was confirmed with fluorescent recovery (**Supplementary Figure S9**). As schematically illustrated in **Figure 1B**, although the synthesis was greatly limited without oxygen, certain low oxygen (10%) was required for the completely functional protein synthesis. However, the high-level oxygen showed the inhibition of protein synthesis in the cell-free system.

Based on the above investigation, the exact oxygen (0–10%) were further investigated for protein synthesis to seek the minimum oxygen concentration. Experiments were performed at the same conditions except for the gradient of oxygen tension, which was ranging from 0 to 10% (**Figure 1C** and **Supplementary Figure S7B**). As before, there was little protein synthesis in 0% oxygen condition, while an increasing protein yield appeared at 2%. Until at 4% oxygen, the protein yields were recovered to the high level, suggesting that only a small amount of oxygen (higher than 2%) was required for protein synthesis in *E. coli*-based cell-free systems. This phenomenon of low oxygen concentration requirement of protein synthesis is consistent with previous literature focusing on the *in vivo* systems. Earlier study had found that as little as 0.5% of oxygen could lead to a slow but significant recovered cell growth with protein synthesis in *E. coli*, and a full recovery was at the oxygen tension of ∼7%. However, the condition of 0% oxygen showed a stopping division and protein synthesis (Polinkovsky et al., 2009). From the above investigation and discussion, the tube-in-tube reactor with the excellent gas controlling feature showed great potential for the fundamental investigation of the effect of oxygen concentration on biological processes in the microscale environment with the cell-free system. Moreover, this flexible CFPS device potentially can be applied to different protein synthesis requirements for tunable synthesis rate regulated by the oxygen concentration.

In general, there were several essential observations in the above section. (i) At the 0% oxygen condition, the protein synthesis was limited. (ii) A small amount (larger than 2%) of oxygen was sufficient for the complete function of protein synthesis, and there was a maximum yield in 21% oxygen. (iii) The yield of protein synthesis was not highly dependent on the oxygen at a range of 4–70%, but a substantial decline at 100%. These findings pose a question of what oxygen affects the process of protein synthesis. There might be some changes in the oxygen metabolic pathway or the direct effects on the transcriptional and translational processes.

Transcription in O_2-Tuned Cell-Free Systems

To gain a more systematic insight into the effect of oxygen concentration on CFPS processes, the transcriptional level was next quantitatively characterized among O_2-tuned cell-free systems. As *in vivo* systems, the transcription in cell-free systems was a multienzyme reaction that synthesizes RNA from DNA templates for the later translation to proteins (**Figure 2A**). According to the protein expression observations discovered above, the transcriptional features of experiments on 0, 21, and 100% oxygen treatments were explored. The total RNA was extracted from cell-free systems at varied time points. The concentrations of total RNA were generally between 500 and 700 ng/μL in samples before 2 h (**Supplementary Figure S10**). Because there was no internal reference in cell-free systems, the sfGFP transcription level was analyzed with absolute quantification real-time PCR, in which the known sfGFP PCR product was used to make a standard curve of CT values attained from qRT-PCR and target DNA concentrations (**Supplementary Figure S11**). With the incubation going, the mRNA in 0 and 21% oxygen treatments followed a single-phase exponential decay rate (**Figure 2B**), as reported previously (Moore et al., 2018). The mRNA concentration increased first and decreased along with the incubation going, as the cell-free reagents running out. However, the transcription in 100% treatment appeared in a low state during the oxygen incubation. As the protein synthesis discussed above, both the 0 and 100% oxygen condition showed comparatively low expression. Coupled with the protein yield, the low oxygen concentration showed little influence in transcription but had a substantial effect on the translation process. However, the high oxygen concentration had a direct negative impact on transcription.

ATP Analysis of the Cell-Free System in Different Oxygen Concentrations

Based on the results above, the direct influences on the transcription and translation regarding different oxygen concentrations are still required to be found out. The primary function of oxygen *in vivo* was the generation of ATP (adenosine triphosphate) through the aerobic respiration in aerobes (Heffner et al., 2013). In the cell-free system, ATP was an essential energy

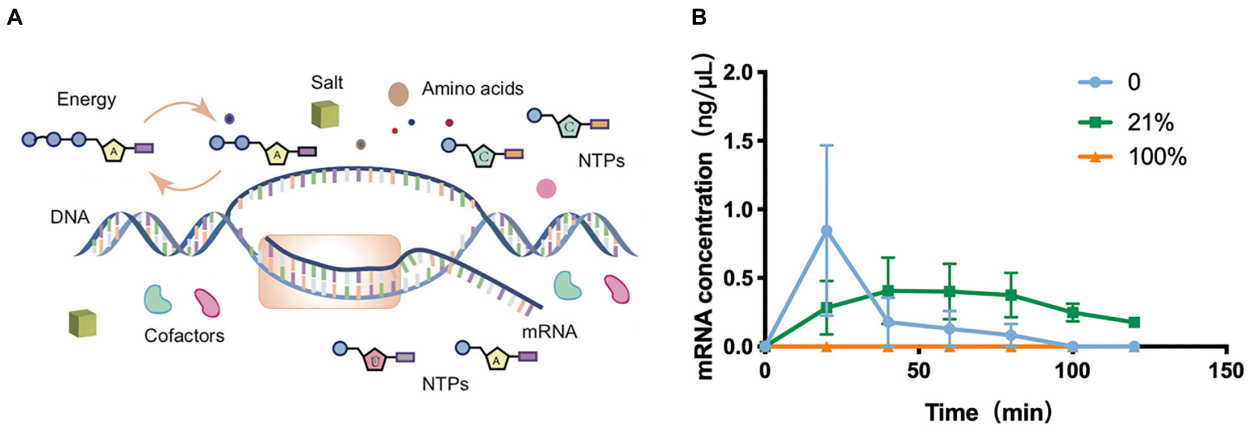


FIGURE 2 | Transcription in cell-free systems. **(A)** Schematic diagram of the transcription process in the cell-free protein synthesis system. **(B)** Transcriptional mRNA analysis of samples incubated in 0, 21, and 100%. Values were calculated as the mean of biological replicates.

substrate for the protein synthesis that was from the initial addition or energy regeneration. ATP concentrations were then analyzed, which was considered as a critical limiting factor for cell-free systems (Moon et al., 2019) and oxygen-derived products in natural physiological systems. In the cell-free systems, 1.2 mM ATP was initially added to provide energy for the protein synthesis along with PEP (phosphoenolpyruvate) as the energy source for the ATP regeneration (Kim and Swartz, 2000). Although PEP-based cell-free systems were not dependent on the oxygen for the ATP regeneration, the usage of ATP in protein synthesis probably involved with the oxygen.

ATP quantification relies on a series of energy-related reactions to transfer ATP concentration to the luminescence signal. Given concentrations of ATP were used to draw the standard curve to couple the luminescence intensity with ATP concentration for the ATP quantification (**Supplementary Figure S12**). The ATP assay was performed at varied time points of the oxygen treatment at different levels (**Figure 3A** and **Supplementary Figure S13**). As the protein synthesis reaction going, the ATP substrates were continuously and quickly consumed after 20 min. The quick decline could be mainly attributed to the protein synthesis. In the following incubation, ATP was in a tendency of decreasing or fluctuating, and it reached a rather low level at 2 h that had limited the synthesis reactions. It seemed that ATPs were continuously regenerated from degradative PEP energy source (Kim and Kim, 2009) and exhausted gradually in 2 h. Coupled with the protein yield results, the reduction of ATP was strongly related to the synthesis process. The higher protein yield corresponded to more rapid ATP consumption. However, from the ATP analysis at varied oxygen concentrations, it seemed that oxygen was not the decisive substrate for the ATP generation or consumption.

Reactive Oxygen Species (ROS) Analysis

According to the analysis above, ATP was the limitation substrate in the protein synthesis in all of the oxygen level. To gain more insight into the effects of the oxygen concentration on the protein synthesis, the oxygen byproducts that might affect the protein synthesis were further analyzed. The level of ROS (reactive oxygen species, O^{2-}, H_2O_2, and OH) was investigated in the oxygen-tuned cell-free system, which was considered as the metabolic intermediate of oxygen and affecting many physiological processes (Schieber and Chandel, 2014; Ezraty et al., 2017; Reichmann et al., 2018; Zhu and Dai, 2019).

The treatments of 0, 21, and 100% oxygen were chosen for the characterization, as they were representative in oxygen-tuned CFPS. The ROS levels were characterized by DCFH-DA (2,7-dichlorodihydrofluorescein diacetate). In this assay, DCFH-DA was catalyzed to DCFH (2,7-dichlorodihydrofluorescein), which could be oxidized by ROS to form fluorescent DCF (2,7-dichlorofluorescein) (**Figure 3B**). In this way, the ROS levels were compared by measuring the fluorescence levels of DCF (excitation at 488 nm, emission at 525 nm). Cell-free reactions with DCFH-DA were injected into the tube-in-tube reactors for incubation without fluorescent protein synthesis. Fluorescence was measured every 20 min in 2 h incubation to reflect the amount of ROS (**Figure 3B**). With the increase of the oxygen concentrations in the cell-free reactions, the ROS levels increased with the incubation time. It showed a low level in 0% oxygen treatment, higher in 21% oxygen treatment, and maximum in 100% oxygen treatment along with the incubation. It seemed that the fluorescence intensity was not apparent without oxygen treating, which indicated that there was little ROS generated under this condition. Compared to 0% oxygen treatment, the fluorescence in 21 and 100% increased rapidly, indicating that ROS was generated along with the reaction, which might have some harmful effects on protein synthesis.

Proteomics Analysis in Cell-Free Systems Between Different Oxygen Tension Treatments

To parameterize the protein transformation in the oxygen-tuned cell-free system, the proteomes of the CFPS system in 0, 21, and 100% after 4-h incubation were then analyzed.

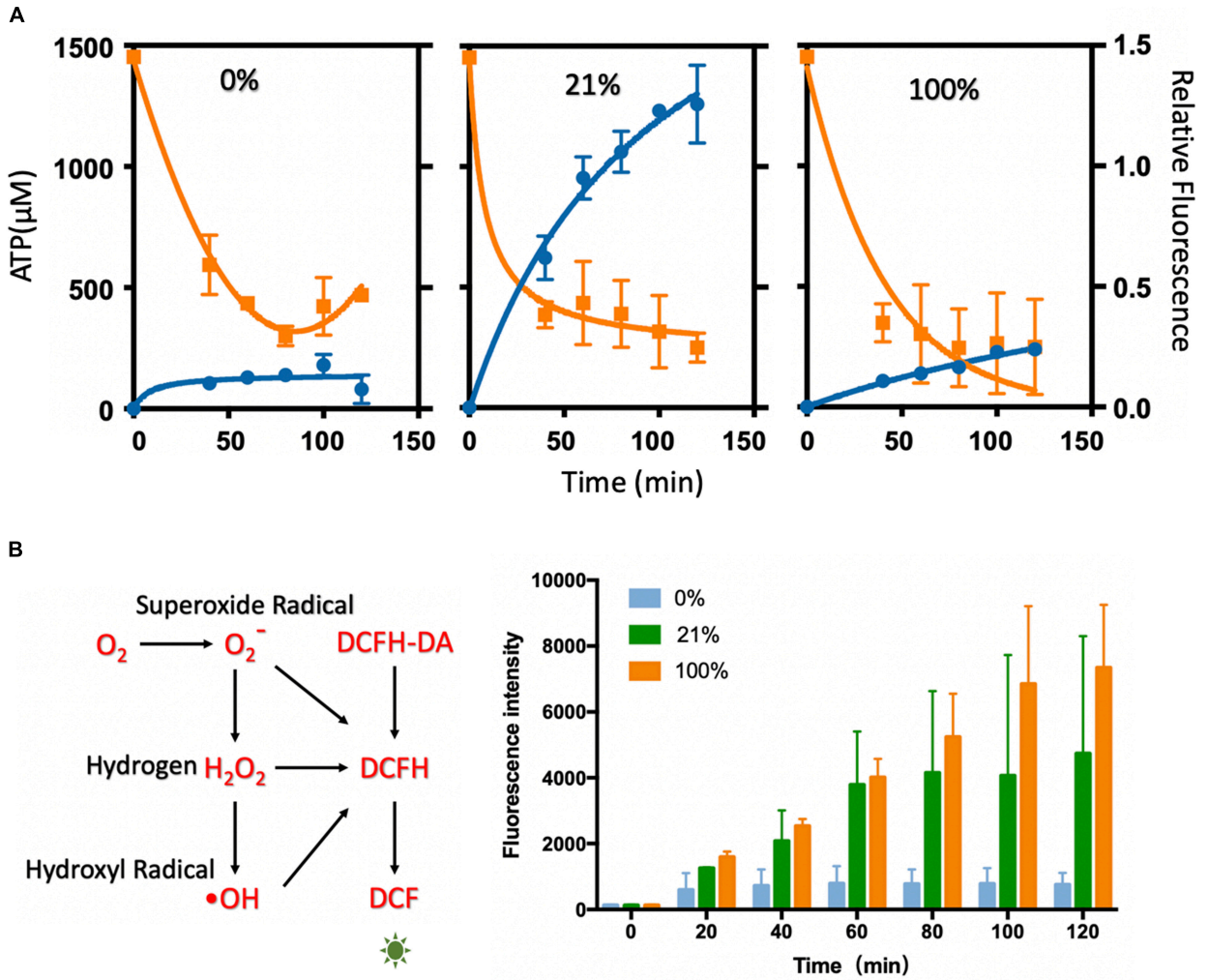


FIGURE 3 | Analysis of ATP level and ROS level in the cell-free systems under different oxygen concentrations. **(A)** ATP analysis of cell-free systems between different gradients of oxygen treatments (0, 21, and 100%). The ATP concentration (orange) was decreasing rapidly along with the protein synthesis (blue). The relative fluorescence was related to **Figure 1B**. **(B)** The ROS level between the oxygen of 0, 21, and 100%. With the incubation, the fluorescence rose rapidly in 21 and 100%, while little changes in 0% oxygen condition. Values were calculated as the mean of biological replicates.

Each treatment was in three replicates. The analysis was completed by LC-MS/MS. Approximately 1,800–2,100 proteins were validated in each sample, and the common 1,915 proteins were identified for further analysis (**Figures 4A,B**). Samples in the same condition group showed high repetitive (**Supplementary Figures S14–S17**). A total of 52 proteins with different abundance were selected out with ANOVA (Anscombe, 1948). Most of them were related to energy metabolism, transcription, translation, and protein folding (**Supplementary Figure S18**).

Efficient transcription and translation were the core processes in protein synthesis. Although additional T7 operon was mainly in charge of the sfGFP synthesis in this cell-free system, the complex endogenous transcription and translation components also played an essential role (Shin and Noireaux, 2012; Garamella et al., 2016). The GO (gene ontology) analysis was used for the protein function annotation in biological process, cellular component, and molecular function. Among the identified 1,915 proteins, 812 proteins were commonly selected for further analysis. Most of them were related to energy regeneration and protein translation (**Supplementary Figure S19**). For biological process, it showed that 40 proteins were annotated as "translation" and 15 proteins were annotated as "protein folding," which were enriched in cell-free systems for efficient protein synthesis. For cellular component, 62% of proteins were enriched in the term of "cytosol." For molecular function, most proteins were considered for protein binding for ATP, RNA, NAD, ribosome, and so on (**Supplementary Figure S18**). From the global proteomic analysis, the cell-free system involved with sufficient transcription and translation key protein complex.

Among the oxygen treatments, 52 proteins were screened out from the samples. Among the proteins in differential abundance, most of them were related to "translation," "rRNA binding,"

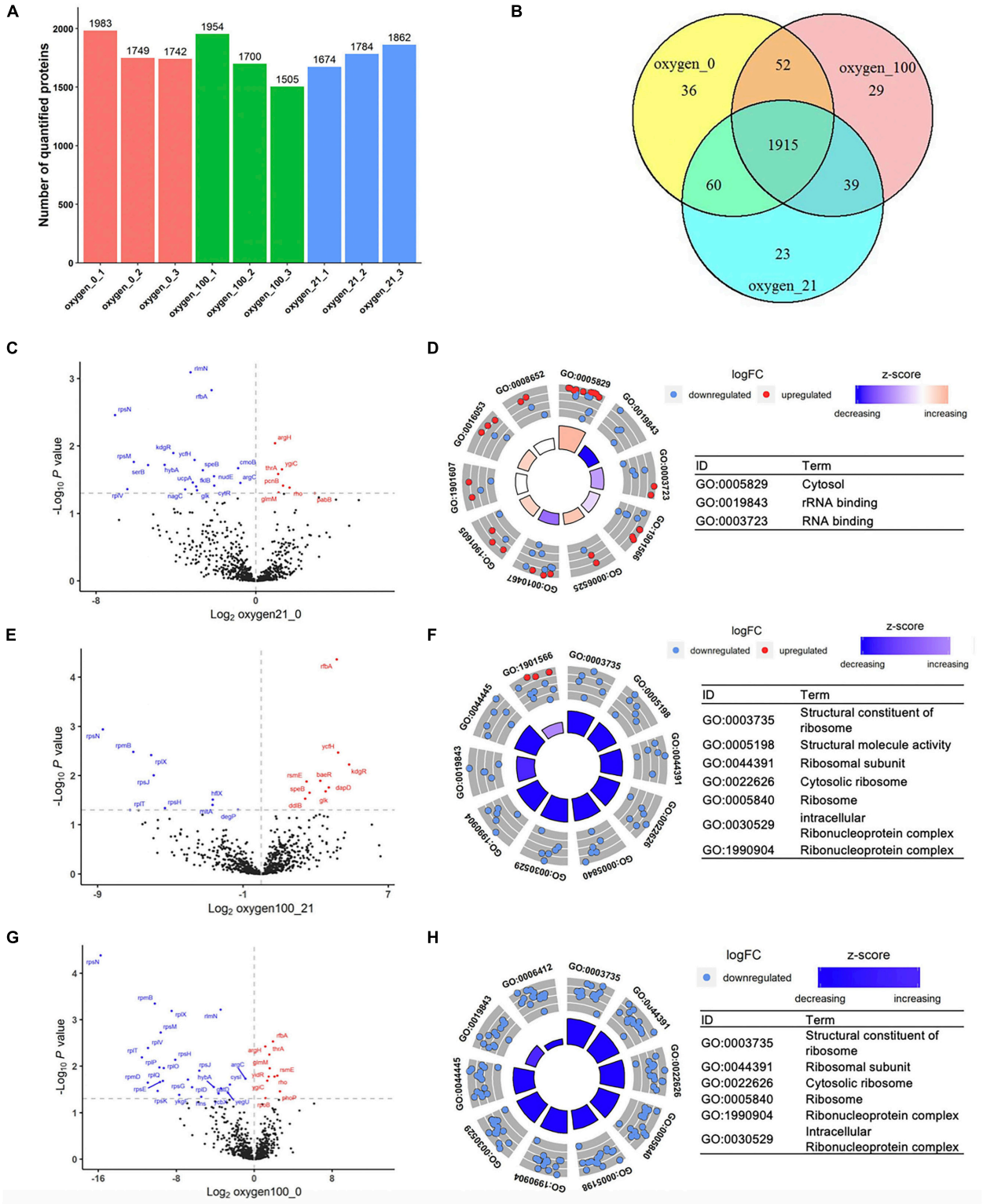


FIGURE 4 | Proteomic analysis of the cell-free systems under different oxygen concentrations (Oxygen_0 represented the samples of cell-free reactions conducted in 0% oxygen, oxygen_21 in 21% oxygen, and oxygen_100 in 100% oxygen. Each treatment was in three replicates.). **(A)** The number of quantified proteins among samples. **(B)** Number of common and unique proteins among samples in Venn Diagram. **(C,D)** Volcano plot and GOCircle plot of proteins in different abundance from the comparison from 21 to 0% oxygen condition. **(E,F)** Volcano plot and GOCircle plot of proteins in different abundance from the comparison from 100 to 21% oxygen condition. **(G,H)** Volcano plot and GOCircle plot of proteins in different abundance from the comparison from 100 to 0% oxygen condition. (More details of the GOCircle annotation were shown in **Supplementary Figures S23–S25**).

and "ribosomal subunit" (**Supplementary Figure S19**). With the increased oxygen concentration, proteins associated with the categories mentioned above were likely to be downregulated, that there was degradation in translation-related proteins in high oxygen concentration (**Figures 4C–H**). Compared to 0% oxygen treatment, in the 21% treatment samples, several genes were upregulated in the term of "cytosol" (**Figure 4D** and **Supplementary Figure S20**). In particular, most of the down-regulated proteins were in terms related to "rRNA binding" and "RNA binding," which were the key components in the translation processes. In the comparison from 100 to 21%, most proteins in the differential abundance were significantly enriched in ribonucleoprotein complex members (**Figure 4F** and **Supplementary Figure S21**), and also the proteins involved in translation. Familiar with the previous comparison, in the comparison from 100 to 0%, most proteins were also primarily enriched in ribosomal protein family as well as translation-related such as "tRNA binding" (**Figure 4H** and **Supplementary Figure S22**). According to the analysis discussed above, the higher oxygen samples showed a downregulation of translation-related proteins. Although the higher oxygen (21%) already resulted in the protein downregulation of translation-related proteins, the remaining cell-free systems were still sufficient for the protein synthesis. However, much higher oxygen (100%) would lead to further degradation that limited the protein synthesis. Therefore, the treatment by a high concentration of oxygen might lead to the degradation of translation-related proteins, which could significantly affect the CFPS processes.

CONCLUSION

The unknown behind the roles of oxygen in life continues to be a challenge for the fundamental physiological science. During the past decade, efforts made in microfluidics have demonstrated it is possible to explore the oxygen physiological effects in precise oxygen concentrations for the applications of fundamental biological or medical research. However, the ineffective methods for intracellular oxygen tension detection are still the limitations in the oxygen bioprocess analysis, mainly for the gas generation and gas concentration sensing system. To address the challenges mentioned above, an emerging tool that is recently widely used for the study of gas-liquid reactions were introduced. Without the intact cells, cell-free reactions are more likely to chemical reactions, where the substrate concentrations involved in bioprocess are concentrations in liquids, avoiding the limitation from the physical measurement.

Cell-free systems were employed for the analysis of the oxygen-tuned protein synthesis, coupled with a tube-in-tube reactor. The tube-in-tube reactor is regarded as a powerful tool for studying the gas-liquid reaction system, which allows gas to penetrate rapidly into the liquid in about 30 s. It could precisely control and alter the oxygen tension gradients. From this measurement platform, there appeared a maximum yield in the 21% oxygen condition, as the natural atmosphere condition. Meanwhile, no <4% oxygen tension can provide sufficient oxygen for the protein synthesis. Further analysis in mRNA, ATP, ROS, and proteome was conducted for the study of the internal connection between oxygen concentration and protein synthesis. The target mRNA in cell-free reactions seemed that there was little effect in low oxygen concentration (0%) but repressive in high concentration (100%) for nearly no transcription. With the protein synthesis, ATP showed a sharp degradation in the early protein synthesis process. Along with the increased oxygen concentration, ROS levels in cell-free reactions showed a related increase. Many translation-related proteins were downregulated in higher oxygen condition, which was of significant influence in protein synthesis. That is, a high concentration of oxygen seemed to result in a degradation of translational-related proteins and therefore limit the protein synthesis.

Oxygen-tuned protein synthesis machinery could synthesize proteins at a certain speed on demand by the regulation of the oxygen level. Coupled with great gas transfer performance in the tube-in-tube reactor and flexible CFPS system, this platform would be further well-adopted to explore gas-regulated biological cellular processes.

DATA AVAILABILITY STATEMENT

The datasets generated from this study are available in request to the corresponding authors: YL (yuanlu@tsinghua.edu.cn) and JZ (jiszhang@tsinghua.edu.cn). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the iProX partner repository (Ma et al., 2019) with the dataset identifier PXD018058.

AUTHOR CONTRIBUTIONS

YL and JZ conceived the project, supervised the research, assisted in analyzing data, preparing the figures, and writing the manuscript. XL and CZ designed and performed the experiments, analyzed the results, prepared the figures, and wrote the manuscript. SZ and HD assisted in bioinformatic analysis.

FUNDING

This work was supported by the National Key R&D Program of China (2018YFA0901700, 2019YFA0905100) and National Natural Science Foundation of China (21978146).

ACKNOWLEDGMENTS

We thank the Protein Chemistry Facility at the Center for Biomedical Analysis of Tsinghua University for sample analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00312/full#supplementary-material

REFERENCES

Abaci, H. E., Devendra, R., Soman, R., Drazer, G., and Gerecht, S. (2012). Microbioreactors to manipulate oxygen tension and shear stress in the microenvironment of vascular stem and progenitor cells. *Biotechnol. Appl. Biochem.* 59, 97–105. doi: 10.1002/bab.1010

Adler, M., Polinkovsky, M., Gutierrez, E., and Groisman, A. (2010). Generation of oxygen gradients with arbitrary shapes in a microfluidic device. *Lab. Chip* 10, 388–391. doi: 10.1039/B920401F

Anscombe, F. J. (1948). The validity of comparative experiments. *J. R. Stat. Soc. Ser. A* 111:181. doi: 10.2307/2984159

Brahimi-Horn, M. C., and Pouysségur, J. (2007). Oxygen, a source of life and stress. *FEBS Lett.* 581, 3582–3591. doi: 10.1016/J.FEBSLET.2007.06.018

Brennan, M. D., Rexius-Hall, M. L., Elgass, L. J., and Eddington, D. T. (2014). Oxygen control with microfluidics. *Lab. Chip* 14, 4305–4318. doi: 10.1039/C4LC00853G

Brzozowski, M., O'Brien, M., Ley, S. V., and Polyzos, A. (2015). Flow chemistry: intelligent processing of gas–liquid transformations using a tube-in-tube reactor. *Acc. Chem. Res.* 48, 349–362. doi: 10.1021/ar500359m

Buerk, D. G. (2004). Measuring tissue PO2 with microelectrodes. *Methods Enzymol.* 381, 665–690. doi: 10.1016/S0076-6879(04)81043-7

Burkholder, M., Gilliland, S., Luxon, A., Tang, C., and Gupton, B. (2019). Improving productivity of multiphase flow aerobic oxidation using a tube-in-tube membrane contactor. *Catalysts* 9:95. doi: 10.3390/catal9010095

Carmeliet, P., and Jain, R. K. (2000). Angiogenesis in cancer and other diseases. *Nature* 407, 249–257. doi: 10.1038/35025220

Clark, L. C., Wolf, R., Granger, D., and Taylor, Z. (1953). Continuous recording of blood oxygen tensions by polarography. *J. Appl. Physiol.* 6, 189–193. doi: 10.1152/jappl.1953.6.3.189

Ezraty, B., Gennaris, A., Barras, F., and Collet, J.-F. (2017). Oxidative stress, protein damage and repair in bacteria. *Nat. Rev. Microbiol.* 15, 385–396. doi: 10.1038/nrmicro.2017.26

Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016). The all *E. coli* TX-TL Toolbox 2.0: a platform for cell-free synthetic biology. *ACS Synthetic Biol.* 5, 344–355. doi: 10.1021/acssynbio.5b00296

Heffner, J. E., Casaburi, R., Toledo, A., Sampaio, L. M., Silva, T., and da Kunikushita, L. (2013). The story of oxygen. *Respirat. Care* 58, 18–31. doi: 10.4187/respcare.01831

Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., et al. (2001). HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science* 292, 464–468. doi: 10.1126/science.1059817

Jaakkola, P., Mole, D. R., Tian, Y.-M., Wilson, M. I., Gielbert, J., Gaskell, S. J., et al. (2001). Targeting of HIF-alpha to the von hippel-lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* 292, 468–472. doi: 10.1126/science.1059796

Jiao, X., Sherman, B. T., Huang, D. W., Stephens, R., Baseler, M. W., Lane, H. C., et al. (2012). DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics* 28, 1805–1806. doi: 10.1093/bioinformatics/bts251

Ke, Q., and Costa, M. (2006). Hypoxia-inducible factor-1 (HIF-1). *Mol. Pharmacol.* 70, 1469–1480. doi: 10.1124/mol.106.027029

Kim, D. M., and Swartz, J. R. (2000). Prolonging cell-free protein synthesis by selective reagent additions. *Biotechnol. Prog.* 16, 385–390. doi: 10.1021/bp000031y

Kim, H.-C., and Kim, D.-M. (2009). Methods for energizing cell-free protein synthesis. *J. Biosci. Bioeng.* 108, 1–4. doi: 10.1016/J.JBIOSC.2009.02.007

Koos, P., Gross, U., Polyzos, A., O'Brien, M., Baxendale, I., and Ley, S. V. (2011). Teflon AF-2400 mediated gas–liquid contact in continuous flow methoxycarbonylations and in-line FTIR measurement of CO concentration. *Org. Biomol. Chem.* 9:6903. doi: 10.1039/c1ob06017a

Lam, R. H. W., Kim, M.-C., and Thorsen, T. (2009). Culturing aerobic and anaerobic bacteria and mammalian cells with a microfluidic differential oxygenator. *Anal. Chem.* 81, 5918–5924. doi: 10.1021/ac9006864

Leclerc, E., Sakai, Y., and Fujii, T. (2004). Microfluidic PDMS (polydimethylsiloxane) bioreactor for large-scale culture of hepatocytes. *Biotechnol. Prog.* 20, 750–755. doi: 10.1021/bp0300568

Lindahl, S. G. E. (2008). Oxygen and life on earth. *Anesthesiology* 109, 7–13. doi: 10.1097/ALN.0b013e31817b5a7e

Lu, Y. (2017). Cell-free synthetic biology: engineering in an open world. *Synthetic Syst. Biotechnol.* 2, 23–27. doi: 10.1016/J.SYNBIO.2017.02.003

Ma, J., Chen, T., Wu, S., Yang, C., Bai, M., Shu, K., et al. (2019). iProX: an integrated proteome resource. *Nucleic Acids Res.* 47, D1211–D1217. doi: 10.1093/nar/gky869

Marshall, R., Maxwell, C. S., Collins, S. P., Jacobsen, T., Luo, M. L., Begemann, M. B., et al. (2018). Rapid and scalable characterization of CRISPR technologies using an *E. coli* cell-free transcription-translation system. *Mol. Cell* 69:146-157.e3. doi: 10.1016/J.MOLCEL.2017.12.007

Mohyeldin, A., Garzón-Muvdi, T., and Quiñones-Hinojosa, A. (2010). Oxygen in stem cell biology: a critical component of the stem cell niche. *Cell Stem Cell* 7, 150–161. doi: 10.1016/J.STEM.2010.07.007

Moon, B.-J., Lee, K.-H., and Kim, D.-M. (2019). Effects of ATP regeneration systems on the yields and solubilities of cell-free synthesized proteins. *J. Indus. Eng. Chem.* 70, 276–280. doi: 10.1016/J.JIEC.2018.10.027

Moore, S. J., MacDonald, J. T., Wienecke, S., Ishwarbhai, A., Tsipa, A., Aw, R., 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.* 2017:1715806115. doi: 10.1073/pnas.1715806115

Nozue, M., Lee, I., Yuan, F., Teicher, B. A., Brizel, D. M., Dewhirst, M. W., et al. (1997). Interlaboratory variation in oxygen tension measurement by Eppendorf "Histograph" and comparison with hypoxic marker. *J. Surg. Oncol.* 66, 30–38. doi: 10.1002/(sici)1096-9098(199709)66:1<30::aid-jso7>3.0.co;2-o

O'brien, M., Taylor, N., Polyzos, A., Baxendale, I. R., and Ley, S. V. (2011). Hydrogenation in flow: Homogeneous(and)heterogeneous catalysis using Teflon AF-2400 to effect gas-liquid contact at elevated pressure. *Chem. Sci.* 2, 1250–1257. doi: 10.1039/c1sc00055a

Palmer, B. F., and Clegg, D. J. (2014). Oxygen sensing and metabolic homeostasis. *Mol. Cell. Endocrinol.* 397, 51–58. doi: 10.1016/J.MCE.2014.08.001

Perez, J. G., Stark, J. C., and Jewett, M. C. (2016). Cell-free synthetic biology: engineering beyond the cell. *Cold Spring Harb. Perspect. Biol.* 8:a023853. doi: 10.1101/cshperspect.a023853

Petersen, T. P., Polyzos, A., O'Brien, M., Ulven, T., Baxendale, I. R., and Ley, S. V. (2012). The oxygen-mediated synthesis of 1,3-butadiynes in continuous flow: using teflon AF-2400 to effect gas/liquid contact. *ChemSusChem* 5, 274–277. doi: 10.1002/cssc.201100339

Polinkovsky, M., Gutierrez, E., Levchenko, A., and Groisman, A. (2009). Fine temporal control of the medium gas content and acidity and on-chip generation of series of oxygen concentrations for cell cultures. *Lab Chip* 9:1073. doi: 10.1039/b816191g

Polyzos, A., O'Brien, M., Petersen, T. P., Baxendale, I. R., and Ley, S. V. (2011). The continuous-flow synthesis of carboxylic acids using CO2 in a tube-in-tube gas permeable membrane reactor. *Angewandte Chem. Int. Edn.* 50, 1190–1193. doi: 10.1002/anie.201006618

Reichmann, D., Voth, W., and Jakob, U. (2018). Maintaining a Healthy Proteome during Oxidative Stress. *Mol. Cell.* 69, 203–213. doi: 10.1016/j.molcel.2017.12.021

Ringborg, R. H., Toftgaard Pedersen, A., and Woodley, J. M. (2017). *Automated Determination of Oxygen-Dependent Enzyme Kinetics in a Tube-in-Tube Flow Reactor*. Hoboken, NJ: Wiley Blackwell, doi: 10.1002/cctc.201701295

Roussakis, E., Li, Z., Nichols, A. J., and Evans, C. L. (2015). Oxygen-sensing methods in biomedicine from the macroscale to the microscale. *Angewandte Chem. Int. Edn.* 54, 8340–8362. doi: 10.1002/anie.201410646

Schieber, M., and Chandel, N. S. (2014). ROS function in redox signaling and oxidative stress. *Curr. Biol.* 24, R453–R462. doi: 10.1016/j.cub.2014.03.034

Shin, J., and Noireaux, V. (2012). An *E. coli* cell-free expression toolbox: application to synthetic gene circuits and artificial cells. *ACS Synthetic Biol.* 1, 29–41. doi: 10.1021/sb200016s

Sud, D., Mehta, G., Mehta, K., Linderman, J., Takayama, S., and Mycek, M.-A. (2006). Optical imaging in microfluidic bioreactors enables oxygen monitoring for continuous cell culture. *J. Biomed. Opt.* 11:050504. doi: 10.1117/1.2355665

Sun, Z. Z., Hayes, C. A., Shin, J., Caschera, F., Murray, R. M., and Noireaux, V. (2013). Protocols for Implementing an *Escherichia coli* Based TX-TL cell-free expression system for synthetic biology. *J. Vis. Exp.* 79:e50762. doi: 10.3791/50762

Wang, L. V., and Hu, S. (2012). Photoacoustic tomography: in vivo imaging from organelles to organs. *Science* 335, 1458–1462. doi: 10.1126/science.1216210

Yang, L., and Jensen, K. F. (2013). Mass transport and reactions in the tube-in-tube reactor. *Org. Process Res. Dev.* 17, 927–933. doi: 10.1021/op400085a

Zhang, J., Teixeira, A. R., Zhang, H., and Jensen, K. F. (2017). Automated in situ measurement of gas solubility in liquids with a simple

tube-in-tube reactor. *Anal. Chem.* 89, 8524–8530. doi: 10.1021/acs.analchem.7b02264

Zhang, J., Teixeira, A. R., Zhang, H., and Jensen, K. F. (2019). Flow toolkit for measuring gas diffusivity in liquids. *Anal. Chem.* 91, 4004–4009. doi: 10.1021/acs.analchem.8b05396

Zhou, C., Lin, X., Lu, Y., and Zhang, J. (2020). Flexible on-demand cell-free protein synthesis platform based on a tube-in-tube reactor. *React. Chem. Eng.* 5, 270–277. doi: 10.1039/c9re00394k

Zhu, M., and Dai, X. (2019). Maintenance of translational elongation rate underlies the survival of *Escherichia coli* during oxidative stress. *Nucleic Acids Res.* 47, 7592–7604. doi: 10.1093/nar/gkz467

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

REVIEW
published: 12 May 2020
doi: 10.3389/fbioe.2020.00399

Biological Materials: The Next Frontier for Cell-Free Synthetic Biology

*Richard J. R. Kelwick[1], Alexander J. Webb[1] and Paul S. Freemont[1,2,3]**

[1] Section of Structural and Synthetic Biology, Department of Infectious Disease, Imperial College London, London, United Kingdom, [2] The London Biofoundry, Imperial College Translation & Innovation Hub, London, United Kingdom, [3] UK Dementia Research Institute Care Research and Technology Centre, Imperial College London, London, United Kingdom

Advancements in cell-free synthetic biology are enabling innovations in sustainable biomanufacturing, that may ultimately shift the global manufacturing paradigm toward localized and ecologically harmonized production processes. Cell-free synthetic biology strategies have been developed for the bioproduction of fine chemicals, biofuels and biological materials. Cell-free workflows typically utilize combinations of purified enzymes, cell extracts for biotransformation or cell-free protein synthesis reactions, to assemble and characterize biosynthetic pathways. Importantly, cell-free reactions can combine the advantages of chemical engineering with metabolic engineering, through the direct addition of co-factors, substrates and chemicals –including those that are cytotoxic. Cell-free synthetic biology is also amenable to automatable design cycles through which an array of biological materials and their underpinning biosynthetic pathways can be tested and optimized in parallel. Whilst challenges still remain, recent convergences between the materials sciences and these advancements in cell-free synthetic biology enable new frontiers for materials research.

Keywords: cell-free synthetic biology, biological materials, biomaterials, biomimetics, metabolic engineering

OPEN ACCESS

Edited by:
Jian Li,
ShanghaiTech University, China

Reviewed by:
Lei Kai,
Jiangsu Normal University, China
Chris John Myers,
The University of Utah, United States

***Correspondence:**
Paul S. Freemont
p.freemont@imperial.ac.uk

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

Received: *10 March 2020*
Accepted: *08 April 2020*
Published: *12 May 2020*

Citation:
Kelwick RJR, Webb AJ and Freemont PS (2020) Biological Materials: The Next Frontier for Cell-Free Synthetic Biology. Front. Bioeng. Biotechnol. 8:399. doi: 10.3389/fbioe.2020.00399

INTRODUCTION

We live in a material world (Callén Moreu and López Gómez, 2019). There are more than five trillion plastic pieces in the world's oceans (Eriksen et al., 2014), an estimated 15 billion trees are cut down each year (Crowther et al., 2015) and global natural fiber production is estimated to have already exceeded 30 million tons per annum (Townsend and Sette, 2016). Clearly, the mass production, processing and disposal of an array of materials has significantly shaped the global ecosystem. Whilst we continue to approach irreversible climate change (Lenton et al., 2019), it is of increasing importance that positive advances in the materials sciences are balanced against any negative environmental consequences that relate to material consumption. Arguably, these challenges warrant significant shifts in the manufacturing paradigm, from global mass production to local, sustainable and personalized manufacturing strategies (Hu, 2013; Stock and Seliger, 2016; Kleer and Piller, 2019). Likewise, instead of the chemical industries, the next generation of materials may come through developments in synthetic biology, sustainable biotechnology and the burgeoning bioeconomy (Le Feuvre and Scrutton, 2018; Philp, 2018; Freemont, 2019; French, 2019).

The field of synthetic biology has emerged over the last twenty years into a highly dynamic community of interdisciplinary researchers and societal stakeholders, that are working toward the

responsible development and implementation of cutting-edge biotechnologies (Cameron et al., 2014; El Karoui et al., 2019; Lai et al., 2019). Synthetic biology combines knowledge across numerous scientific disciplines including molecular biology, biochemistry, biophysics and the social sciences, whilst also integrating them within an engineering design framework (Anderson et al., 2019; Trump et al., 2019). Ultimately, the synthetic biology approach is geared toward the rational engineering or repurposing of biological parts, devices and systems into solutions that can help address societal challenges (Ausländer et al., 2017). Synthetic biologists envision a broad application space for the field and anticipate positive societal impacts in medical technologies (Pardee, 2018; Hicks et al., 2020) (e.g., biosensors and therapeutics), food security and sustainable energy (Russo et al., 2019; Roell and Zurbriggen, 2020) (e.g., biofuels and synthetic photosynthesis), bioremediation (Dvořák et al., 2017; Wan et al., 2019) (e.g., pollution monitoring and sequestration), education (Kelwick et al., 2015a; Huang et al., 2018; Dy et al., 2019) (e.g., iGEM competition) and biomanufacturing (Le Feuvre and Scrutton, 2018; Choi K.R. et al., 2019; Gilbert and Ellis, 2019; Roberts et al., 2019) (e.g., fine chemicals and materials production).

However, biological systems are highly complex and initial attempts at engineering biological systems to fulfill specific application goals, are often only partially successful. To help overcome these challenges, synthetic biology employs the concept of the design-cycle, through which biotechnologies are iteratively designed, built and tested (Arpino et al., 2013; Church et al., 2014; Kelwick et al., 2014). Learning how to improve a biological system may require multiple attempts, which could be made easier by more rapid and systematic workflows. One potential solution is to utilize cell-free synthetic biology for rapid prototyping (Moore et al., 2017b; Kelwick et al., 2018). Typically, cell-free reactions make use of isolated cellular components and machinery (e.g., ribosomes and recombinant proteins), rather than live whole-cells. Whilst, some cell-free reaction components may be cell-derived (e.g., cell extracts), once prepared, cell-free workflows can be completed within hours (Chappell et al., 2013). In contrast, typical whole-cell experiments may involve several days or weeks of delays that are associated with plasmid cloning, transformation and cell growth (Sun et al., 2014). Furthermore, cell-free reactions are accessible and can combine the advantages of chemical engineering with metabolic engineering, through the direct addition of enzyme co-factors, substrates and chemicals – including those that are cytotoxic (Dudley et al., 2015; Karim and Jewett, 2016; Kay and Jewett, 2020). These advantages are increasingly being exploited for cell-free applications including biopart prototyping, cell-free metabolic engineering, medical or environmental biosensors and on-demand therapeutics production (Kightlinger et al., 2019; Silverman et al., 2019). Based upon an expanding repertoire of examples in the literature, we envision that biological materials and bio-functionalized smart materials are the next frontier for cell-free synthetic biology (**Table 1**). To this end, this review will introduce key concepts and recent developments in cell-free synthetic biology, with a focus on examples relevant to the materials sciences. Examples will be given of industrially and societally important biological materials that have been generated using cell-free synthetic biology. Cell-free synthetic biology can also be utilized to bio-functionalize materials, which may further enable the emergence of new types of smart materials. This review will also explore future trends and challenges in cell-free synthetic biology and speculate on their potential impact on biological materials of the future.

CELL-FREE SYNTHETIC BIOLOGY REACTION FORMATS AND STRATEGIES

Cell-free synthetic biology is a broad term that encompasses many different *in vitro* biotechnologies. Broadly, the term cell-free synthetic biology refers to different methods and technologies for engineering or using biological processes outside of a cell. For example, cell-free protein synthesis reactions enable the production of proteins within biochemical reactions. Thus, cell-free reactions typically make use of isolated cellular components (e.g., recombinant proteins) and/or cell extracts, rather than live whole-cells. In the context of this review four commonly used cell-free reaction formats will be discussed (**Figure 1**). We describe these cell-free reaction formats as either (i) recombinant enzyme-based, (ii) protein synthesis using recombinant elements (PURE)-based cell-free protein synthesis, (iii) wildtype and/or engineered cell extract biotransformation or (iv) cell extract-based cell-free protein synthesis.

Recombinant enzyme-based reaction formats utilize purified enzymes, along with any required co-factors and pathway substrates, to produce fine chemicals, polymer monomers or other molecules of interest. The PURE-based cell-free protein synthesis format reconstitutes the transcription and translation machinery from *Escherichia coli* using purified histidine (His)-tagged proteins (Shimizu et al., 2001, 2005). In this reaction format, the exact components are known, including the co-factors, substrates and energy mixes. Since PURE reaction components are known they can be standardized and rationally optimized. However, PURE cell-free reactions typically produce lower protein yields than cell-free protein synthesis reactions that use *E. coli* extracts (Shimizu et al., 2005). The third cell-free reaction format uses cell extracts from lysed wildtype and/or engineered cells, which can be mixed together along with relevant required enzyme co-factors and substrates to form multicomponent biosynthetic pathways. Finally, the last format, cell extract-based cell-free protein synthesis (CFPS), uses the transcription and translation machinery from lysed cells, along with added co-factors and energy mixes to produce proteins *in vitro*. Cellular extract-based cell-free reactions use the host cells native transcription and translation machinery as well as other metabolic components, including energy providing enzymes such as those involved in glycolysis and the Krebs cycle, which are released when the cells are lysed either mechanically (French pressure cell press), by sonication or osmotically (Gregorio et al., 2019). These reactions were originally developed as experimental tools to enable the fundamental understanding of aspects of cellular biochemistry, molecular biology and *in vitro* production of various proteins of interest (Gagoski et al., 2016). A range of

TABLE 1 | Cell-free strategies for biological material biomanufacturing or material bio-functionalization.

Cell-free reaction format	Material	Application	References
Recombinant enzymes	Polyhydroxyalkanoates (PHAs)	Biopolymer production	Thomson et al., 2009; Han et al., 2011; Tomizawa et al., 2012; Opgenorth et al., 2016
	Lactic acid	Platform material for polymer production	Kopp et al., 2019
Cell extract biotransformation	Bio-cellulose	Bio-cellulose production	Ullah et al., 2015
	Chitin	Chitin synthesis	Jaworski et al., 1963
	Poly-3-hydroxybutyrate [P(3HB)]	Optimizing PHAs biopolymer production	Kelwick et al., 2018
	Gold nanoparticles (AuNPs)	Medical and industrial	Chauhan et al., 2011; Krishnan et al., 2016
	Silver nanoparticles (AgNPs)	Nanobiotechnology, therapeutic development	Costa Silva et al., 2017
Cell-free protein synthesis	Bacteriophages	*De novo* synthesis and phage engineering	Garamella et al., 2016; Rustad et al., 2018
	Chitin	Chitinase expression	Endoh et al., 2006
	Clay microgels	Protein production	Jiao et al., 2018
	DNA hydrogels/Protein-producing gels (P-gel)	Protein production	Park et al., 2009a; Ruiz et al., 2012
	Elastin-like polypeptides (ELPs)	Biopolymer with non-canonical amino acids	Martin et al., 2018
	Extracellular vesicles (EVs)	Therapeutics/EV biogenesis research	Shurtleff et al., 2016; García-Manrique et al., 2018
	Freeze-dried pellets	*In vitro* diagnostics or therapeutic production	Pardee et al., 2016b; Salehi et al., 2016, 2017
	Liposomes and nanodiscs	Membrane protein production, drug discovery or protocell production	Garamella et al., 2016; Rues et al., 2016; Shinoda et al., 2016; Contreras-Llano and Tan, 2018; Gessesse et al., 2018; Dubuc et al., 2019; Shelby et al., 2019
	Microfluidic devices (various)	Antibody development and protein microarrays	Kilb et al., 2014; Georgi et al., 2016; Contreras-Llano and Tan, 2018
	Microparticles/nanoparticles	On-demand functional biomaterials/therapeutics	Lim et al., 2009; Benítez-Mateos et al., 2018
	Paper	*In vitro* diagnostics	Pardee et al., 2014, 2016a; Duyen et al., 2017; Gräwe et al., 2019; Thavarajah et al., 2020
	PEG hydrogels	Education	Huang et al., 2018
	Poly-3-hydroxybutyrate (P(3HB))	Polyhydroxyalkanoates (PHAs) biosynthetic operon prototyping	Kelwick et al., 2018
	Protein biologics	Cancer therapeutics, protein therapeutics	Zawada et al., 2011; Sullivan et al., 2016; Salehi et al., 2017; Kightlinger et al., 2019
	Silk fibroin	Silk fibroin production	Greene et al., 1975; Lizardi et al., 1979

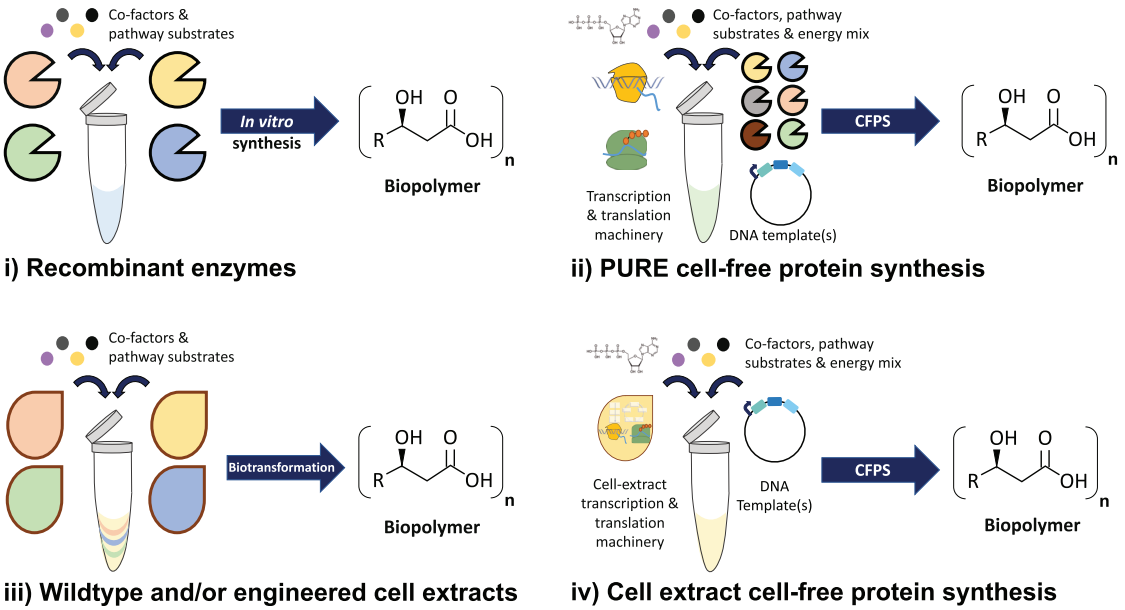


FIGURE 1 | Cell-free synthetic biology reaction formats and strategies. **(i)** Recombinant enzymes can be mixed together along with enzyme co-factors and substrates to form biosynthetic pathways. **(ii)** The PURE cell-free protein synthesis system utilizes reconstituted *Escherichia coli* transcription and translation machinery, DNA templates, purified enzymes and other factors. **(iii)** Cell extracts from lysed wildtype or engineered cells can be mixed together along with enzyme co-factors and substrates to form biosynthetic pathways. **(iv)** Cell extract-based cell-free protein synthesis reactions utilize the transcription and translation machinery within cell lysates, along with exogenously added energy mix components (e.g., amino acids) and DNA templates for *in vitro* protein production.

different host cells have been used to develop these reactions, including bacteria such as *Bacillus subtilis* (Kelwick et al., 2016), *Streptomyces venezuelae* (Moore et al., 2017a; Li et al., 2018) and *E. coli* (Sun et al., 2013) as well as insect (Ezure et al., 2006), wheat germ (Harbers, 2014), yeast (Hodgman and Jewett, 2013; Aw and Polizzi, 2019), protozoans such as *Leishmania tarentolae* (Mureev et al., 2009; Kovtun et al., 2010, 2011) and mammalian cells (Weber et al., 1975; Martin et al., 2017).

It is important to note that these different cell-free reaction formats are not mutually exclusive and can be combined together. Recombinant enzymes or small molecule substrates can also be added into cell-free protein synthesis reactions to complete biosynthetic pathways, or to use exogenous chemistries within the reaction. It is this flexibility that we envision being particularly useful in terms of exploring how cell-free reactions could create novel types of biological materials or bio-functionalized smart materials. In the following sections we discuss exemplars where cell-free synthetic reactions have been utilized to prototype, manufacture or bio-functionalize, biological materials (**Table 1**).

CELL-FREE STRATEGIES FOR SUSTAINABLE MATERIALS BIOMANUFACTURING

Living cells and organisms have evolved highly complex enzymes and metabolic processes that generate extremely diverse biochemistries. Exploring these natural biochemistries may lead to important foundational advances in our understanding of natural product synthesis. Foundational discoveries in functional genomics, cellular metabolism and natural product synthesis are also important, because they might inspire novel biosynthetic pathway designs for biological materials production. In synthetic biology, cell-free metabolic engineering (CF-ME) approaches can reconstitute entire biosynthetic pathways using either cell extracts from diverse species, engineered cells and/or cell-free synthesized recombinant enzymes (Karim and Jewett, 2018; Martin et al., 2018; Yim et al., 2019; Bowie et al., 2020) (**Figure 1**). Also, cell-free protein synthesis and cell extract biotransformation reactions can be combined to create more complex cell-free reactions (Karim and Jewett, 2018; Kelwick et al., 2018). Another important advantage in using cell-free approaches is that pathway reaction bottlenecks can be identified, through the direct addition of the required recombinant enzymes, enzyme co-factors or chemical substrates needed for each stage of a biosynthetic pathway (Dudley et al., 2015). Increasingly sophisticated combinatorial CF-ME strategies, together with high-throughput automation, deep data omics and design of experiments (DoE) approaches to cell-free reaction optimization have been deployed (Caschera et al., 2018; Jiang et al., 2018; Dopp et al., 2019). These advancements have considerably improved the feasibility of refactoring and optimizing fine chemical or natural product biosynthetic pathways within short timeframes (Dudley et al., 2015; Korman et al., 2017; Moore et al., 2017a; Wilding et al., 2018).

Cell-free synthetic biology approaches have also been directed toward the *de novo* bioproduction of biological materials, including biopolymers or their monomers, cellulosic materials and nanoparticles (**Table 1**). However, the maximum cell-free

bioproduction yields or reaction efficiencies of several reported materials were generally low or unspecified. Examples of cell-free produced materials and their reported maximum production yields and reaction efficiencies include bio-cellulose (3.726 ± 0.05 g/L; 57.68%) (Ullah et al., 2015), chitin (yields not stated) (Jaworski et al., 1963; Endoh et al., 2006), lactic acid (6.6 ± 0.1 mM; 47.4 ± 3.9%) (Kopp et al., 2019), gold nanoparticles (yields not stated) (Chauhan et al., 2011; Krishnan et al., 2016), (*R*)-3-hydroxybutyrate-CoA (32.87 ± 6.58 μM) (Kelwick et al., 2018), silver nanoparticles (yields not specified) (Costa Silva et al., 2017) and silk fibroin (yields not specified) (Greene et al., 1975; Lizardi et al., 1979). Poor cell-free production yields and efficiencies can be due to a variety of factors including rapid depletion of reaction energy mix components (e.g., ATP, amino acids), the formation of inhibitory waste products (e.g., inorganic phosphates) or unwanted side reactions that divert reaction fluxes away from desirable pathways (Caschera and Noireaux, 2014). Because of these limitations, cell-free synthetic biology may not be an ideal production method for some biological materials. Nevertheless, whilst actual cell-free material production yields can be relatively low, these approaches are still beneficial for prototyping different biosynthetic pathways, substrates or reaction conditions to boost both *in vitro* and whole-cell production yields. An exemplar is the use of cell-free assays to characterize polyhydroxyalkanoates (PHAs) biosynthetic pathways from *phaCAB* operons that also enhanced *in vivo* PHAs production (Kelwick et al., 2018). Furthermore, the same study also demonstrated that the cell-extract biotransformation of whey permeate into 3-hydroxybutyrate (3HB), could be simultaneously coupled with the cell-free protein synthesis of a potential Acetyl-CoA recycling enzyme (Kelwick et al., 2018). Thus, highlighting that combinatorial cell-free reaction formats can be a useful strategy for bioplastic pathway prototyping and optimization.

Interestingly, in some cases, cell-free bioproduction may actually be a more desirable manufacturing route. For instance, several *in vitro* gold or silver nanoparticle production studies reported desirable nanoparticle characteristics (e.g., size/zeta potential) and/or easier purification protocols within cell-free bioproduction reactions than whole-cell production methods (Krishnan et al., 2016; Costa Silva et al., 2017). Cell-free bioproduction can also be carried out at industrially relevant scales, as illustrated by Sutro biopharma who have developed a highly scalable good manufacturing practices (GMP)-compliant cell-free protein synthesis platform, for producing therapeutic proteins within 100 L bioreactor reaction volumes (Zawada et al., 2011). For cell-free materials production, a highly efficient synthetic biochemistry module was developed to convert glucose into bio-based chemicals, including the PHA bioplastic monomer polyhydroxybutyrate (PHB) (Opgenorth et al., 2016). To achieve this, purified recombinant enzymes were used to reconstitute core elements of the pentose, bifido, glycolysis and PHB pathways (Opgenorth et al., 2016). Cell-free PHB production yields (40 g/L) and efficiencies (90%) were impressive and are promisingly close to industrially attractive scales (Opgenorth et al., 2016). These improvements in PHB production are also welcome since PHB, as well as other PHAs biopolymers, are biodegradable and can potentially be used as 'drop-in' replacements for oil-derived plastics (e.g., food packaging) or as biomaterials for tissue engineering (Choi S.Y. et al., 2019; Tarrahi et al., 2020). PHAs are also an interesting example because of their industrial importance and the diversity of cell-free strategies that have been applied to PHAs research (**Table 1**). Building upon these examples, CFME approaches could be used to explore a greater diversity of PHAs biopolymers given that PHA biopolymers can be composed of a variety of different wildtype and/or synthetic monomers (∼160 different monomers exist) to create complex co-polymers, with an array of material characteristics (Choi S.Y. et al., 2019). Future cell-free synthetic biology explorations of PHAs are likely to unlock novel PHAs biopolymers with unique characteristics (Chen and Hajnal, 2015) and therefore, accelerate bioplastic materials development.

Microbial (*in vivo*) PHAs production has been commercially manufactured at industrial scales over the last several decades. Unfortunately, the commercial impact of PHA-based bioplastics has been historically prohibited by their higher production costs than oil-derived plastics (Chen et al., 2020). However, more efficient PHAs production processes have been devised through the rational design of *phaCAB* biosynthetic pathways (Hiroe et al., 2012; Kelwick et al., 2015b; Li et al., 2016; Tao et al., 2017; Zhang X. et al., 2019), key metabolite recycling processes (e.g., Acetyl-CoA) (Matsumoto et al., 2013; Beckers et al., 2016), alternative microbial production hosts (e.g., *Halomonas* sp., Tan et al., 2011) and the use of industrially sourced, low-cost feedstocks (e.g., whey permeate) (Wong and Lee, 1998; Ahn et al., 2000; Kim, 2000; Nikel et al., 2006; Cui et al., 2016; Nielsen et al., 2017). Interestingly, several of these microbial PHAs production strategies are also compatible with cell-free synthetic biology reactions. In particular, using locally sourced, low-cost feedstocks (e.g., whey permeate) may help to make cell-extract based PHAs production more economically viable (Kelwick et al., 2018). A similar approach has already been piloted for cell-free lactic acid production from spent coffee grounds (Kopp et al., 2019) and could become a generalized strategy for sustainable cell-free materials bioproduction (**Figure 2**). We would argue that combining cell-free extracts with local feedstocks enables immediate access to highly diverse cellular biochemistries and low-cost substrates (e.g., waste feedstocks), that could potentially be used for the sustainable biomanufacturing of a diverse array of biological materials (Yan and Fong, 2015; Le Feuvre and Scrutton, 2018). Furthermore, just as lyophilized cell-free reactions enable the on-demand production of biotherapeutics (Pardee et al., 2016b), we likewise envision that cell-free reactions might one day lead to rapid and distributed, on-demand biological materials production or bio-functionalization.

CELL-FREE MATERIAL BIO-FUNCTIONALIZATION AND BIOMIMETICS

Cell-free systems are still being applied to understand the foundational principles and molecular mechanisms of transcription and translation (Hecht et al., 2017; Borkowski

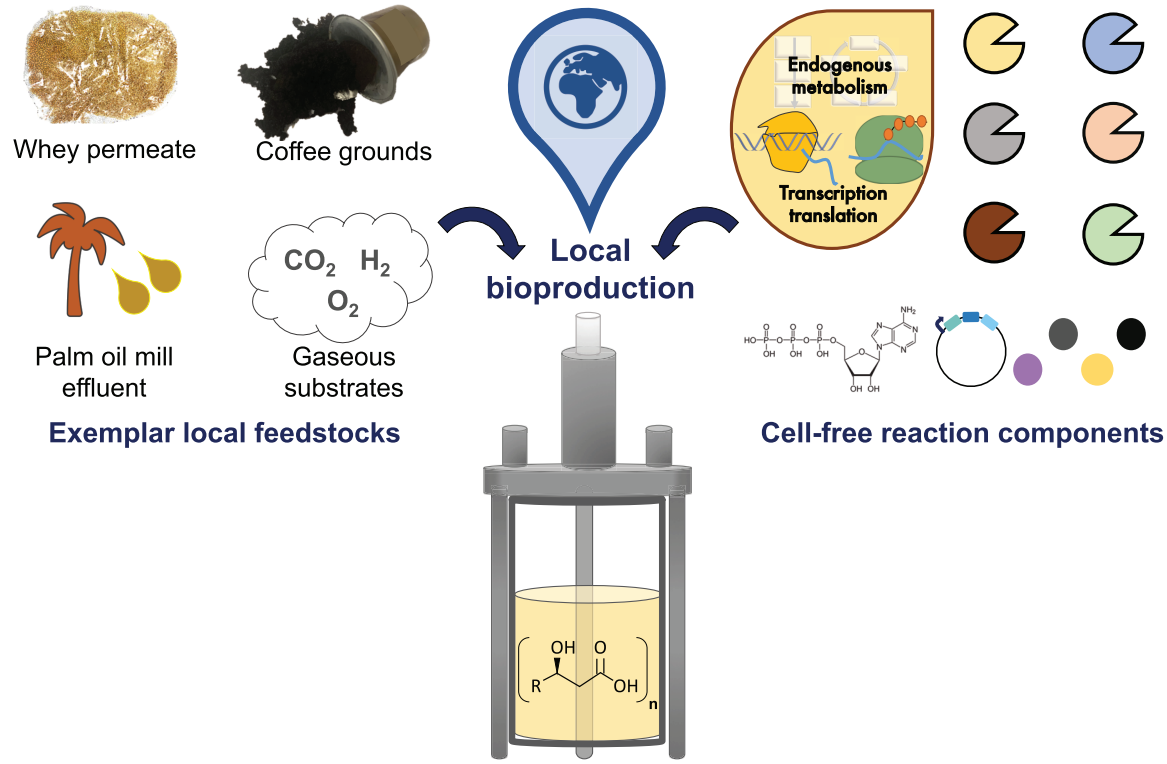


FIGURE 2 | Sustainable cell-free biomanufacturing of biological materials. Schematic depicts the local, on-demand cell-free mediated, biomanufacturing of biological materials. Local feedstocks can potentially be utilized as replacements for expensive reaction energy mix components, or to provide the enzymatic co-factors and biosynthetic pathway substrates that are required to produce biological materials of interest.

et al., 2018). Unlocking these principles can also potentially identify novel strategies to enhance *in vitro* protein production, which continues to be a core application for cell-free synthetic biology (Gagoski et al., 2016; Khambhati et al., 2019). Simplified or improved cell-free methodologies (Sun et al., 2013; Kwon and Jewett, 2015; Krinsky et al., 2016; Katsura et al., 2017; Lavickova and Maerkl, 2019), coupled with lower-cost and rationally optimized cell-free reaction energy mixes (Cai et al., 2015), has enabled the implementation of *in vitro* therapeutic protein production at industrial scales (Ogonah et al., 2017; Garenne and Noireaux, 2019). Interestingly, the integration of cell-free reactions within materials can also enhance cell-free protein synthesis yields. For instance, the Luo group has developed the protein-producing gel (P-gel) platform, which integrates protein-producing cell-free reactions within enzyme-catalyzed, 3D DNA hydrogel matrices (Park et al., 2009a). P-gels have been used to produce a panel of model proteins including green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT) and *Renilla reniformis* luciferase protein (Park et al., 2009a,b). Optimized P-gels reportedly generated between 1.87 and 5 mg/ml^{-1} of functional luciferase protein, which was significantly higher (>90-fold) than comparative, standard liquid-format cell-free reactions (Park et al., 2009a,b). However, the authors compared P-gels against linear DNA templates within standard, liquid-format cell-free reactions. In comparison to plasmid DNA, linear DNA fragments can be relatively unstable in cell-free reactions and this instability can result in relatively low cell-free protein production yields (Chappell et al., 2013; Sun et al., 2014). Nevertheless, the P-gel platform is potentially a flexible and versatile platform for cell-free protein production. Furthermore, P-gel droplets can also be rapidly fabricated using microfluidics – thus making scale-up more feasible (Ruiz et al., 2012). In a separate study, a microfluidic fabrication strategy has been developed to embed cell-free protein synthesis reactions within clay microgels that could produce >1 mg/ml of GFP protein (Jiao et al., 2018). Importantly, P-gels and clay microgels illustrate that the cell-free functionalization of materials can lead to novel approaches for protein biomanufacturing.

Combining cell-free synthetic biology with materials can also be important for other applications. For example, by combining the cell-free protein synthesis of membrane proteins with liposomes or nanodiscs can potentially facilitate the production and stable integration of membrane proteins within lipid bilayers – thus enabling structure-function studies or drug discovery applications (Shinoda et al., 2016; Gessesse et al., 2018; Shelby et al., 2019). In a separate study, the Schekman group recreated aspects of exosome biogenesis *in vitro*, by using cell-free reactions to examine exosome membrane protein

topology and exosome-associated miRNA sorting (Shurtleff et al., 2016). These cell-free approaches also enable directed membrane functionalization - potentially leading to rationally engineered protocells and designer exosomes. Combining cell-free synthetic biology with materials has led to the bottom-up engineering of biomimetics, including synthetic protocells (Shin and Noireaux, 2012; Huang et al., 2013; Garamella et al., 2016; Dubuc et al., 2019; Yue et al., 2019), entire phages (Rustad et al., 2018) and more recently, a partially self-replicating *in vitro* translation system was reported that functions by activating a 116 kb genome and is an important step toward a living, synthetic cell (Libicher et al., 2020).

CELL-FREE SYNTHETIC BIOLOGY ENABLED SMART MATERIALS

Cell-free synthetic biology reactions also enable smart materials and biosensor applications. Cell-free synthetic biology reactions can be programmed with plasmid-encoded gene circuits, lyophilized and embedded within paper (Pardee et al., 2014, 2016a; Smith and Berkheimer, 2014; Pardee, 2018), as well as potentially other materials such P-gels or clay microgels. Cell-free paper-based biosensors can be activated, post-lyophilization, using water or liquid samples and have been shown to maintain activity even after several months of storage at room temperature (Pardee et al., 2014; Smith and Berkheimer, 2014). Once activated, these paper-based cell-free reactions enable *in vitro* biosensor applications where these cell-free reactions are programmed to generate detectable signal outputs in response to the presence of relevant molecules (e.g., Mercury) or disease biomarkers (Lee and Kim, 2019; Silverman et al., 2019). Beneficially, these cell-free smart materials also provide greater flexibilities in terms of their usage beyond the laboratory and outside in the field (Pardee et al., 2016a,b). The complexity of synthetic biology genetic circuits and the strategies used to devise them has increased significantly in recent years (Nielsen et al., 2016; Grunberg and Del Vecchio, 2020). This has resulted in the development of an array of cell-free compatible genetic circuits that incorporate a variety of different regulatory elements that are applicable to different biosensing applications. For example, cell-free compatible gene networks can be transcription-based such that the presence of a small molecule induces reporter gene expression, through binding to, and activation or repression of, a transcriptional regulator. These types of transcriptional gene circuits have been used to develop cell-free biosensors for detecting heavy metals (Gräwe et al., 2019; Gupta et al., 2019), a date-rape drug (Gräwe et al., 2019), metabolites (Voyvodic et al., 2019) and quorum sensing molecules from *Pseudomonas aeruginosa*-infected respiratory samples (Wen et al., 2017). Small molecules can also regulate transcription by binding to endogenous or engineered RNA aptamer-regions within 5′ untranslated (UTR) mRNA regions – termed riboswitches. Small molecule binding to the riboswitch facilitates the emergence of stable RNA structures that permit continued reporter gene transcription. Small molecules can regulate translation via a different mechanism, whereby binding of the small molecule to the riboswitch influences the mRNA structure such that it occludes ribosome binding and downstream translation (Nahvi et al., 2002). Utilizing these principles a cell-free biosensor was developed that exploits a riboswitch to detect environmental fluoride (Thavarajah et al., 2020). Cell-free biosensors can also utilize toeholds to detect RNAs (e.g., viral RNA). Essentially, toehold aptamers function in a similar way to riboswitches except that the presence of a complementary RNA is responsible for a conformational change in the mRNA that enables ribosome access and reporter protein translation (**Figure 3**). Engineered cell-free toehold biosensors can also differentiate between different Ebola (Pardee et al., 2014) and Zika (Pardee et al., 2016a) strains. Conceivably, RNA aptamer-based approaches could be adapted to detect Coronaviruses such as SARS-CoV and 2019 n-CoV (Ahn et al., 2009). Indeed, the 2020 COVID-19 pandemic might be a catalyst for the development of cell-free viral biosensors and the accompanying clinical studies that will be required for comparative testing against existing technologies (e.g., quantitative reverse transcription polymerase chain reaction, RT-qPCR) (Liu et al., 2020). Post-translational cell-free biosensors have also been developed to detect glucose. In this example, cell-free protein synthesis was used to produce fusion protein pairs that elicit changes in Fluorescence Resonance Energy Transfer (FRET) signals in response to bound glucose (Pardee et al., 2014). Furthermore, purified fusion proteins can also enable other types of smart material applications. For instance, Wagner et al. (2019) developed information-processing materials that function by using protease-based signal-amplifying cascades, that integrate both proteolytic activities and ligand-receptor sensing within its logic circuits. As an exemplar, the same authors reported on the development of smart materials, inspired by synthetic biology logic circuits, that can detect novobiocin antibiotics (Wagner et al., 2019). These, as well as many other transcriptional and translational regulatory mechanisms can be combined into highly complex cell-free executable circuit designs (Jeong et al., 2019). Therefore, it is conceivable that these studies might inspire future efforts to embed cell-free executable logic circuits within a broad array of synthetic biology-based smart materials.

AUTOMATED DESIGN-CYCLES FOR CELL-FREE BIOLOGICAL MATERIALS

An important future trend in synthetic biology is the maturation of rational, engineering-led strategies for designing and implementing complex biological systems. Essentially, the field of synthetic biology envisions a future where model-guided, forward-engineering strategies will be routinely used to iterate design-build-test-learn cycles toward the final biotechnology production and application (Kelwick et al., 2014; Bultelle et al., 2016; Misirli et al., 2019). However, the scale of experiments needed to realize this vision may require continued advancements in synthetic biology machine-learning strategies, coupled with automation equipment (e.g., liquid handling robots) (Bultelle et al., 2016; Rajakumar et al., 2019). Indeed, automation capabilities can greatly expand the scale,

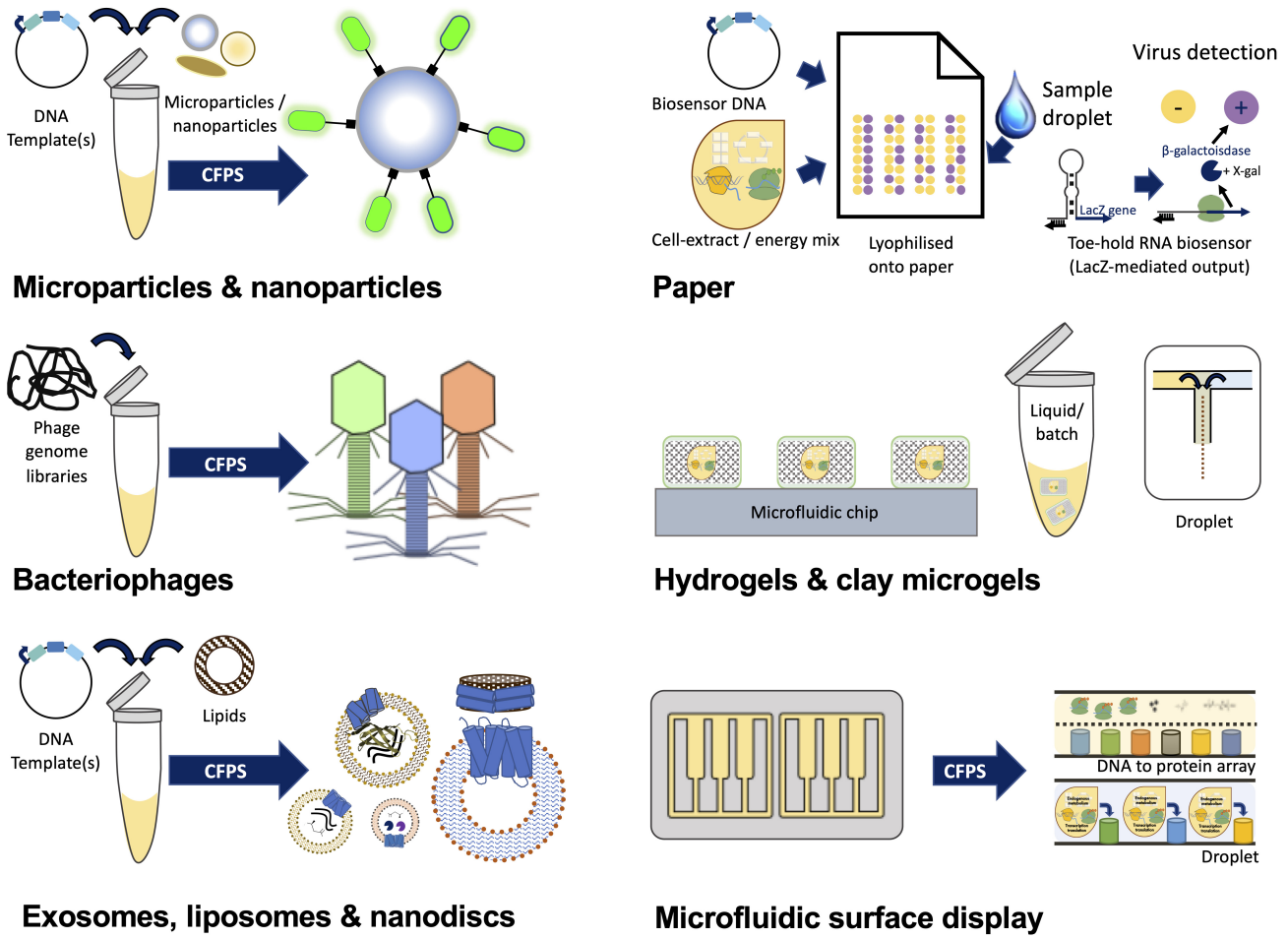


FIGURE 3 | Cell-free synthetic biology-based material functionalization. Schematic depicts examples of materials that have been bio-functionalized using cell-free protein synthesis (CFPS) reactions.

sophistication and scope of synthetic biology design cycles (Kelwick et al., 2014; Carbonell et al., 2019; El Karoui et al., 2019; Hillson et al., 2019; Jessop-Fabre and Sonnenschein, 2019). Co-ordinated efforts toward improving inter-laboratory data reproducibility, standardizing experimental metrology and an emphasis on industrially scalable biotechnology are also driving the adoption of automation in synthetic biology workflows (Kelly et al., 2009; Beal et al., 2016, 2018a,b; de Lorenzo and Schmidt, 2018; Carbonell et al., 2019; Exley et al., 2019). In a broader sense, automation is also becoming more accessible through reductions in gene synthesis costs (Carlson, 2014), advancements in automated DNA assembly protocols (Kanigowska et al., 2016; Rajakumar et al., 2019; Storch et al., 2019; Walsh et al., 2019), rapid mass spectrometry of complex biological samples (Gowers et al., 2019; Miguez et al., 2019; O'Kane et al., 2019) and through the emergence of academic biofoundries (Chambers et al., 2016; Hillson et al., 2019).

Cell-free prototyping strategies can be readily integrated into design-cycles, including those applications that are intended to be functional in living cells (*in vivo*). Several studies have described comparability between DNA regulatory elements and genetic circuits tested *in vitro* (cell-free) and *in vivo* (whole-cells) across several model (e.g., *E. coli*, *Bacillus subtilis*) (Chappell et al., 2013; Sun et al., 2014; Kelwick et al., 2016) and non-model (e.g., *Bacillus megaterium, Vibrio natriegens*) (Failmezger et al., 2018; Moore et al., 2018) organisms. Thus, cell-free rapid prototyping strategies can also be applied to speed up the development of *in vivo* (whole-cell) applications. Cell-free workflows are also amenable to automation. Indeed, several studies have successfully utilized acoustic liquid handling robots to rapidly setup large-scale, low-volume ($\leq$10 μl) prototyping cell-free reactions in a 384 well plate format (Moore et al., 2018; Kopniczky et al., 2020). Microfluidic (Swank et al., 2019), droplet array (Zhang Y. et al., 2019) or multiplex (Yim et al., 2019) strategies have also been used to enable high-throughput cell-free experiments. These approaches enable the testing of large numbers of regulatory elements or enzymes, which could potentially inform material biosynthetic pathway optimization. Conceivably, cell-free optimized material biosynthetic pathways could also be implemented within *in vivo* production strategies

(e.g., microbial cell factories). These data sets also enable accompanying quantitative modeling that can potentially predict unknown model parameters, such as transcription factor binding affinities or cell-free energy utilization (Bujara and Panke, 2012; Moore et al., 2018). These data can potentially be entered into biopart data repositories (Ham et al., 2012; Bultelle et al., 2016; McLaughlin et al., 2018) for use in machine learning-enhanced design of experiments (DoE) approaches, to speed up materials development (Liu et al., 2017; Exley et al., 2019; Madsen et al., 2019). High-throughput cell-free experiments can also tease apart where cell-free reactions are fundamentally different to native intracellular environments. Molecular crowding, metabolic fluxes, co-factor regeneration rates and other biophysiochemical characteristics may all differ between cell-free reactions and *in vivo* (whole-cell) contexts. Also, cell extract processing breaks down cellular organelles and other intracellular compartments, thus perturbing native biomolecule localization. Yet, despite these potential limitations complex genetic circuits (e.g., oscillators), that are functional *in vivo*, have been prototyped using cell-free forward-engineering design cycles (Niederholtmeyer et al., 2015). This review has also explored several other examples where cell-free approaches have been used to produce or bio-functionalize materials (**Table 1**). It is therefore conceivable, that cell-free synthetic biology design-cycles may strengthen future efforts to accelerate the development of an array of bio-functionalized smart materials (**Figure 4**).

SUMMARY AND OUTLOOK

In an era of extreme climate events there is an increasing global consensus to collectively implement more sustainable and carbon-neutral policies (Gills and Morgan, 2019). However, there continues to be an insatiable global demand for materials and the natural resources that are used to manufacture them. These activities have, however, negatively impacted the global ecosystem through for example deforestation, industrial pollution and biodiversity destruction. To partially address some of these challenges, there is increasing support for more locally based bioeconomies underpinned by sustainable practices. Cell-free biomanufacturing could have an important role in the emerging bioeconomy by enabling biological materials to be produced locally, on-demand and more sustainably from waste feedstocks (e.g., whey permeate). Furthermore, centrally produced cell-free reaction components can be lyophilized and then distributed to different geographical locations (Pardee et al., 2016b). Once distributed, these cell-free reactions could

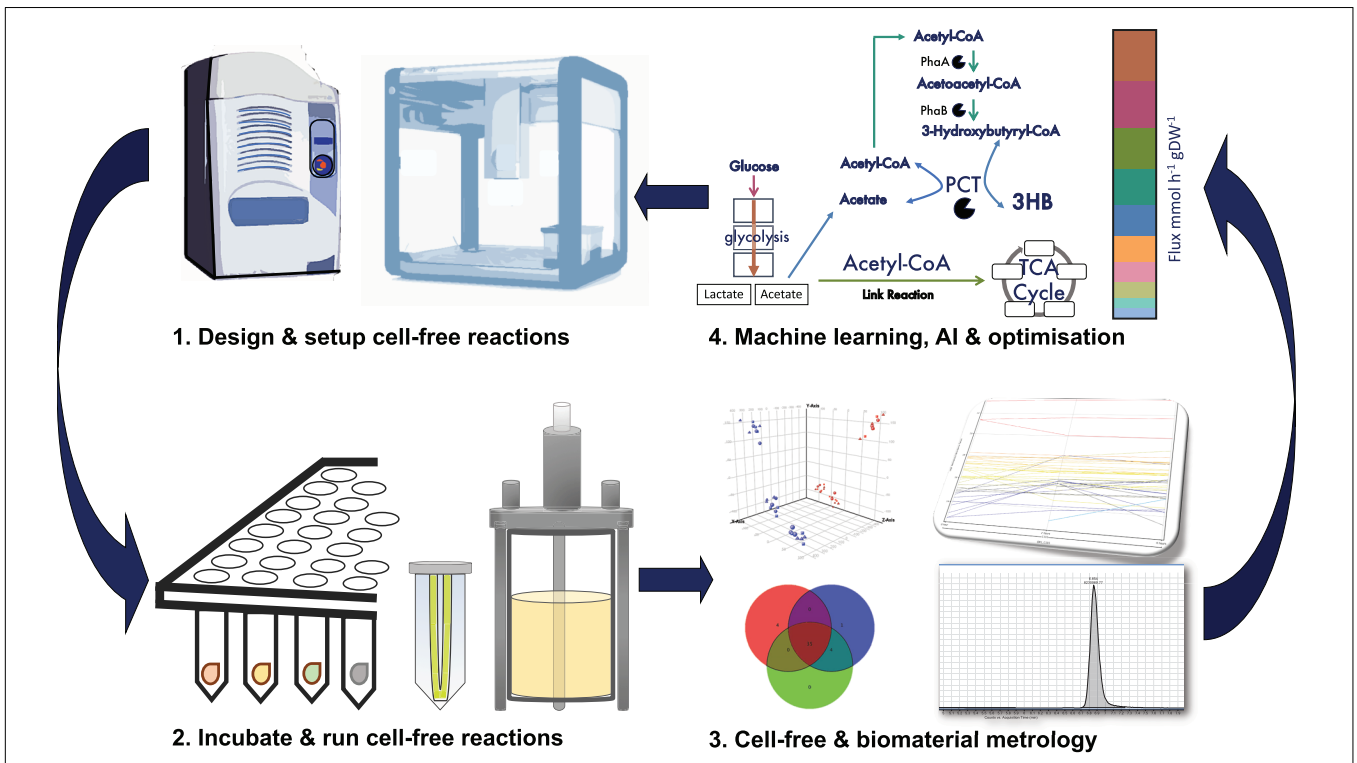


FIGURE 4 | Automated cell-free design-cycles for biological materials development. Schematic depicts a cell-free design-cycle for prototyping or biomanufacturing biological materials. Firstly, cell-free reactions are designed and then setup using liquid handling robots. Secondly, these cell-free reactions are incubated in an array of reaction formats including batch or continuous feed, at a range of scales (μl to L). Thirdly, cell-free reactions and biological materials are assayed and characterized. Finally, these data inform cell-free reaction models (e.g., metabolic flux models) or material quality control benchmarks that ultimately inform the next iteration of biosynthetic pathways and cell-free reaction parameters to be tested. Sophisticated design-cycle workflows may also utilize machine learning and design-of-experiment (DOE) approaches, to rationally iterate design cycles toward the desired biological material.

potentially make use of highly customized DNA-encoded biosynthetic operons to produce personalized materials that suit local needs. Conceivably, these cell-free produced biological materials could be fabricated using synthetic biochemistries [e.g., unnatural amino acids (Martin et al., 2018; Des Soye et al., 2019; Gao et al., 2019) and xeno nucleic acids (Glasscock et al., 2016; Hu et al., 2019)], mixed cell extracts from diverse bacterial species (Yim et al., 2019), or *de novo* biological components [e.g., engineered ribosomes (Caschera et al., 2018; d'Aquino et al., 2020) and rationally designed proteins (Huang et al., 2016; Rolf et al., 2019)]. These synthetic components might confer cell-free produced materials with unique physical characteristics or other attributes that are not typically associated with natural fibers. Likewise, we envision that future cell-free materials will have integrated smart-features including biosensing capabilities, the ability to change material properties in response to specific stimuli or self-healing capabilities when damaged. Such capabilities may require further advancements in cell-free genetic circuits and the methods used to embed them within different materials.

A panoply of recent examples, including those discussed in this review, are indicative that biological materials are the next frontier for cell-free synthetic biology – but challenges remain. A deeper understanding of the compositions and biochemical activities of processed cell extracts and cell-free reactions are needed (Foshag et al., 2018) to mitigate cell-free extract batch variability and increase protein synthesis yields. Compounding these challenges are the need for improved cell-free metrology and cell-free biomanufacturing quality control standards, which are especially important for industrial applications (de Lorenzo and Schmidt, 2018). Equally, the acceptability of cell-free produced biological materials may need to be considered through stakeholder engagement, consumer awareness activities and through, where necessary, the establishment of specialist recycling/waste management infrastructure. The scalability of cell free reactions to industrial manufacturing is also extremely challenging. Despite these challenges, Genomatica, Greenlight Biosciences, Sutro Biopharma, Tierra Biosciences and other companies have already successfully developed scalable cell-free synthetic biology platforms that have at least pilot tested materials production (e.g., Genomatica's cell-free platform produces polymer chemicals including 1,4-butanediol). In addition, the synthetic biology companies Bolt Threads, Spiber, Colorifix, and Ginkgo Bioworks are also working toward a future sustainable fashion industry. Bolt Threads produce spinnable recombinant spider silk in yeast and are also developing scalable biomanufacturing processes for mushroom-based leather. Spiber are also producing recombinant silk, as well as other synthetic protein materials. Both Bolt Threads and Spiber have successfully manufactured pilot batches of clothing products for a future sustainable fashion industry. Colorifix have developed a biological textile dye process that is designed to be less polluting than traditional chemical methods. Meanwhile, Ginkgo Bioworks have made significant investments into their organism-engineering foundry to support an array of manufacturing applications – including the production of materials. Ginkgo Bioworks also occasionally uses cell-free metabolic engineering strategies within its innovation pipelines.

In summary, cell-free synthetic biology is a powerful, highly customizable and promising biotechnology that is beginning to have a positive impact on the industrialization of sustainable materials production. Importantly, synthetic biology companies continue to champion interdisciplinary collaborations with designers and materials scientists as part of the development process. We envision that a continuation of these trends will result in a new frontier for sustainable cell-free materials production and the growing bioeconomy.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

RK was supported by a BBSRC-funded Royal Society of Edinburgh Enterprise Fellowship. AW was supported by the UK Government's Global Challenges Research Fund (GCRF) through the EPSRC grant (EP/P028519/1) as part of the WISER project. We acknowledge the support of Imperial Confidence in Concept (MRC/EPSRC), Imperial College London EPSRC Impact Acceleration Account (EP/R511547/1), EPSRC grant (EP/L011573/1), BBSRC grant (BB/L027852/1), and the Imperial College London CRUK Development Fund.

ACKNOWLEDGMENTS

We thank colleagues in the Section of Structural and Synthetic Biology, WISER project colleagues and Dr. Amélie Heliot.

REFERENCES

Ahn, D.-G., Jeon, I.-J., Kim, J. D., Song, M.-S., Han, S.-R., Lee, S.-W., et al. (2009). RNA aptamer-based sensitive detection of SARS coronavirus nucleocapsid protein. *Analyst* 134:1896. doi: 10.1039/b906788d

Ahn, W. S., Park, S. J., and Lee, S. Y. (2000). Production of Poly(3-hydroxybutyrate) by fed-batch culture of recombinant *Escherichia coli* with a highly concentrated whey solution. *Appl. Environ. Microbiol.* 66, 3624–3627. doi: 10.1128/aem.66.8.3624-3627.2000

Anderson, D. A., Jones, R. D., Arkin, A. P., and Weiss, R. (2019). Principles of synthetic biology: a MOOC for an emerging field. *Synth. Biol.* 4:ysz010. doi: 10.1093/synbio/ysz010

Arpino, J. A. J., Hancock, E. J., Anderson, J., Barahona, M., Stan, G.-B. V., Papachristodoulou, A., et al. (2013). Tuning the dials of synthetic biology. *Microbiology* 159, 1236–1253. doi: 10.1099/mic.0.067975-0

Ausländer, S., Ausländer, D., and Fussenegger, M. (2017). synthetic biology-the synthesis of biology. *Angew. Chemie Int. Ed.* 56, 6396–6419. doi: 10.1002/anie.201609229

Aw, R., and Polizzi, K. M. (2019). Biosensor-assisted engineering of a high-yield Pichia pastoris cell-free protein synthesis platform. *Biotechnol. Bioeng.* 116, 656–666. doi: 10.1002/bit.26901

Beal, J., Haddock-Angelli, T., Baldwin, G., Gershater, M., Dwijayanti, A., Storch, M., et al. (2018a). Quantification of bacterial fluorescence using independent calibrants. *PLoS One* 13:e0199432. doi: 10.1371/journal.pone.0199432

Beal, J., Haddock-Angelli, T., Farny, N., and Rettberg, R. (2018b). Time to get serious about measurement in synthetic biology. *Trends Biotechnol.* 36, 869–871. doi: 10.1016/j.tibtech.2018.05.003

Beal, J., Haddock-Angelli, T., Gershater, M., de Mora, K., Lizarazo, M., Hollenhorst, J., et al. (2016). Reproducibility of fluorescent expression from engineered biological constructs in *E. coli*. *PLoS One* 11:e0150182. doi: 10.1371/journal.pone.0150182

Beckers, V., Poblete-Castro, I., Tomasch, J., and Wittmann, C. (2016). Integrated analysis of gene expression and metabolic fluxes in PHA-producing *Pseudomonas* putida grown on glycerol. *Microb. Cell Fact.* 15:73. doi: 10.1186/s12934-016-0470-2

Benítez-Mateos, A. I., Llarena, I., Sánchez-Iglesias, A., and López-Gallego, F. (2018). Expanding one-pot cell-free protein synthesis and immobilization for on-demand manufacturing of biomaterials. *ACS Synth. Biol.* 7, 875–884. doi: 10.1021/acssynbio.7b00383

Borkowski, O., Bricio, C., Murgiano, M., Rothschild-Mancinelli, B., Stan, G.-B., and Ellis, T. (2018). Cell-free prediction of protein expression costs for growing cells. *Nat. Commun.* 9:1457. doi: 10.1038/s41467-018-03970-x

Bowie, J. U., Sherkhanov, S., Korman, T. P., Valliere, M. A., Opgenorth, P. H., and Liu, H. (2020). Synthetic biochemistry: the bio-inspired cell-free approach to commodity chemical production. *Trends Biotechnol.* doi: 10.1016/j.tibtech.2019.12.024

Bujara, M., and Panke, S. (2012). In silico assessment of cell-free systems. *Biotechnol. Bioeng.* 109, 2620–2629. doi: 10.1002/bit.24534

Bultelle, M., de Murieta, I. S., and Kitney, R. (2016). "Introducing SynBIS - the synthetic biology information system," in *Proceedings of the IET/SynbiCITE Engineering Biology Conference*, Michael Faraday House.

Cai, Q., Hanson, J. A., Steiner, A. R., Tran, C., Masikat, M. R., Chen, R., et al. (2015). A simplified and robust protocol for immunoglobulin expression in *E scherichia coli* cell-free protein synthesis systems. *Biotechnol. Prog.* 31, 823–831. doi: 10.1002/btpr.2082

Callén Moreu, B., and López Gómez, D. (2019). Intimate with your junk! A waste management experiment for a material world. *Sociol. Rev.* 67, 318–339. doi: 10.1177/0038026119830318

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

Carbonell, P., Radivojevic, T., and García Martín, H. (2019). Opportunities at the intersection of synthetic biology, machine learning, and automation. *ACS Synth. Biol.* 8, 1474–1477. doi: 10.1021/acssynbio.8b00540

Carlson, R. (2014). *Time for New DNA Sequencing And Synthesis Cost Curves*. Available online at: https://synbiobeta.com/time-new-dna-synthesis-sequencing-cost-curves-rob-carlson/ (accessed February 1, 2020).

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. doi: 10.1021/acssynbio.8b00276

Caschera, F., and Noireaux, V. (2014). Synthesis of 2.3 mg/ml of protein with an all *Escherichia coli* cell-free transcription–translation system. *Biochimie* 99, 162–168. doi: 10.1016/j.biochi.2013.11.025

Chambers, S., Kitney, R., and Freemont, P. (2016). The foundry: the DNA synthesis and construction foundry at imperial college. *Biochem. Soc. Trans.* 44, 687–688. doi: 10.1042/BST20160007

Chappell, J., Jensen, K., and Freemont, P. S. (2013). Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res.* 41, 3471–3481. doi: 10.1093/nar/gkt052

Chauhan, A., Zubair, S., Tufail, S., Sherwani, A., Sajid, M., Raman, S. C., et al. (2011). Fungus-mediated biological synthesis of gold nanoparticles: potential in detection of liver cancer. *Int. J. Nanomedicine* 2305:195. doi: 10.2147/IJN.S23195

Chen, G.-Q., Chen, X.-Y., Wu, F.-Q., and Chen, J.-C. (2020). Polyhydroxyalkanoates (PHA) toward cost competitiveness and functionality. *Adv. Ind. Eng. Polym. Res.* 3, 1–7. doi: 10.1016/j.aiepr.2019.11.001

Chen, G.-Q., and Hajnal, I. (2015). The 'PHAome.'. *Trends Biotechnol.* 33, 559–564. doi: 10.1016/j.tibtech.2015.07.006

Choi, K. R., Jang, W. D., Yang, D., Cho, J. S., Park, D., and Lee, S. Y. (2019). Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* 37, 817–837. doi: 10.1016/j.tibtech.2019.01.003

Choi, S. Y., Rhie, M. N., Kim, H. T., Joo, J. C., Cho, I. J., Son, J., et al. (2019). Metabolic engineering for the synthesis of polyesters: a 100-year journey from polyhydroxyalkanoates to non-natural microbial polyesters. *Metab. Eng.* 58, 47–81. doi: 10.1016/j.ymben.2019.05.009

Church, G. M., Elowitz, M. B., Smolke, C. D., Voigt, C. A., and Weiss, R. (2014). Realizing the potential of synthetic biology. *Nat. Rev. Mol. Cell Biol.* 15, 289–294. doi: 10.1038/nrm3767

Contreras-Llano, L. E., and Tan, C. (2018). High-throughput screening of biomolecules using cell-free gene expression systems. *Synth. Biol.* 3, 1–13. doi: 10.1093/synbio/ysy012

Costa Silva, L. P., Pinto Oliveira, J., Keijok, W. J., da Silva, A. R., Aguiar, A. R., Guimarães, M. C. C., et al. (2017). Extracellular biosynthesis of silver nanoparticles using the cell-free filtrate of nematophagous fungus duddingtonia flagrans. *Int. J. Nanomedicine Volume* 12, 6373–6381. doi: 10.2147/IJN.S137703

Crowther, T. W., Glick, H. B., Covey, K. R., Bettigole, C., Maynard, D. S., Thomas, S. M., et al. (2015). Mapping tree density at a global scale. *Nature* 525, 201–205. doi: 10.1038/nature14967

Cui, Y.-W., Zhang, H.-Y., Lu, P.-F., and Peng, Y.-Z. (2016). Effects of carbon sources on the enrichment of halophilic polyhydroxyalkanoate-storing mixed microbial culture in an aerobic dynamic feeding process. *Sci. Rep.* 6:30766. doi: 10.1038/srep30766

d'Aquino, A. E., Azim, T., Aleksashin, N. A., Hockenberry, A. J., Krüger, A., and Jewett, M. C. (2020). Mutational characterization and mapping of the 70S ribosome active site. *Nucleic Acids Res.* 48, 2777–2789. doi: 10.1093/nar/gkaa001

de Lorenzo, V., and Schmidt, M. (2018). Biological standards for the knowledge-based bioeconomy: what is at stake. *N. Biotechnol.* 40, 170–180. doi: 10.1016/j.nbt.2017.05.001

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.e–1754.e. doi: 10.1016/j.chembiol.2019.10.008

Dopp, J. L., Jo, Y. R., and Reuel, N. F. (2019). Methods to reduce variability in *E. Coli*-based cell-free protein expression experiments. *Synth. Syst. Biotechnol.* 4, 204–211. doi: 10.1016/j.synbio.2019.10.003

Dubuc, E., Pieters, P. A., van der Linden, A. J., van Hest, J. C., Huck, W. T., and de Greef, T. F. (2019). Cell-free microcompartmentalised transcription–translation for the prototyping of synthetic communication networks. *Curr. Opin. Biotechnol.* 58, 72–80. doi: 10.1016/j.copbio.2018.10.006

Dudley, Q. M., Karim, A. S., and Jewett, M. C. (2015). Cell-free metabolic engineering: biomanufacturing beyond the cell. *Biotechnol. J.* 10, 69–82. doi: 10.1002/biot.201400330

Duyen, T. T. M., Matsuura, H., Ujiie, K., Muraoka, M., Harada, K., and Hirata, K. (2017). Paper-based colorimetric biosensor for antibiotics inhibiting bacterial protein synthesis. *J. Biosci. Bioeng.* 123, 96–100. doi: 10.1016/j.jbiosc.2016.07.015

Dvořák, P., Nikel, P. I., Damborský, J., and de Lorenzo, V. (2017). Bioremediation 3. 0: Engineering pollutant-removing bacteria in the times of systemic biology. *Biotechnol. Adv.* 35, 845–866. doi: 10.1016/j.biotechadv.2017.08.001

Dy, A. J., Aurand, E. R., and Friedman, D. C. (2019). YouTube resources for synthetic biology education. *Synth. Biol.* 4:ysz022. doi: 10.1093/synbio/ysz022

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

Endoh, T., Kanai, T., Sato, Y. T., Liu, D. V., Yoshikawa, K., Atomi, H., et al. (2006). Cell-free protein synthesis at high temperatures using the lysate of a hyperthermophile. *J. Biotechnol.* 126, 186–195. doi: 10.1016/j.jbiotec.2006.04.010

Eriksen, M., Lebreton, L. C. M., Carson, H. S., Thiel, M., Moore, C. J., Borerro, J. C., et al. (2014). plastic pollution in the world's oceans: more than 5 Trillion plastic pieces weighing over 250,000 tons afloat at Sea. *PLoS One* 9:111913. doi: 10.1371/journal.pone.0111913

Exley, K., Reynolds, C. R., Suckling, L., Chee, S. M., Tsipa, A., Freemont, P. S., et al. (2019). Utilising datasheets for the informed automated design and build of a synthetic metabolic pathway. *J. Biol. Eng.* 13:8. doi: 10.1186/s13036-019-0141-z

Ezure, T., Suzuki, T., Higashide, S., Shintani, E., Endo, K., Kobayashi, S.-I., et al. (2006). Cell-free protein synthesis system prepared from insect cells by freeze-thawing. *Biotechnol. Prog.* 22, 1570–1577. doi: 10.1021/bp060110v

Failmezger, J., Scholz, S., Blombach, B., and Siemann-Herzberg, M. (2018). Cell-free protein synthesis from fast-growing vibrio natriegens. *Front. Microbiol.* 9:1146. doi: 10.3389/fmicb.2018.01146

Foshag, D., Henrich, E., Hiller, E., Schäfer, M., Kerger, C., Burger-Kentischer, A., et al. (2018). The *E. coli S*30 lysate proteome: a prototype for cell-free protein production. *N. Biotechnol.* 40, 245–260. doi: 10.1016/j.nbt.2017.09.005

Freemont, P. S. (2019). Synthetic biology industry: data-driven design is creating new opportunities in biotechnology. *Emerg. Top. Life Sci.* 3, 651–657. doi: 10.1042/ETLS20190040

French, K. E. (2019). Harnessing synthetic biology for sustainable development. *Nat. Sustain.* 2, 250–252. doi: 10.1038/s41893-019-0270-x

Gagoski, D., Polinkovsky, M. E., Mureev, S., Kunert, A., Johnston, W., Gambin, Y., et al. (2016). Performance benchmarking of four cell-free protein expression systems. *Biotechnol. Bioeng.* 113, 292–300. doi: 10.1002/bit.25814

Gao, W., Cho, E., Liu, Y., and Lu, Y. (2019). Advances and challenges in cell-free incorporation of unnatural amino acids into proteins. *Front. Pharmacol.* 10:611. doi: 10.3389/fphar.2019.00611

Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016). The all *E. coli* TX-TL toolbox 2.0: a platform for cell-free synthetic biology. *ACS Synth. Biol.* 5, 344–355. doi: 10.1021/acssynbio.5b00296

García-Manrique, P., Matos, M., Gutiérrez, G., Pazos, C., and Blanco-López, M. C. (2018). Therapeutic biomaterials based on extracellular vesicles: classification of bio-engineering and mimetic preparation routes. *J. Extracell. Vesicles* 7:1422676. doi: 10.1080/20013078.2017.1422676

Garenne, D., and Noireaux, V. (2019). Cell-free transcription–translation: engineering biology from the nanometer to the millimeter scale. *Curr. Opin. Biotechnol.* 58, 19–27. doi: 10.1016/j.copbio.2018.10.007

Georgi, V., Georgi, L., Blechert, M., Bergmeister, M., Zwanzig, M., Wüstenhagen, D. A., et al. (2016). On-chip automation of cell-free protein synthesis: new opportunities due to a novel reaction mode. *Lab. Chip.* 16, 269–281. doi: 10.1039/C5LC00700C

Gessesse, B., Nagaike, T., Nagata, K., Shimizu, Y., and Ueda, T. (2018). G-protein coupled receptor protein synthesis on a lipid bilayer using a reconstituted cell-free protein synthesis system. *Life* 8:54. doi: 10.3390/life8040054

Gilbert, C., and Ellis, T. (2019). Biological engineered living materials: growing functional materials with genetically programmable properties. *ACS Synth. Biol.* 8, 1–15. doi: 10.1021/acssynbio.8b00423

Gills, B., and Morgan, J. (2019). Global climate emergency: after COP24, climate science, urgency, and the threat to humanity. *Globalizations* 1–18. doi: 10.1080/14747731.2019.1669915

Glasscock, C. J., Lucks, J. B., and DeLisa, M. P. (2016). Engineered protein machines: emergent tools for synthetic biology. *Cell Chem. Biol.* 23, 45–56. doi: 10.1016/j.chembiol.2015.12.004

Gowers, G.-O. F., Cameron, S. J. S., Perdones-Montero, A., Bell, D., Chee, S. M., Kern, M., et al. (2019). Off-colony screening of biosynthetic libraries by rapid laser-enabled mass spectrometry. *ACS Synth. Biol.* 8, 2566–2575. doi: 10.1021/acssynbio.9b00243

Gräwe, A., Dreyer, A., Vornholt, T., Barteczko, U., Buchholz, L., Drews, G., et al. (2019). A paper-based, cell-free biosensor system for the detection of heavy metals and date rape drugs. *PLoS One* 14:e0210940. doi: 10.1371/journal.pone.0210940

Greene, R. A., Morgan, M., Shatkin, A. J., and Gage, L. P. (1975). Translation of silk fibroin messenger RNA in an Ehrlich ascites cell-free extract. *J. Biol. Chem.* 250, 5114–5121.

Gregorio, N. E., Levine, M. Z., and Oza, J. P. (2019). A User's guide to cell-free protein synthesis. *Methods Protoc.* 2:24. doi: 10.3390/mps2010024

Grunberg, T. W., and Del Vecchio, D. (2020). Modular analysis and design of biological circuits. *Curr. Opin. Biotechnol.* 63, 41–47. doi: 10.1016/j.copbio.2019.11.015

Gupta, S., Sarkar, S., Katranidis, A., and Bhattacharya, J. (2019). Development of a cell-free optical biosensor for detection of a broad range of mercury contaminants in water: a plasmid DNA-Based approach. *ACS Omega* 4, 9480–9487. doi: 10.1021/acsomega.9b00205

Ham, T. S., Dmytriv, Z., Plahar, H., Chen, J., Hillson, N. J., and Keasling, J. D. (2012). Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. *Nucleic Acids Res.* 40:e141. doi: 10.1093/nar/gks531

Han, X., Satoh, Y., Satoh, T., Matsumoto, K., Kakuchi, T., Taguchi, S., et al. (2011). Chemo-enzymatic synthesis of polyhydroxyalkanoate (PHA) incorporating 2-hydroxybutyrate by wild-type class i PHA synthase from *Ralstonia eutropha*. *Appl. Microbiol. Biotechnol.* 92, 509–517. doi: 10.1007/s00253-011-3362-8

Harbers, M. (2014). Wheat germ systems for cell-free protein expression. *FEBS Lett.* 588, 2762–2773. doi: 10.1016/j.febslet.2014.05.061

Hecht, A., Glasgow, J., Jaschke, P. R., Bawazer, L. A., Munson, M. S., Cochran, J. R., et al. (2017). Measurements of translation initiation from all 64 codons in *E. coli*. *Nucleic Acids Res.* 45, 3615–3626. doi: 10.1093/nar/gkx070

Hicks, M., Bachmann, T. T., and Wang, B. (2020). Synthetic biology enables programmable cell-based biosensors. *Chemphyschem* 21, 132–144. doi: 10.1002/cphc.201900739

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

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

Hodgman, C. E., and Jewett, M. C. (2013). Optimized extract preparation methods and reaction conditions for improved yeast cell-free protein synthesis. *Biotechnol. Bioeng.* 110, 2643–2654. doi: 10.1002/bit.24942

Hu, J., Xiao, K., Jin, B., Zheng, X., Ji, F., and Bai, D. (2019). Paper-based point-of-care test with xeno nucleic acid probes. *Biotechnol. Bioeng.* 116, 2764–2777. doi: 10.1002/bit.27106

Hu, S. J. (2013). Evolving paradigms of manufacturing: from mass production to mass customization and personalization. *Procedia CIRP* 7, 3–8. doi: 10.1016/j.procir.2013.05.002

Huang, A., Nguyen, P. Q., Stark, J. C., Takahashi, M. K., Donghia, N., Ferrante, T., et al. (2018). BiobitsTM explorer: a modular synthetic biology education kit. *Sci. Adv.* 4, 1–10. doi: 10.1126/sciadv.aat5105

Huang, P.-S., Boyken, S. E., and Baker, D. (2016). The coming of age of de novo protein design. *Nature* 537, 320–327. doi: 10.1038/nature19946

Huang, X., Li, M., Green, D. C., Williams, D. S., Patil, A. J., and Mann, S. (2013). Interfacial assembly of protein–polymer nano-conjugates into stimulus-responsive biomimetic protocells. *Nat. Commun.* 4:2239. doi: 10.1038/ncomms3239

Jaworski, E., Wang, L., and Marco, G. (1963). Synthesis of chitin in cell-free extracts of prodenia eridania. *Nature* 198, 790–790. doi: 10.1038/198790a0

Jeong, D., Klocke, M., Agarwal, S., Kim, J., Choi, S., Franco, E., et al. (2019). Cell-free synthetic biology platform for engineering synthetic biological circuits and systems. *Methods Protoc.* 2:39. doi: 10.3390/mps2020039

Jessop-Fabre, M. M., and Sonnenschein, N. (2019). Improving reproducibility in synthetic biology. *Front. Bioeng. Biotechnol.* 7:18. doi: 10.3389/fbioe.2019.00018

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. doi: 10.1016/j.synbio.2018.02.003

Jiao, Y., Liu, Y., Luo, D., Huck, W. T. S., and Yang, D. (2018). Microfluidic-assisted fabrication of clay microgels for cell-free protein synthesis. *ACS Appl. Mater. Interfaces* 10, 29308–29313. doi: 10.1021/acsami.8b09324

Kanigowska, P., Shen, Y., Zheng, Y., Rosser, S., and Cai, Y. (2016). Smart DNA fabrication using sound waves. *J. Lab. Autom.* 21, 49–56. doi: 10.1177/2211068215593754

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. doi: 10.1016/j.ymben.2016.03.002

Karim, A. S., and Jewett, M. C. (2018). Cell-free synthetic biology for pathway prototyping. *Methods Enzymol.* 608, 31–57. doi: 10.1016/bs.mie.2018.04.029

Katsura, K., Matsuda, T., Tomabechi, Y., Yonemochi, M., Hanada, K., Ohsawa, N., et al. (2017). A reproducible and scalable procedure for preparing bacterial

extracts for cell-free protein synthesis. *J. Biochem.* 162, 357–369. doi: 10.1093/jb/mvx039

Kay, J. E., and Jewett, M. C. (2020). A cell-free system for production of 2,3-butanediol is robust to growth-toxic compounds. *Metab. Eng. Commun.* 10:e00114. doi: 10.1016/j.mec.2019.e00114

Kelly, J. R., Rubin, A. J., Davis, J. H., Ajo-Franklin, C. M., Cumbers, J., Czar, M. J., et al. (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.* 3:4. doi: 10.1186/1754-1611-3-4

Kelwick, R., Bowater, L., Yeoman, K. H., and Bowater, R. P. (2015a). Promoting microbiology education through the iGEM synthetic biology competition. *FEMS Microbiol. Lett.* 362:fnv129. doi: 10.1093/femsle/fnv129

Kelwick, R., Kopniczky, M., Bower, I., Chi, W., Chin, M. H. W., Fan, S., et al. (2015b). A Forward-design approach to increase the production of poly-3-hydroxybutyrate in genetically engineered *Escherichia coli*. *PLoS One* 10:e0117202. doi: 10.1371/journal.pone.0117202

Kelwick, R., MacDonald, J. T., Webb, A. J., and Freemont, P. (2014). Developments in the tools and methodologies of synthetic biology. *Front. Bioeng. Biotechnol.* 2:60. doi: 10.3389/fbioe.2014.00060

Kelwick, R., Ricci, L., Chee, S. M., Bell, D., Webb, A. J., and Freemont, P. S. (2018). Cell-free prototyping strategies for enhancing the sustainable production of polyhydroxyalkanoates bioplastics. *Synth. Biol.* 3:ysy016. doi: 10.1093/synbio/ysy016

Kelwick, R., Webb, A. J., MacDonald, J. T., and Freemont, P. S. (2016). Development of a *Bacillus subtilis* cell-free transcription-translation system for prototyping regulatory elements. *Metab. Eng.* 38, 370–381. doi: 10.1016/j.ymben.2016.09.008

Khambhati, K., Bhattacharjee, G., Gohil, N., Braddick, D., Kulkarni, V., and Singh, V. (2019). Exploring the potential of cell-free protein synthesis for extending the abilities of biological systems. *Front. Bioeng. Biotechnol.* 7:248. doi: 10.3389/fbioe.2019.00248

Kightlinger, W., Duncker, K. E., Ramesh, A., Thames, A. H., Natarajan, A., Stark, J. C., et al. (2019). A cell-free biosynthesis platform for modular construction of protein glycosylation pathways. *Nat. Commun.* 10, 1–13. doi: 10.1038/s41467-019-12024-9

Kilb, N., Burger, J., and Roth, G. (2014). Protein microarray generation by in situ protein expression from template DNA. *Eng. Life Sci.* 14, 352–364. doi: 10.1002/elsc.201300052

Kim, B. S. (2000). Production of poly(3-hydroxybutyrate) from inexpensive substrates. *Enzyme Microb. Technol.* 27, 774–777. doi: 10.1016/s0141-0229(00)00299-4

Kleer, R., and Piller, F. T. (2019). Local manufacturing and structural shifts in competition: market dynamics of additive manufacturing. *Int. J. Prod. Econ.* 216, 23–34. doi: 10.1016/j.ijpe.2019.04.019

Kopniczky, M. B., Canavan, C., McClymont, D. W., Crone, M. A., Suckling, L., Goetzmann, B., et al. (2020). Cell-free protein synthesis as a prototyping platform for mammalian synthetic biology. *ACS Synth. Biol.* 9, 144–156. doi: 10.1021/acssynbio.9b00437

Kopp, D., Willows, R. D., and Sunna, A. (2019). Cell-free enzymatic conversion of spent coffee grounds into the platform chemical lactic acid. *Front. Bioeng. Biotechnol.* 7:389. doi: 10.3389/fbioe.2019.00389

Korman, T. P., Opgenorth, P. H., and Bowie, J. U. (2017). A synthetic biochemistry platform for cell free production of monoterpenes from glucose. *Nat. Commun.* 8:15526. doi: 10.1038/ncomms15526

Kovtun, O., Mureev, S., Johnston, W., and Alexandrov, K. (2010). Towards the construction of expressed proteomes using a *Leishmania tarentolae* based cell-free expression system. *PLoS One* 5:e14388. doi: 10.1371/journal.pone.0014388

Kovtun, O., Mureev, S., Jung, W., Kubala, M. H., Johnston, W., and Alexandrov, K. (2011). Leishmania cell-free protein expression system. *Methods* 55, 58–64. doi: 10.1016/j.ymeth.2011.06.006

Krinsky, N., Kaduri, M., Shainsky-Roitman, J., Goldfeder, M., Ivanir, E., Benhar, I., et al. (2016). A simple and rapid method for preparing a cell-free bacterial lysate for protein synthesis. *PLoS One* 11:e0165137. doi: 10.1371/journal.pone.0165137

Krishnan, S., Narayan, S., and Chadha, A. (2016). Whole resting cells vs. cell free extracts of *Candida parapsilosis* ATCC 7330 for the synthesis of gold nanoparticles. *AMB Express* 6:92. doi: 10.1186/s13568-016-0268-y

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. doi: 10.1038/srep08663

Lai, H.-E., Canavan, C., Cameron, L., Moore, S., Danchenko, M., Kuiken, T., et al. (2019). Synthetic biology and the united nations. *Trends Biotechnol.* 37, 1146–1151. doi: 10.1016/j.tibtech.2019.05.011

Lavickova, B., and Maerkl, S. J. (2019). A simple, robust, and low-cost method to produce the pure cell-free system. *ACS Synth. Biol.* 8, 455–462. doi: 10.1021/acssynbio.8b00427

Le Feuvre, R. A., and Scrutton, N. S. (2018). A living foundry for synthetic biological materials: a synthetic biology roadmap to new advanced materials. *Synth. Syst. Biotechnol.* 3, 105–112. doi: 10.1016/j.synbio.2018.04.002

Lee, K.-H., and Kim, D.-M. (2019). In vitro use of cellular synthetic machinery for biosensing applications. *Front. Pharmacol.* 10:1166. doi: 10.3389/fphar.2019.01166

Lenton, T. M., Rockström, J., Gaffney, O., Rahmstorf, S., Richardson, K., Steffen, W., et al. (2019). Climate tipping points — too risky to bet against. *Nature* 575, 592–595. doi: 10.1038/d41586-019-03595-0

Li, J., Wang, H., and Jewett, M. C. (2018). Expanding the palette of Streptomyces -based cell-free protein synthesis systems with enhanced yields. *Biochem. Eng. J.* 130, 29–33. doi: 10.1016/j.bej.2017.11.013

Li, T., Ye, J., Shen, R., Zong, Y., Zhao, X., Lou, C., et al. (2016). Semirational approach for ultrahigh Poly(3-hydroxybutyrate) accumulation in *Escherichia coli* by combining one-step library construction and high-throughput screening. *ACS Synth. Biol.* 5, 1308–1317. doi: 10.1021/acssynbio.6b00083

Libicher, K., Hornberger, R., Heymann, M., and Mutschler, H. (2020). In vitro self-replication and multicistronic expression of large synthetic genomes. *Nat. Commun.* 11:904. doi: 10.1038/s41467-020-14694-2

Lim, S. Y., Kim, K.-O., Kim, D.-M., and Park, C. B. (2009). Silica-coated alginate beads for in vitro protein synthesis via transcription/translation machinery encapsulation. *J. Biotechnol.* 143, 183–189. doi: 10.1016/j.jbiotec.2009.07.006

Liu, R., Fu, A., Deng, Z., Li, Y., and Liu, T. (2020). Promising methods for detection of novel coronavirus SARS-CoV-2. *View* 1:e4. doi: 10.1002/viw2.4

Liu, Y., Zhao, T., Ju, W., Shi, S., Shi, S., and Shi, S. (2017). Materials discovery and design using machine learning. *J. Mater.* 3, 159–177. doi: 10.1016/j.jmat.2017.08.002

Lizardi, P. M., Mahdavi, V., Shields, D., and Candelas, G. (1979). Discontinuous translation of silk fibroin in a reticulocyte cell-free system and in intact silk gland cells. *Proc. Natl. Acad. Sci.* 76, 6211–6215. doi: 10.1073/pnas.76.12.6211

Madsen, C., Goñi Moreno, A., Umesh, P., Palchick, Z., Roehner, N., Atallah, C., et al. (2019). Synthetic biology open language (SBOL) Version 2.3. *J. Integr. Bioinform.* 16:25. doi: 10.1515/jib-2019-0025

Martin, R. W., Des Soye, B. J., Kwon, Y.-C., Kay, J., Davis, R. G., Thomas, P. M., et al. (2018). Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat. Commun.* 9:1203. doi: 10.1038/s41467-018-03469-5

Martin, R. W., Majewska, N. I., Chen, C. X., Albanetti, T. E., Jimenez, R. B. C., Schmelzer, A. E., et al. (2017). Development of a CHO-based cell-free platform for synthesis of active monoclonal antibodies. *ACS Synth. Biol.* 6, 1370–1379. doi: 10.1021/acssynbio.7b00001

Matsumoto, K. K., Okei, T., Honma, I., Ooi, T., Aoki, H., and Taguchi, S. (2013). Erratum to: Efficient (R)-3-hydroxybutyrate production using acetyl CoA-regenerating pathway catalyzed by coenzyme A transferase. *Appl. Microbiol. Biotechnol.* 97, 439–439. doi: 10.1007/s00253-012-4563-5

McLaughlin, J. A., Myers, C. J., Zundel, Z., Mısırlı, G., Zhang, M., Ofiteru, I. D., et al. (2018). SynBioHub: a standards-enabled design repository for synthetic biology. *ACS Synth. Biol.* 7, 682–688. doi: 10.1021/acssynbio.7b00403

Miguez, A. M., McNerney, M. P., and Styczynski, M. P. (2019). Metabolic profiling of *Escherichia coli* -based cell-free expression systems for process optimization. *Ind. Eng. Chem. Res.* 58, 22472–22482. doi: 10.1021/acs.iecr.9b03565

Misirli, G., Nguyen, T., McLaughlin, J. A., Vaidyanathan, P., Jones, T. S., Densmore, D., et al. (2019). A computational workflow for the automated generation of models of genetic designs. *ACS Synth. Biol.* 8, 1548–1559. doi: 10.1021/acssynbio.7b00459

Moore, S. J., Lai, H.-E., Needham, H., Polizzi, K. M., and Freemont, P. S. (2017a). *Streptomyces venezuelae* TX-TL - a next generation cell-free synthetic biology tool. *Biotechnol. J.* 12:1600678. doi: 10.1002/biot.201600678

Moore, S. J., MacDonald, J. T., and Freemont, P. S. (2017b). Cell-free synthetic biology for in vitro prototype engineering. *Biochem. Soc. Trans.* 45, 785–791. doi: 10.1042/BST20170011

Moore, S. J., MacDonald, J. T., Wienecke, S., Ishwarbhai, A., Tsipa, A., Aw, R., 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. doi: 10.1073/pnas.1715806115

Mureev, S., Kovtun, O., Nguyen, U. T. T., and Alexandrov, K. (2009). Species-independent translational leaders facilitate cell-free expression. *Nat. Biotechnol.* 27, 747–752. doi: 10.1038/nbt.1556

Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., and Breaker, R. R. (2002). Genetic control by a metabolite binding mRNA. *Chem. Biol.* 9, 1043–1049. doi: 10.1016/S1074-5521(02)00224-7

Niederholtmeyer, H., Sun, Z. Z., Hori, Y., Yeung, E., Verpoorte, A., Murray, R. M., et al. (2015). Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife* 4:e09771. doi: 10.7554/eLife.09771

Nielsen, A. A. K., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., et al. (2016). Genetic circuit design automation. *Science* 352:7341. doi: 10.1126/science.aac7341

Nielsen, C., Rahman, A., Rehman, A. U., Walsh, M. K., and Miller, C. D. (2017). Food waste conversion to microbial polyhydroxyalkanoates. *Microb. Biotechnol.* 2:12776. doi: 10.1111/1751-7915.12776

Nikel, P. I., de Almeida, A., Melillo, E. C., Galvagno, M. A., and Pettinari, M. J. (2006). New recombinant *Escherichia coli* strain tailored for the production of poly(3-hydroxybutyrate) from agroindustrial by-products. *Appl. Environ. Microbiol.* 72, 3949–3954. doi: 10.1128/AEM.00044-06

Ogonah, O. W., Polizzi, K. M., and Bracewell, D. G. (2017). Cell free protein synthesis: a viable option for stratified medicines manufacturing? *Curr. Opin. Chem. Eng.* 18, 77–83. doi: 10.1016/j.coche.2017.10.003

O'Kane, P. T., Dudley, Q. M., McMillan, A. K., Jewett, M. C., and Mrksich, M. (2019). High-throughput mapping of CoA metabolites by SAMDI-MS to optimize the cell-free biosynthesis of HMG-CoA. *Sci. Adv.* 5:eaaw9180. doi: 10.1126/sciadv.aaw9180

Opgenorth, P. H., Korman, T. P., and Bowie, J. U. (2016). A synthetic biochemistry module for production of bio-based chemicals from glucose. *Nat. Chem. Biol.* 12, 1–4. doi: 10.1038/nchembio.2062

Pardee, K. (2018). Perspective: solidifying the impact of cell-free synthetic biology through lyophilization. *Biochem. Eng. J.* 138, 91–97. doi: 10.1016/j.bej.2018.07.008

Pardee, K., Green, A. A., Ferrante, T., Cameron, D. E., DaleyKeyser, A., Yin, P., et al. (2014). Paper-based synthetic gene networks. *Cell* 159, 940–954. doi: 10.1016/j.cell.2014.10.004

Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., et al. (2016a). Rapid, low-cost detection of zika virus using programmable biomolecular components. *Cell* 165, 1255–1266. doi: 10.1016/j.cell.2016.04.059

Pardee, K., Slomovic, S., Nguyen, P. Q., Lee, J. W., Donghia, N., Burrill, D., et al. (2016b). Portable, on-demand biomolecular manufacturing. *Cell* 167, 248.e–259.e. doi: 10.1016/j.cell.2016.09.013

Park, N., Kahn, J. S., Rice, E. J., Hartman, M. R., Funabashi, H., Xu, J., et al. (2009a). High-yield cell-free protein production from P-gel. *Nat. Protoc.* 4, 1759–1770. doi: 10.1038/nprot.2009.174

Park, N., Um, S. H., Funabashi, H., Xu, J., and Luo, D. (2009b). A cell-free protein-producing gel. *Nat. Mater.* 8, 432–437. doi: 10.1038/nmat2419

Philp, J. (2018). The bioeconomy, the challenge of the century for policy makers. *N. Biotechnol.* 40, 11–19. doi: 10.1016/j.nbt.2017.04.004

Rajakumar, P. D., Gowers, G.-O. F., Suckling, L., Foster, A., Ellis, T., Kitney, R. I., et al. (2019). Rapid prototyping platform for *Saccharomyces cerevisiae* using computer-aided genetic design enabled by parallel software and workcell platform development. *SLAS Technol. Transl. Life Sci. Innov.* 24, 291–297. doi: 10.1177/2472630318798304

Roberts, A. D., Finnigan, W., Wolde-Michael, E., Kelly, P., Blaker, J. J., Hay, S., et al. (2019). Synthetic biology for fibers, adhesives, and active camouflage materials in protection and aerospace. *MRS Commun.* 9, 486–504. doi: 10.1557/mrc.2019.35

Roell, M.-S., and Zurbriggen, M. D. (2020). The impact of synthetic biology for future agriculture and nutrition. *Curr. Opin. Biotechnol.* 61, 102–109. doi: 10.1016/j.copbio.2019.10.004

Rolf, J., Rosenthal, K., and Lütz, S. (2019). Application of cell-free protein synthesis for faster biocatalyst development. *Catalysts* 9:190. doi: 10.3390/catal9020190

Rues, R.-B., Henrich, E., Boland, C., Caffrey, M., and Bernhard, F. (2016). Cell-free production of membrane proteins in *Escherichia coli* lysates for functional and structural studies. *Methods Mol. Biol.* 1432, 1–21. doi: 10.1007/978-1-4939-3637-3_1

Ruiz, R. C. H., Kiatwuthinon, P., Kahn, J. S., Roh, Y. H., and Luo, D. (2012). Cell-free protein expression from DNA-based hydrogel (P-Gel) droplets for scale-up production. *Ind. Biotechnol.* 8, 372–377. doi: 10.1089/ind.2012.0024

Russo, D. A., Zedler, J. A. Z., and Jensen, P. E. (2019). A force awakens: exploiting solar energy beyond photosynthesis. *J. Exp. Bot.* 70, 1703–1710. doi: 10.1093/jxb/erz054

Rustad, M., Eastlund, A., Jardine, P., and Noireaux, V. (2018). Cell-free TXTL synthesis of infectious bacteriophage T4 in a single test tube reaction. *Synth. Biol.* 3, 1–7. doi: 10.1093/synbio/ysy002

Salehi, A. S. M., Shakalli Tang, M. J., Smith, M. T., Hunt, J. M., Law, R. A., Wood, D. W., et al. (2017). Cell-Free protein synthesis approach to biosensing hTRβ-specific endocrine disruptors. *Anal. Chem.* 89, 3395–3401. doi: 10.1021/acs.analchem.6b04034

Salehi, A. S. M., Smith, M. T., Bennett, A. M., Williams, J. B., Pitt, W. G., and Bundy, B. C. (2016). Cell-free protein synthesis of a cytotoxic cancer therapeutic: onconase production and a just-add-water cell-free system. *Biotechnol. J.* 11, 274–281. doi: 10.1002/biot.201500237

Shelby, M. L., He, W., Dang, A. T., Kuhl, T. L., and Coleman, M. A. (2019). Cell-free co-translational approaches for producing mammalian receptors: expanding the cell-free expression toolbox using nanolipoproteins. *Front. Pharmacol.* 10:744. doi: 10.3389/fphar.2019.00744

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., et al. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755. doi: 10.1038/90802

Shimizu, Y., Kanamori, T., and Ueda, T. (2005). Protein synthesis by pure translation systems. *Methods* 36, 299–304. doi: 10.1016/j.ymeth.2005.04.006

Shin, J., and Noireaux, V. (2012). An *E. coli* cell-free expression toolbox: application to synthetic gene circuits and artificial cells. *ACS Synth. Biol.* 1, 29–41. doi: 10.1021/sb200016s

Shinoda, T., Shinya, N., Ito, K., Ishizuka-Katsura, Y., Ohsawa, N., Terada, T., et al. (2016). Cell-free methods to produce structurally intact mammalian membrane proteins. *Sci. Rep.* 6, 1–15. doi: 10.1038/srep30442

Shurtleff, M. J., Temoche-Diaz, M. M., Karfilis, K. V., Ri, S., and Schekman, R. (2016). Y-box protein 1 is required to sort microRNAs into exosomes in cells and in a cell-free reaction. *Elife* 5, 1–23. doi: 10.7554/eLife.19276

Silverman, A. D., Karim, A. S., and Jewett, M. C. (2019). Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* 21, 151–170. doi: 10.1038/s41576-019-0186-3

Smith, M. T., and Berkheimer, S. D. (2014). Lyophilized *Escherichia coli* -based cell-free systems for robust, high-density, long-term storage. *Biotechniques* 56:158. doi: 10.2144/000114158

Stock, T., and Seliger, G. (2016). Opportunities of sustainable manufacturing in industry 4.0. *Proc. CIRP* 40, 536–541. doi: 10.1016/j.procir.2016.01.129

Storch, M., Haines, M. C., and Baldwin, G. S. (2019). DNA-BOT: a low-cost, automated DNA assembly platform for synthetic biology. *bioRxiv* [Preprint], doi: 10.1101/832139

Sullivan, C. J., Pendleton, E. D., Sasmor, H. H., Hicks, W. L., Farnum, J. B., Muto, M., et al. (2016). A cell-free expression and purification process for rapid production of protein biologics. *Biotechnol. J.* 11, 238–248. doi: 10.1002/biot.201500214

Sun, Z. Z., Hayes, C. A., Shin, J., Caschera, F., Murray, R. M., and Noireaux, V. (2013). Protocols for implementing an *Escherichia coli* based TX-TL cell-free expression system for synthetic biology. *J. Vis. Exp.* e50762. doi: 10.3791/50762

Sun, Z. Z., Yeung, E., Hayes, C. A., Noireaux, V., and Murray, R. M. (2014). Linear DNA for rapid prototyping of synthetic biological circuits in an *Escherichia coli* Based TX-TL cell-free system. *ACS Synth. Biol.* 3, 387–397. doi: 10.1021/sb400131a

Swank, Z., Laohakunakorn, N., and Maerkl, S. J. (2019). Cell-free gene-regulatory network engineering with synthetic transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* 116, 5892–5901. doi: 10.1073/pnas.1816591116

Tan, D., Xue, Y.-S., Aibaidula, G., and Chen, G.-Q. (2011). Unsterile and continuous production of polyhydroxybutyrate by halomonas TD01. *Bioresour. Technol.* 102, 8130–8136. doi: 10.1016/j.biortech.2011.05.068

Tao, W., Lv, L., and Chen, G.-Q. (2017). Engineering halomonas species TD01 for enhanced polyhydroxyalkanoates synthesis via CRISPRi. *Microb. Cell Fact.* 16:48. doi: 10.1186/s12934-017-0655-3

Tarrahi, R., Fathi, Z., Seydibeyoğlu, M. Ö, Doustkhah, E., and Khataee, A. (2020). Polyhydroxyalkanoates (PHA): from production to nanoarchitecture. *Int. J. Biol. Macromol.* 146, 596–619. doi: 10.1016/j.ijbiomac.2019.12.181

Thavarajah, W., Silverman, A. D., Verosloff, M. S., Kelley-Loughnane, N., Jewett, M. C., and Lucks, J. B. (2020). Point-of-use detection of environmental fluoride via a cell-free riboswitch-based biosensor. *ACS Synth. Biol.* 9, 10–18. doi: 10.1021/acssynbio.9b00347

Thomson, N., Roy, I., Summers, D., and Sivaniah, E. (2009). In vitro production of polyhydroxyalkanoates: achievements and applications. *J. Chem. Technol. Biotechnol.* 85, 760–767. doi: 10.1002/jctb.2299

Tomizawa, S., Yoshioka, M., Ushimaru, K., and Tsuge, T. (2012). Preparative synthesis of Poly(R)-3-hydroxybutyrate monomer for enzymatic cell-free polymerization. *Polym. J.* 44, 982–985. doi: 10.1038/pj.2012.34

Townsend, T., and Sette, J. (2016). "Natural fibres and the world economy," in *Natural Fibres: Advances in Science and Technology Towards Industrial Applications*, eds R. Fangueiro and S. Rana (Dordrecht: Springer), 381–390. doi: 10.1007/978-94-017-7515-1_30

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

Ullah, M. W., Ul-Islam, M., Khan, S., Kim, Y., and Park, J. K. (2015). Innovative production of bio-cellulose using a cell-free system derived from a single cell line. *Carbohydr. Polym.* 132, 286–294. doi: 10.1016/j.carbpol.2015.06.037

Voyvodic, P. L., Pandi, A., Koch, M., Conejero, I., Valjent, E., Courtet, P., et al. (2019). Plug-and-play metabolic transducers expand the chemical detection space of cell-free biosensors. *Nat. Commun.* 10:1697. doi: 10.1038/s41467-019-09722-9

Wagner, H. J., Engesser, R., Ermes, K., Geraths, C., Timmer, J., and Weber, W. (2019). Synthetic biology-inspired design of signal-amplifying materials systems. *Mater. Today* 22, 25–34. doi: 10.1016/j.mattod.2018.04.006

Walsh, D. I., Pavan, M., Ortiz, L., Wick, S., Bobrow, J., Guido, N. J., et al. (2019). Standardizing automated DNA assembly: best practices, metrics, and protocols using robots. *SLAS Technol. Transl. Life Sci. Innov.* 24, 282–290. doi: 10.1177/2472630318825335

Wan, X., Volpetti, F., Petrova, E., French, C., 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

Weber, L. A., Feman, E. R., and Baglioni, C. (1975). Cell free system from HeLa cells active in initiation of protein synthesis. *Biochemistry* 14, 5315–5321. doi: 10.1021/bi00695a015

Wen, K. Y., Cameron, L., Chappell, J., Jensen, K., Bell, D. J., Kelwick, R., et al. (2017). A Cell-free biosensor for detecting quorum sensing molecules in *P. aeruginosa* -infected respiratory samples. *ACS Synth. Biol.* 6, 2293–2301. doi: 10.1021/acssynbio.7b00219

Wilding, K. M., Schinn, S.-M., Long, E. A., and Bundy, B. C. (2018). The emerging impact of cell-free chemical biosynthesis. *Curr. Opin. Biotechnol.* 53, 115–121. doi: 10.1016/j.copbio.2017.12.019

Wong, H. H., and Lee, S. Y. (1998). Poly-(3-hydroxybutyrate) production from whey by high-density cultivation of recombinant *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 50, 30–33. doi: 10.1007/s002530051252

Yan, Q., and Fong, S. S. (2015). Bacterial chitinase: nature and perspectives for sustainable bioproduction. *Bioresour. Bioprocess.* 2:31. doi: 10.1186/s40643-015-0057-5

Yim, S. S., Johns, N. I., Park, J., Gomes, A. L., McBee, R. M., Richardson, M., et al. (2019). Multiplex transcriptional characterizations across diverse bacterial species using cell-free systems. *Mol. Syst. Biol.* 15:875. doi: 10.15252/msb.20198875

Yue, K., Zhu, Y., and Kai, L. (2019). Cell-free protein synthesis: chassis toward the minimal cell. *Cells* 8:315. doi: 10.3390/cells8040315

Zawada, J. F., Yin, G., Steiner, A. R., Yang, J., Naresh, A., Roy, S. M., et al. (2011). Microscale to manufacturing scale-up of cell-free cytokine production-a new approach for shortening protein production development timelines. *Biotechnol. Bioeng.* 108, 1570–1578. doi: 10.1002/bit.23103

Zhang, X., Lin, Y., Wu, Q., Wang, Y., and Chen, G.-Q. (2019). Synthetic biology and genome-editing tools for improving pha metabolic engineering. *Trends Biotechnol.* doi: 10.1016/j.tibtech.2019.10.006

Zhang, Y., Minagawa, Y., Kizoe, H., Miyazaki, K., Iino, R., Ueno, H., et al. (2019). Accurate high-throughput screening based on digital protein synthesis in a massively parallel femtoliter droplet array. *Sci. Adv.* 5:eaav8185. doi: 10.1126/sciadv.aav8185

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

ORIGINAL RESEARCH
published: 18 June 2020
doi: 10.3389/fbioe.2020.00536

Establishing a Eukaryotic *Pichia pastoris* Cell-Free Protein Synthesis System

*Lingkai Zhang, Wan-Qiu Liu and Jian Li**

School of Physical Science and Technology, ShanghaiTech University, Shanghai, China

In recent years, cell-free protein synthesis (CFPS) systems have been used to synthesize proteins, prototype genetic elements, manufacture chemicals, and diagnose diseases. These exciting, novel applications lead to a new wave of interest in the development of new CFPS systems that are derived from prokaryotic and eukaryotic organisms. The eukaryotic *Pichia pastoris* is emerging as a robust chassis host for recombinant protein production. To expand the current CFPS repertoire, we report here the development and optimization of a eukaryotic CFPS system, which is derived from a protease-deficient strain *P. pastoris* SMD1163. By developing a simple crude extract preparation protocol and optimizing CFPS reaction conditions, we were able to achieve superfolder green fluorescent protein (sfGFP) yields of 50.16 ± 7.49 μg/ml in 5 h batch reactions. Our newly developed *P. pastoris* CFPS system fits to the range of the productivity achieved by other eukaryotic CFPS platforms, normally ranging from several to tens of micrograms protein per milliliter in batch mode reactions. Looking forward, we believe that our *P. pastoris* CFPS system will not only expand the CFPS toolbox for synthetic biology applications, but also provide a novel platform for cost-effective, high-yielding production of complex proteins that need post-translational modification and functionalization.

Keywords: cell-free protein synthesis, *Pichia pastoris*, yeast, eukaryote, protein expression, cell-free synthetic biology

OPEN ACCESS

Edited by:
Tao Chen,
Tianjin University, China

Reviewed by:
Kristina Hedfalk,
University of Gothenburg, Sweden
Jiazhang Lian,
Zhejiang University, China

****Correspondence:***
Jian Li
lijian@shanghaitech.edu.cn

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

Received: *24 February 2020*
Accepted: *04 May 2020*
Published: *18 June 2020*

Citation:
Zhang L, Liu W-Q and Li J (2020) Establishing a Eukaryotic Pichia pastoris Cell-Free Protein Synthesis System. Front. Bioeng. Biotechnol. 8:536. doi: 10.3389/fbioe.2020.00536

INTRODUCTION

Cell-free protein synthesis (CFPS) systems are emerging as effective platforms for *in vitro* synthetic biology and biotechnology applications from fundamental research to biomanufacturing (Carlson et al., 2012; Bundy et al., 2018; Li et al., 2018b; Swartz, 2018; Khambhati et al., 2019; Liu et al., 2019; Silverman et al., 2020). Such platforms separate the cell growth and the protein synthesis into two stages, which can alleviate the cell's metabolic burden and enhance the productivity. Due to the open nature of CFPS, cell-free reactions can bypass limitations on mass transfer and are more tolerant of toxic protein products. Additionally, the process of CFPS without cell walls can be easily manipulated, controlled, and optimized. Therefore, CFPS systems have recently attracted considerable attention as a robust approach for the production of various proteins, for example, membrane proteins (Henrich et al., 2015; Sonnabend et al., 2017), therapeutic proteins (Min et al., 2016; Wilding et al., 2019), unnatural amino acid modified proteins (Martin et al., 2018; Gao et al., 2019), and difficult-to-express proteins (Li et al., 2016; Jin and Hong, 2018). With the advances of synthetic biology, CFPS technology

has also been used to construct protein-based biosensors (Pardee et al., 2016; Thavarajah et al., 2020), metabolic pathways (Goering et al., 2017; Zhuang et al., 2020), high-throughput screening platforms (Sawasaki et al., 2002; Swank et al., 2019), bottom-up synthetic cells (Karzbrun et al., 2014; van Nies et al., 2018), and classroom education kits (Huang et al., 2018; Stark et al., 2018), among others.

Due to the aforementioned emerging applications of CFPS systems, many previous efforts have been focused on the optimization and enhancement of a selected few model systems like the *Escherichia coli* and wheat germ platforms (Carlson et al., 2012; Perez et al., 2016). Unfortunately, these well-developed CFPS systems may have their own disadvantages and drawbacks such as the lack of post-translational modifications (e.g., glycosylation), incorrect protein folding without suitable chaperones, and low protein yields (Zemella et al., 2015). In order to tackle these problems, several new CFPS systems have recently been developed to better mimic the physicochemical environment of native hosts for synthetic biology and biotechnology applications. However, the newly developed CFPS systems are mainly derived from prokaryotic microorganisms, including some from *Streptomyces* species (Li et al., 2017, 2018a; Moore et al., 2017), *Bacillus subtilis* (Kelwick et al., 2016), *Pseudomonas putida* (Wang et al., 2018), and *Vibrio natriegens* (Des Soye et al., 2018; Failmezger et al., 2018; Wiegand et al., 2018). Although a couple of eukaryote-based CFPS systems are available, they are mostly prepared from plant (e.g., wheat germ), insect (e.g., *Spodoptera frugiperda*), and mammalian (e.g., Chinese hamster ovary, CHO) cells (Tarui et al., 2001; Takai et al., 2010; Brödel et al., 2014), which often need laborious and expensive cell extract preparation approaches. For example, it takes 4–5 days to prepare wheat germ extracts (ca. 5 ml) from 5 to 6 kg seeds with the steps of grinding, sieving, extensive washing, and eye selection of the embryo particles (Takai et al., 2010). To date, only a few eukaryotic microorganisms (e.g., *Saccharomyces cerevisiae*), which can be easily cultivated in the laboratory, have been used to develop eukaryotic CFPS platforms (Hodgman and Jewett, 2013; Gan and Jewett, 2014). Despite its success, the protein yield of the *S. cerevisiae*-based CFPS system is relatively low (<10 μg protein/ml) (Hodgman and Jewett, 2013). Therefore, it is highly desirable to develop more eukaryotic microorganism-based CFPS systems to expand the protein expression toolkit for the rapid synthesis, study, and engineering of proteins.

The methylotrophic yeast *Pichia pastoris*, a generally recognized as safe (GRAS) eukaryotic microorganism, has emerged as a reliable and robust chassis host for biotechnological applications in both laboratory and industry (Bill, 2014). Specifically, *P. pastoris* has been well-documented as a cell factory to produce recombinant products such as therapeutic proteins, industrial enzymes, and antimicrobial peptides (Ahmad et al., 2014; Kim et al., 2015; Peña et al., 2018; Yang and Zhang, 2018). The use of *P. pastoris* as an attractive expression system is largely due to its rapid growth on simple media (Darby et al., 2012), readily genetic manipulation tools (e.g., CRISPR-Cas technology) (Raschmanová et al., 2018), and proper eukaryotic post-translational modifications (e.g., humanized *N*-linked glycosylation) (Hamilton et al., 2006). In addition, the genome sequence of *P. pastoris* is available (De Schutter et al., 2009), which provides more opportunities to engineer the organism for desired goals (Peña et al., 2018; Yang and Zhang, 2018). For example, disruption of protease genes in *P. pastoris* generates protease-deficient strains that can prevent recombinant protein degradation and thus increase the product yield (Gleeson et al., 1998; Ni et al., 2008; Wu et al., 2013).

In this work, we aim to establish a eukaryotic microorganism-based CFPS system that is derived from a protease-deficient yeast strain *P. pastoris* SMD1163. After showing the baseline ability to synthesize a reporter protein, we set out to investigate cell lysis procedures to obtain highly active cell extracts, which contain the necessary catalytic components for transcription, translation, and protein folding (e.g., aminoacyl-tRNA synthetases, ribosomes, elongation factors, chaperones, etc.). Then, we assessed the effect of cultivation time, energy conditions, and other physicochemical parameters on protein synthesis yields. Finally, we achieved a ∼55-fold increase in protein yields as compared to the initial yield of 0.91 ± 0.12 μg/ml. This work establishes a robust and easy to use eukaryotic CFPS system, which we anticipate that it will serve as an alternative platform for the synthesis of "difficult-to-express" proteins that need, for example, glycosylation, as well as for broad synthetic biology applications.

MATERIALS AND METHODS

Strains and Culture Medium

The protease-deficient yeast strain *P. pastoris* SMD1163 (*pep4 prb1*) was used in this work. Yeast cells were cultivated in a liquid YPD medium consisting of (per liter) 10 g yeast extract, 20 g peptone, and 20 g glucose.

Plasmid Construction

The superfolder green fluorescent protein (sfGFP) is used as a reporter protein. All plasmids were constructed by modifying the pJL1 expression vector with sfGFP, which is a gift from Michael Jewett (Addgene plasmid # 69496). The cloning was performed according to the Gibson assembly method (Gibson et al., 2009). All gene fragments and sequences used in this study were synthesized by GENEWIZ (Suzhou, China). Initially, a codon optimized sfGFP sequence according to the codon usage in *P. pastoris* was synthesized to replace the original sfGFP gene in pJL1 between the restriction sites *Nde*I and *Sal*I. Then, a synthetic 50 bp poly(A) tail was inserted to the 3′ end of the sfGFP gene. Afterwards, several internal ribosome entry site (IRES) sequences were synthesized and individually cloned to the vector in front of the sfGFP gene. In addition, when the cricket paralysis virus (CrPV) IRES was used, a Kozak sequence (GAAACG) was included after CrPV. All constructs were verified by DNA sequencing (GENEWIZ, Suzhou, China). IRES and codon optimized sfGFP sequences are shown in **Table S1**.

Cell Cultivation and Harvest

Yeast cells were cultivated in liquid YPD medium at 30°C in an orbital shaker at 250 rpm. An overnight culture of *P. pastoris* was used to inoculate 1 L fresh YPD medium in a

2.5 L baffled Ultra Yield™ flask (Thomson Instrument Company, USA) with an initial OD_{600} of 0.05. After 18 h cultivation (mid-exponential phase, an OD_{600} of ∼6), the cells were harvested by centrifugation at 3,000 g and 4°C for 15 min. Cell pellets were then washed three times with cold washing buffer (30 mM HEPES pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate, 2 mM dithiolthretol). After the final wash and centrifugation, the pelleted cells were weighed, flash-frozen in liquid nitrogen, and stored at −80°C until further use. Alternatively, cells can be lysed immediately to make cell extracts.

Cell Extract Preparation

Frozen cells were thawed on ice for 30 min before lysis. The thawed cells were resuspended in 1.5 ml of cold lysis buffer (30 mM HEPES pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate, 2 mM dithiolthretol, 0.5 mM PMSF) per gram of wet cell weight. Cell disruption was performed using three approaches: sonication, high-pressure homogenization, and 0.5 mm glass beads.

For sonication lysis, the cells were disrupted by using a Q125 Sonicator (Qsonica, Newtown, USA) with 45 s On/60 s Off for five cycles (3 mm diameter probe, 50% of amplitude). For high-pressure homogenization lysis, the smooth suspended cells were lysed by a UH-06 homogenizer (Union-Biotech, Shanghai, China) with two passes at a pressure of 1,200 bar. For glass beads disruption, the cell suspension was mixed with 0.5 mm glass beads (Tansoole, Shanghai, China) in a 50 ml falcon tube at a mass ratio of 1:1 (cell:bead, g/g). Then, the mixture was vortexed vigorously using a vortex mixer (Vortex-Genie 2, New York, USA) for 40 cycles with 1 min on vortex and 1 min on ice.

After cell disruption, the lysate was centrifuged at 30,000 g and 4°C for 30 min. The supernatant was transferred to a fresh tube and clarified again with the same condition. The resultant supernatant was carefully removed and underwent buffer exchange by dialysis with a 3.5 kDa molecular weight cut-off (MWCO) membrane. The lysate was dialyzed against four exchanges of 50-volumes of fresh lysis buffer (30 mM HEPES pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate, 2 mM dithiolthretol, 0.5 mM PMSF) for 30 min each at 4°C. After dialysis, the extract was centrifuged at 21,000 g and 4°C for 30 min. The resultant supernatant as cell extract was collected, aliquoted, immediately flash-frozen in liquid nitrogen, and finally stored at −80°C until use.

Cell-Free Protein Synthesis

Coupled cell-free transcription and translation reactions were performed as described previously (Hodgman and Jewett, 2013) with some modifications. Standard CFPS reactions were carried out in 1.5 ml microcentrifuge tubes. Each reaction (15 μl) contains the following components: 25 mM HEPES-KOH pH 7.4, 120 mM potassium glutamate, 6 mM magnesium glutamate, 1.5 mM of each ATP, GTP, CTP, and UTP, 0.1 mM of each of 20 amino acids, 25 mM creatine phosphate, 1.7 mM DTT, 1 mM putrescine, 0.5 mM spermidine, 0.27 mg/ml creatine phosphokinase (from rabbit muscle; Sigma–Aldrich), 16.7 μg/ml plasmid, 60 U T7 RNA polymerase (Thermo Fisher Scientific), and 50% (v/v) cell extract. All reactions were mixed using above conditions and incubated at 23°C for 5 h unless otherwise noted.

Protein Quantification

The reporter protein sfGFP was used to measure and optimize protein synthesis activity of the *P. pastoris*-based CFPS system. After the reactions, two microliters of the CFPS sample were mixed with 48 μl nuclease-free water and placed in a 96-well plate with flat bottom. The fluorescence of sfGFP was measured using a BioTek SYNETGY H1 plate reader with excitation and emission wavelength at 485 and 528 nm, respectively. sfGFP fluorescence units were converted to concentration (μg/ml) according to a linear standard curve made in house with purified sfGFP. For each protein quantification, at least three independent experiments were carried out using the same cell extract. Then, the protein concentration of each independent reaction was technically measured in triplicate.

RESULTS AND DISCUSSION

Development of an Initial *P. pastoris*-Based CFPS System

In general, during the initial development of a new CFPS system, two primary requirements that need to be considered are the choice of a suitable strain and the construction of an efficient expression vector (Brödel et al., 2014; Gan and Jewett, 2014; Kelwick et al., 2016; Li et al., 2017; Des Soye et al., 2018). For example, a protease-deficient *B. subtilis* strain cell extracts produced notably higher and more consistent yields of a reporter protein than a wild-type strain with endogenous proteases (Kelwick et al., 2016). In addition, when constructing an expression vector for eukaryotic CFPS systems, IRES sequences that can recruit eukaryotic ribosomes to initiate cap-independent translation were often investigated to enable combined cell-free transcription and translation (Brödel et al., 2014; Gan and Jewett, 2014). In order to establish a robust *P. pastoris*-based CFPS system, we began our study by trying to adopt the protocol used for the *S. cerevisiae* CFPS system (Hodgman and Jewett, 2013; Gan and Jewett, 2014). Our first step was to find a proper *P. pastoris* strain. Although the strain *P. pastoris* GS115 was reported as a robust host to express recombinant proteins (Gurramkonda et al., 2009; Nie et al., 2014; Zheng et al., 2019; Wang et al., 2020), cell extracts prepared from this strain showed relatively little protein synthesis activity with very low and unreliable yields of sfGFP (data not shown). Then, we switched to the strain *P. pastoris* SMD1163 (*pep4 prb1*), which is derived from GS115 with knocking out of two protease genes (*pep4* encodes protease A and *prb1* encodes protease B) (Gleeson et al., 1998). By doing this, SMD1163 cell extracts were active to synthesize sfGFP. A representative time course of sfGFP synthesis was monitored by online fluorescence measurement (**Figure 1A**). Our data indicated that the synthesis rate of sfGFP was the highest during the first 1 h reaction with a nearly linear increase manner. Then, the protein synthesis rate declined from 1 to 3 h and no obvious increase of the fluorescence was observed between 3 and 5 h. Therefore, all following CFPS reactions were terminated after 5 h of incubation.

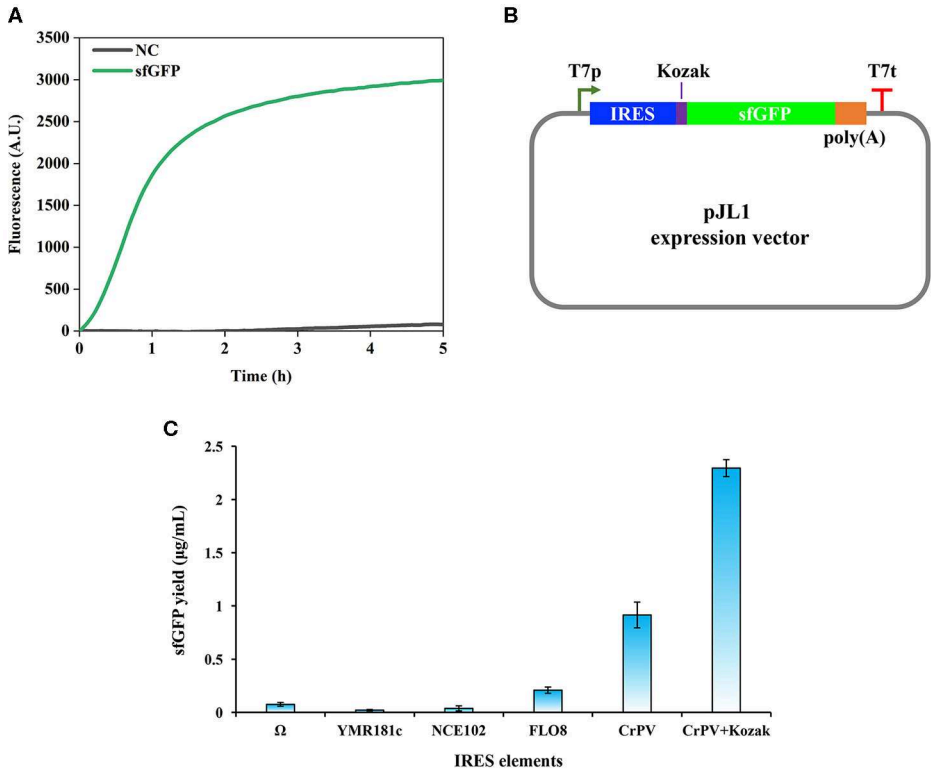


FIGURE 1 | Cell-free protein synthesis of superfolder green fluorescent protein (sfGFP) using *Pichia pastoris* extract. **(A)** Time course of sfGFP synthesis with online fluorescence measurement. NC, negative control without plasmid in the reaction. **(B)** Design of expression vectors based on the pJL1 plasmid. IRES, internal ribosome entry site; Kozak, a 6 bp sequence (GAAACG); poly(A), a 50 bp poly(A) tail; T7p, T7 promoter; T7t, T7 terminator. **(C)** Effects of IRES elements on the cell-free synthesis of sfGFP in a *P. pastoris* CFPS system. Values show means with error bars representing standard deviations (s.d.) of at least three independent experiments.

Having validated the combined cell-free transcription and translation, we next set out to evaluate the impact of IRES sequences on the protein synthesis. IRES elements are commonly used by viruses when they infect eukaryotic cells to recruit cellular ribosomes to start cap-independent translation for their own protein synthesis (Baird et al., 2006). This advantage has been taken to develop eukaryotic CFPS systems (Takai et al., 2010; Brödel et al., 2014; Gan and Jewett, 2014), which can facilitate the protein translation process without the laborious preparation of capped mRNA templates. In eukaryotic CFPS systems, two IRES elements often used are the Ω sequence from tobacco mosaic virus (TMV) and the cricket paralysis virus (CrPV) IRES sequence. We, therefore, chose Ω and CrPV IRES sequences to construct our expression vectors. In addition, several IRES sequences have also been identified in *S. cerevisiae* such as *YMR181c*, *NCE102*, and *FLO8* IRES sequences (Gilbert et al., 2007). Since *P. pastoris* and *S. cerevisiae* are similar yeast strains, the functionality of above three IRES sequences was also evaluated together with Ω and CrPV IRES sequences. The design of expression vectors is shown in **Figure 1B**. CFPS reactions with different expression vectors were performed in 15 μl batch reactions at 23°C for 5 h. As shown in **Figure 1C**, the CrPV IRES sequence showed the highest activity among all tested cap-independent translation sequences. The Ω sequence was found to be the best IRES in the *S. cerevisiae* CFPS reaction (Gan and Jewett, 2014), however, it showed low activity in our *P. pastoris* CFPS system. While the three native IRES sequences from *S. cerevisiae* were able to initiate protein translation, their activities were significantly lower than that of the CrPV IRES sequence.

In eukaryotic cells, there are some consensus sequences, which are the so-called Kozak sequences and locate in the upstream of open reading frames (ORF) for ensuring efficient translation initiation (Kozak, 1986, 1987). As reported previously, a Kozak sequence (GAAACG) from the native alcohol oxidase 1 (*AOX1*) gene of *P. pastoris* was often used to construct expression vectors in front of the start codon for enhancing protein translation and recombinant protein yields (Mellitzer et al., 2012; Várnai et al., 2014). The benefit of Kozak sequences was also observed in eukaryotic CFPS systems (Kozak, 1990; Aw and Polizzi, 2019). In order to test if the Kozak sequence (GAAACG) from *AOX1* helps enhance the translation efficiency in our *P. pastoris* CFPS system, it was inserted between the CrPV IRES sequence and the ATG start codon. Our data indicated that the Kozak sequence improved the protein yield to 2.29 ± 0.08 μg/ml, which is about 2.5 times higher than that of the plasmid without the Kozak sequence (the sfGFP yield was 0.91 ± 0.12 μg/ml, **Figure 1C**). Therefore, the expression plasmid with the CrPV IRES element plus the Kozak sequence was used in our following experiments for further optimization.

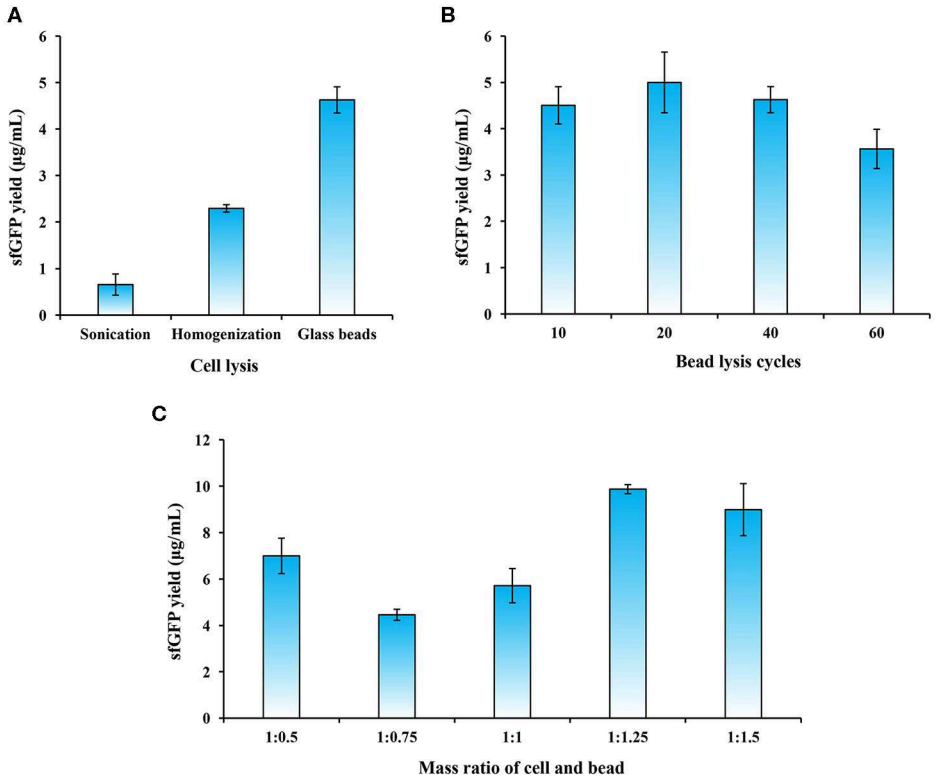


FIGURE 2 | Optimization of cell extract preparation. **(A)** Evaluation of three cell lysis methods. **(B)** Effects of bead lysis cycles on cell extract activity. **(C)** Optimization of the mass ratio of cell and bead (g/g) on sfGFP synthesis. Values show means with error bars representing standard deviations (s.d.) of at least three independent experiments.

Identification of Optimal Procedures for Preparing *P. pastoris* Cell Extracts

Because cell extracts contain essential components like ribosomes, aminoacyl-tRNA synthetases, and chaperons for protein synthesis, it is crucial to prepare highly active cell lysates to support CFPS reactions. For disrupting different types of cells, commonly used cell lysis methods include sonication and high-pressure homogenization (Gan and Jewett, 2014; Kelwick et al., 2016; Li et al., 2017, 2018a; Des Soye et al., 2018; Failmezger et al., 2018; Wang et al., 2018). In addition, glass beads-based disruption has also been used to lyse cells such as *E. coli* (Kigawa, 2010; Sun et al., 2013) and *S. cerevisiae* (Hofbauer et al., 1982; Hodgman and Jewett, 2013) for making cell lysates for CFPS reactions. To identify an optimal cell lysis method, we evaluated the above-mentioned three techniques. Our data demonstrated that the most productive lysate was generated by the bead beating method, which resulted in the highest yield of sfGFP at 4.63 ± 0.28 μg/ml as compared to the other two methods (**Figure 2A**). The use of beads for cell disruption is simple and promising because no expensive equipment is required like the homogenizer and such method can be easily adopted by other laboratories.

Since the bead lysis method utilizes a grinding mechanism of action, the amount of glass beads in the cell suspension plays an important role in the generation of crushing and grinding forces that break up the cells. We, therefore, next sought to optimize the conditions during the bead lysis process including cell lysis cycles and the mass ratio of glass beads and cell biomass. To this end, we first compared the sfGFP yield with different lysis cycles (one cycle is 1 min on vortex mixer and 1 min on ice for cooling). The results indicated that the highest sfGFP yield reached at 4.99 ± 0.66 μg/mL with 20 cycles of cell lysis (**Figure 2B**). However, more cycles (>40) slightly reduced the protein yields. Next, we investigated the effect of mass ratio of glass beads and cell biomass on the sfGFP synthesis. The data suggested that a higher ratio of bead in the mixture was better to grind the cells with a ratio of 1:1.25 (cell:bead, g/g) maximizing protein expression in our experiments (**Figure 2C**). This is likely due to the fact that more beads may provide higher grinding forces and thus can efficiently disrupt cells as reported previously (Hofbauer et al., 1982; Kigawa, 2010; Sun et al., 2013).

Effect of Cell Harvest Time on CFPS Reactions

The activity of crude extract-based CFPS reactions is essentially dependent on the composition of the cellular machinery (e.g., ribosomes) at the time of cell harvest. Previous studies have shown that cells harvested during the exponential growth phase, especially, in the mid-exponential phase, could provide benefits for protein synthesis because the translation machinery is most

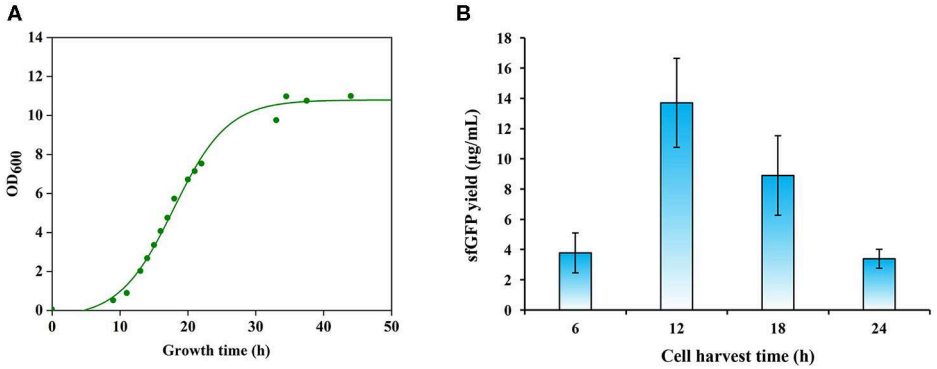


FIGURE 3 | Effects of cell growth phase on protein synthesis. **(A)** A representative growth curve of *Pichia pastoris*. **(B)** Comparison of sfGFP synthesis with cell extracts prepared from different growth phases. Values show means with error bars representing standard deviations (s.d.) of at least three independent experiments.

active (Hodgman and Jewett, 2013; Li et al., 2017; Wang et al., 2018). We were curious to determine the optimal cell harvest time for *P. pastoris* to prepare highly active cell extracts. We grew *P. pastoris* cells in 1 L media with an initial OD_{600} of 0.05 for 45 h at 30°C and 250 rpm. According to the cell growth curve (**Figure 3A**), we harvested cells after the cultivation at 12, 18, and 24 h, which spanned a range of early to late exponential growth phase in our culture media. Then, we prepared cell extracts from each of these cultivations. To compare these extracts, we performed CFPS reactions in parallel and sfGFP yields were determined. We found that cell extracts prepared from the early exponential phase (12 h) with a final OD_{600} of ∼2 were most active and synthesized the highest yield of sfGFP at 13.70 ± 2.94 μg/ml (**Figure 3B**). In addition, we harvested cells before the early exponential phase after 6 h cultivation for comparison. However, the protein yield from such extracts was significantly lower than the highest yield (**Figure 3B**). Not like other CFPS systems, cell extracts derived from the early exponential phase rather than the mid-exponential phase favored protein synthesis in our *P. pastoris* CFPS system. Taken together, our results highlight the importance of cell harvest as a critical factor for development and optimization of CFPS systems.

Optimization of CFPS Reaction Conditions

In order to further increase protein expression yields, we next carried out a systematic optimization of the *P. pastoris* CFPS system. Specifically, we investigated the effects of magnesium ion concentration, energy regeneration, template plasmid concentration, and reaction temperature on protein yields. We started with Mg^{2+} concentration, which is known to be a critical cation used in cell-free systems that influences the assembly and activity of ribosomes (Klein et al., 2004; Yamamoto et al., 2010; Petrov et al., 2012). Previous reports showed that different CFPS systems also need different concentrations of Mg^{2+} ions and the optimal values have to be optimized accordingly (Brödel et al., 2014; Kelwick et al., 2016; Li et al., 2017; Des Soye et al., 2018; Wang et al., 2018). We performed CFPS reactions with a range of Mg^{2+} concentrations from 2 to 12 mM. Our results suggested that 4 mM gave rise to the highest sfGFP yield of 15.04 ± 0.90 μg/ml (**Figure 4A**). Similar yields were obtained at around 13 μg/ml with 6 and 8 mM Mg^{2+} concentrations, respectively.

Energy supply in CFPS systems is an important factor that affects the efficiency of protein synthesis (Kim and Swartz, 1999; Jewett et al., 2008). The energy regeneration system with the pair of creatine phosphate/creatine phosphokinase (CP/CK) is often used in eukaryotic CFPS systems (Tarui et al., 2001; Takai et al., 2010; Hodgman and Jewett, 2013; Brödel et al., 2014). The CP/CK energy system was also employed in our *P. pastoris* CFPS system. Therefore, we next sought to optimize each concentration of CP and CK in the reaction mixture. To do this, we varied the concentrations of CP and CK separately in CFPS reactions. We observed that the optimal concentration for CP and CK were 40 mM and 0.6 mg/ml (**Figures 4B,C**), respectively. The optimized CP/CK system improved the sfGFP yield to 21.94 ± 2.08 μg/ml, which is ∼70% higher than the yield without optimization.

Another key factor needs to be optimized is the reaction temperature, because it affects protein synthesis and protein folding (Gan and Jewett, 2014; Li et al., 2017; Wang et al., 2018). We run CFPS reactions with different temperatures ranging from 20 to 37°C. Our data indicated that the sfGFP yields gradually improved with the increase of temperature from 20 to 30°C (**Figure 4D**). However, the yields notably decreased at temperatures of 33 and 37°C. The reaction temperature at 30°C, which is the same as the preferred growth temperature of *P. pastoris*, remarkably favored protein synthesis with the highest yield of 50.16 ± 7.49 μg/ml that is >2 times higher than the yield at 23°C (**Figure 4D**). Our finding is similar to other CFPS systems like *E. coli* and *B. subtilis* at 30°C (Kelwick et al., 2016; Li et al., 2016), albeit the optimal temperature for the *S. cerevisiae* CFPS reaction is 24°C (Gan and Jewett, 2014).

Finally, we investigated the impact of plasmid concentration on the performance of our *P. pastoris* CFPS system. The reactions were carried out in 15 μl volume by adding 200, 250, or 300 ng of template plasmids. The results showed that the sfGFP expression peaked when 250 ng of plasmid was supplied as template (**Figure 4E**). Taken together, the final, optimized *P. pastoris* CFPS

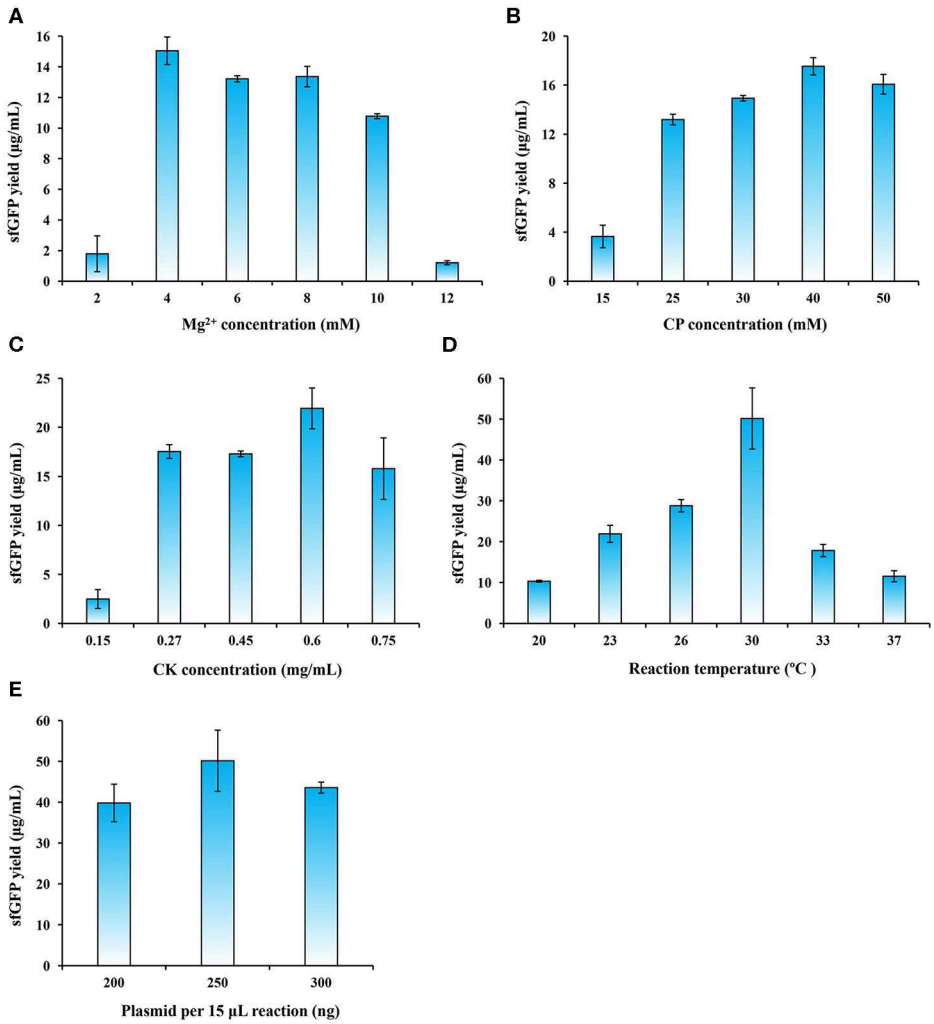


FIGURE 4 | CFPS reaction optimization for enhancing sfGFP expression yields. The CFPS reaction was optimized by altering **(A)** Mg^{2+} concentration, **(B)** CP (creatine phosphate) concentration, **(C)** CK (creatine phosphokinase) concentration, **(D)** reaction temperature, and **(E)** plasmid concentration for sfGFP synthesis. Values show means with error bars representing standard deviations (s.d.) of at least three independent experiments.

system described here is capable of synthesizing ∼50 μg/ml of sfGFP in a 5 h batch reaction, which is a ∼55-fold increase in protein yields as compared to the initial yield of 0.91 ± 0.12 μg/ml. Our yield fits to the range of the yields achieved by other eukaryotic CFPS platforms, normally ranging from several to tens of micrograms protein per milliliter in batch mode reactions (Carlson et al., 2012). Such CFPS platforms include, for example, the insect system (25 μg/ml) (Tarui et al., 2001), the CHO system (∼50 μg/ml) (Brödel et al., 2014), and the *S. cerevisiae* system (<10 μg/ml) (Hodgman and Jewett, 2013). While we successfully established the *P. pastoris*-based CFPS platform, more efforts are still needed in the future to further improve the overall protein yields through upstream and downstream engineering strategies, for example, strain engineering (e.g., using nuclease-deficient strains), bioprocess engineering (e.g., high cell density cultivation in bioreactors), and reaction engineering (e.g., semi-continuous and continuous reaction formats). In addition, a recent report has demonstrated that enhancing ribosome content during cell growth could provide more active cell extracts, giving rise to higher protein yields in a CFPS system (Aw and Polizzi, 2019). Therefore, combining this strategy with above-mentioned solutions might further enhance the productivity of the *P. pastoris* CFPS system.

Having obtained an optimized CFPS system, we sought to demonstrate the utility of the platform. We chose two other eukaryotic proteins, luciferase and human serum albumin (HSA, a therapeutic protein), as our targets, which were expressed and analyzed by western blotting. The results indicated that both proteins were also successfully expressed and more than 90% of total proteins were soluble (**Figure S1**). As a result, the ability to use the newly developed *P. pastoris* CFPS system to express different proteins demonstrated its potential applications for the synthesis of more complex eukaryotic proteins such as membrane and glycosylated proteins.

CONCLUSIONS

In this study, we developed a robust and easy-to-use yeast-based CFPS system, which is derived from the protease-deficient strain *P. pastoris* SMD1163. In particular, we designed a suitable expression vector, identified a simple cell extract preparation protocol, and performed a physicochemical optimization. After the above systematic optimization, we were able to achieve sfGFP yields of ∼50 μg/ml in 5 h batch reactions. The high-yielding CFPS capacity, together with the heterologous protein expression capability of *P. pastoris*, makes it a valuable addition to the current existing CFPS platforms. Looking forward, we believe that this work will not only provide an alternative eukaryotic CFPS platform for the synthesis of "difficult-to-express" proteins like glycoproteins, but also set the stage for exciting, novel applications in biotechnology and synthetic biology.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

LZ, W-QL, and JL designed the experiments. LZ performed the experiments. LZ and JL analyzed the data and prepared the figures. JL conceived the project, supervised the research, and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the Shanghai Pujiang Program (18PJ1408000), the National Natural Science Foundation of China (31971348 and 31800720), and the Natural Science Foundation of Shanghai (19ZR1477200). JL also acknowledges the starting grant of ShanghaiTech University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00536/full#supplementary-material

REFERENCES

Ahmad, M., Hirz, M., Pichler, H., and Schwab, H. (2014). Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.* 98, 5301–5317. doi: 10.1007/s00253-014-5732-5

Aw, R., and Polizzi, K. M. (2019). Biosensor-assisted engineering of a high-yield *Pichia pastoris* cell-free protein synthesis platform. *Biotechnol. Bioeng.* 116, 656–666. doi: 10.1002/bit.26901

Baird, S. D., Turcotte, M., Korneluk, R. G., and Holcik, M. (2006). Searching for IRES. *RNA* 12, 1755–1785. doi: 10.1261/rna.157806

Bill, R. M. (2014). Playing catch-up with *Escherichia coli*: using yeast to increase success rates in recombinant protein production experiments. *Front. Microbiol.* 5:85. doi: 10.3389/fmicb.2014.00085

Brödel, A. K., Sonnabend, A., and Kubick, S. (2014). Cell-free protein expression based on extracts from CHO cells. *Biotechnol. Bioeng.* 111, 25–36. doi: 10.1002/bit.25013

Bundy, B. C., Hunt, J. P., Jewett, M. C., Swartz, J. R., Wood, D. W., Frey, D. D., et al. (2018). Cell-free biomanufacturing. *Curr. Opin. Chem. Eng.* 22, 177–183. doi: 10.1016/j.coche.2018.10.003

Carlson, E. D., Gan, R., Hodgman, C. E., and Jewett, M. C. (2012). Cell-free protein synthesis: applications come of age. *Biotechnol. Adv.* 30, 1185–1194. doi: 10.1016/j.biotechadv.2011.09.016

Darby, R. A. J., Cartwright, S. P., Dilworth, M. V., and Bill, R. M. (2012). Which yeast species shall I choose? *Saccharomyces cerevisiae* versus pichia pastoris (review). *Methods Mol. Biol.* 866, 11–23. doi: 10.1007/978-1-61779-770-5_2

De Schutter, K., Lin, Y. C., Tiels, P., Van Hecke, A., Glinka, S., Weber-Lehmann, J., et al. (2009). Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat. Biotechnol.* 27, 561–566. doi: 10.1038/nbt.1544

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. doi: 10.1021/acssynbio.8b00252

Failmezger, J., Scholz, S., Blombach, B., and Siemann-Herzberg, M. (2018). Cell-free protein synthesis from fast-growing *Vibrio natriegens*. *Front. Microbiol.* 9:1146. doi: 10.3389/fmicb.2018.01146

Gan, R., and Jewett, M. C. (2014). A combined cell-free transcription-translation system from *Saccharomyces cerevisiae* for rapid and robust protein synthesis. *Biotechnol. J.* 9, 641–651. doi: 10.1002/biot.201300545

Gao, W., Bu, N., and Lu, Y. (2019). Efficient incorporation of unnatural amino acids into proteins with a robust cell-free system. *Methods Protoc.* 2:16. doi: 10.3390/mps2010016

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. III., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Gilbert, W. V., Zhou, K., Butler, T. K., and Doudna, J. A. (2007). Cap-independent translation is required for starvation-induced differentiation in yeast. *Science* 317, 1224–1227. doi: 10.1126/science.1144467

Gleeson, M. A., White, C. E., Meininger, D. P., and Komives, E. A. (1998). Generation of protease-deficient strains and their use in heterologous protein expression. *Methods Mol. Biol.* 103, 81–94. doi: 10.1385/0-89603-421-6:81

Goering, A. W., Li, J., McClure, R. A., Thomson, R. J., Jewett, M. C., and Kelleher, N. L. (2017). *In vitro* reconstruction of nonribosomal peptide biosynthesis directly from DNA using cell-free protein synthesis. *ACS Synth. Biol.* 6, 39–44. doi: 10.1021/acssynbio.6b00160

Gurramkonda, C., Adnan, A., Gäbel, T., Lünsdorf, H., Ross, A., Nemani, S. K., et al. (2009). Simple high-cell density fed-batch technique for high-level recombinant protein production with *Pichia pastoris*: application to intracellular production of Hepatitis B surface antigen. *Microb. Cell. Fact.* 8:13. doi: 10.1186/1475-2859-8-13

Hamilton, S. R., Davidson, R. C., Sethuraman, N., Nett, J. H., Jiang, Y. W., Rios, S., et al. (2006). Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* 313, 1441–1443. doi: 10.1126/science.1130256

Henrich, E., Hein, C., Dötsch, V., and Bernhard, F. (2015). Membrane protein production in *Escherichia coli* cell-free lysates. *FEBS Lett.* 589, 1713–1722. doi: 10.1016/j.febslet.2015.04.045

Hodgman, C. E., and Jewett, M. C. (2013). Optimized extract preparation methods and reaction conditions for improved yeast cell-free protein synthesis. *Biotechnol. Bioeng.* 110, 2643–2654. doi: 10.1002/bit.24942

Hofbauer, R., Fessl, F., Hamilton, B., and Ruis, H. (1982). Preparation of a mRNA-dependent cell-free translation system from whole cells of *Saccharomyces cerevisiae*. *Eur. J. Biochem.* 122, 199–203. doi: 10.1111/j.1432-1033.1982.tb05867.x

Huang, A., Nguyen, P. Q., Stark, J. C., Takahasi, M. K., Donghia, N., Ferrante, T., et al. (2018). BioBits^{TM} explorer: a modular synthetic biology education kit. *Sci. Adv.* 4:eaat5105. doi: 10.1126/sciadv.aat5105

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. doi: 10.1038/msb.2008.57

Jin, X., and Hong, S. H. (2018). Cell-free protein synthesis for producing 'difficult-to-express' proteins. *Biochem. Eng. J.* 138, 156–164. doi: 10.1016/j.bej.2018.07.013

Karzbrun, E., Tayar, A. M., Noireaux, V., and Bar-Ziv, R. H. (2014). Programmable on-chip DNA compartments as artificial cells. *Science* 345, 829–832. doi: 10.1126/science.1255550

Kelwick, R., Webb, A. J., MacDonald, J. T., and Freemont, P. S. (2016). Development of a *Bacillus subtilis* cell-free transcription-translation system for prototyping regulatory elements. *Metab. Eng.* 38, 370–381. doi: 10.1016/j.ymben.2016.09.008

Khambhati, K., Bhattacharjee, G., Gohil, N., Braddick, D., Kulkarni, V., and Singh, V. (2019). Exploring the potential of cell-free protein synthesis for extending the abilities of biological systems. *Front. Bioeng. Biotechnol.* 7:248. doi: 10.3389/fbioe.2019.00248

Kigawa, T. (2010). Cell-free protein preparation through prokaryotic transcription-translation methods. *Methods Mol. Biol.* 607, 1–10. doi: 10.1007/978-1-60327-331-2_1

Kim, D. M., and Swartz, J. R. (1999). Prolonging cell-free protein synthesis with a novel ATP regeneration system. *Biotechnol. Bioeng.* 66, 180–188. doi: 10.1002/(SICI)1097-0290(1999)66:3<180::AID-BIT6>3.0.CO;2-S

Kim, H., Yoo, S. J., and Kang, H. A. (2015). Yeast synthetic biology for the production of recombinant therapeutic proteins. *FEMS Yeast Res.* 15, 1–16. doi: 10.1111/1567-1364.12195

Klein, D. J., Moore, P. B., and Steitz, T. A. (2004). The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* 10, 1366–1379. doi: 10.1261/rna.7390804

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44, 283–292. doi: 10.1016/0092-8674(86)90762-2

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15, 8125–8148. doi: 10.1093/nar/15.20.8125

Kozak, M. (1990). Evaluation of the fidelity of initiation of translation in reticulocyte lysates from commercial sources. *Nucleic Acids Res.* 18:2828. doi: 10.1093/nar/18.9.2828

Li, J., Lawton, T. J., Kostecki, J. S., Nisthal, A., Fang, J., Mayo, S. L., et al. (2016). Cell-free protein synthesis enables high yielding synthesis of an active multicopper oxidase. *Biotechnol. J.* 11, 212–218. doi: 10.1002/biot.201500030

Li, J., Wang, H., and Jewett, M. C. (2018a). Expanding the palette of *Streptomyces*-based cell-free protein synthesis systems with enhanced yields. *Biochem. Eng. J.* 130, 29–33. doi: 10.1016/j.bej.2017.11.013

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. doi: 10.1002/bit.26253

Li, J., Zhang, L., and Liu, W. (2018b). Cell-free synthetic biology for *in vitro* biosynthesis of pharmaceutical natural products. *Synth. Syst. Biotechnol.* 3, 83–89. doi: 10.1016/j.synbio.2018.02.002

Liu, W. Q., Zhang, L., Chen, M., and Li, J. (2019). Cell-free protein synthesis: recent advances in bacterial extract sources and expanded applications. *Biochem. Eng. J.* 141, 182–189. doi: 10.1016/j.bej.2018.10.023

Martin, R. W., Des Soye, B. J., Kwon, Y. C., Kay, J., Davis, R. G., Thomas, P. M., et al. (2018). Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat. Commun.* 9:1203. doi: 10.1038/s41467-018-03469-5

Mellitzer, A., Weis, R., Glieder, A., and Flicker, K. (2012). Expression of lignocellulolytic enzymes in *Pichia pastoris*. *Microb. Cell. Fact.* 11:61. doi: 10.1186/1475-2859-11-61

Min, S. E., Lee, K. H., Park, S. W., Yoo, T. H., Oh, C. H., Park, J. H., et al. (2016). Cell-free production and streamlined assay of cytosol-penetrating antibodies. *Biotechnol. Bioeng.* 113, 2107–2112. doi: 10.1002/bit.25985

Moore, S. J., Lai, H. E., Needham, H., Polizzi, K. M., and Freemont, P. S. (2017). *Streptomyces venezuelae* TX-TL – a next generation cell-free synthetic biology tool. *Biotechnol. J.* 12:1600678. doi: 10.1002/biot.201600678

Ni, Z., Zhou, X., Sun, X., Wang, Y., and Zhang, Y. (2008). Decrease of hirudin degradation by deleting the *KEX1* gene in recombinant *Pichia pastoris*. *Yeast* 25, 1–8. doi: 10.1002/yea.1542

Nie, Y., Huang, M., Lu, J., Qian, J., Lin, W., Chu, J., et al. (2014). Impacts of high β-galactosidase expression on central metabolism of recombinant *Pichia pastoris* GS115 using glucose as sole carbon source via ^{13}C metabolic flux analysis. *J. Biotechnol.* 187, 124–134. doi: 10.1016/j.jbiotec.2014.07.011

Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., et al. (2016). Rapid, low-cost detection of Zika virus using programmable biomolecular components. *Cell* 165, 1255–1266. doi: 10.1016/j.cell.2016.04.059

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

Perez, J. G., Stark, J. C., and Jewett, M. C. (2016). Cell-free synthetic biology: engineering beyond the cell. *Cold Spring Harb. Perspect. Biol.* 8:a023853. doi: 10.1101/cshperspect.a023853

Petrov, A. S., Bernier, C. R., Hsiao, C., Okafor, C. D., Tannenbaum, E., Stern, J., et al. (2012). RNA-magnesium-protein interactions in large ribosomal subunit. *J. Phys. Chem. B* 116, 8113–8120. doi: 10.1021/jp304723w

Raschmanová, H., Weninger, A., Glieder, A., Kovar, K., and Vogl, T. (2018). Implementing CRISPR-Cas technologies in conventional and non-conventional yeasts: current state and future prospects. *Biotechnol. Adv.* 36, 641–665. doi: 10.1016/j.biotechadv.2018.01.006

Sawasaki, T., Ogasawara, T., Morishita, R., and Endo, Y. (2002). A cell-free protein synthesis system for high-throughput proteomics. *Proc. Natl. Acad. Sci. U.S.A.* 99, 14652–14657. doi: 10.1073/pnas.232580399

Silverman, A. D., Karim, A. S., and Jewett, M. C. (2020). Cell-free gene expression systems: an expanding repertoire of applications. *Nat. Rev. Genet.* 21, 151–170. doi: 10.1038/s41576-019-0186-3

Sonnabend, A., Spahn, V., Stech, M., Zemella, A., Stein, C., and Kubick, S. (2017). Production of G protein-coupled receptors in an insect-based cell-free system. *Biotechnol. Bioeng.* 114, 2328–2338. doi: 10.1002/bit.26346

Stark, J. C., Huang, A., Nguyen, P. Q., Dubner, R. S., Hsu, K. J., Ferrante, T., et al. (2018). BioBits^{TM} bright: a fluorescent synthetic biology education kit. *Sci. Adv.* 4:eaat5107. doi: 10.1126/sciadv.aat5107

Sun, Z. Z., Hayes, C. A., Shin, J., Caschera, F., Murray, R. M., and Noireaux, V. (2013). Protocols for implementing an *Escherichia coli* based TX-TL cell-free expression system for synthetic biology. *J. Vis. Exp.* 79:e50762. doi: 10.3791/50762

Swank, Z., Laohakunakorn, N., and Maerkl, S. J. (2019). Cell-free gene-regulatory network engineering with synthetic transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* 116, 5892–5901. doi: 10.1073/pnas.1816591116

Swartz, J. R. (2018). Expanding biological applications using cell-free metabolic engineering: an overview. *Metab. Eng.* 50, 156–172. doi: 10.1016/j.ymben.2018.09.011

Takai, K., Sawasaki, T., and Endo, Y. (2010). Practical cell-free protein synthesis system using purified wheat embryos. *Nat. Protoc.* 5, 227–238. doi: 10.1038/nprot.2009.207

Tarui, H., Murata, M., Tani, I., Imanishi, S., Nishikawa, S., and Hara, T. (2001). Establishment and characterization of cell-free translation/glycosylation in insect cell (*Spodoptera frugiperda* 21) extract prepared with high pressure treatment. *Appl. Microbiol. Biotechnol.* 55, 446–453. doi: 10.1007/s002530000534

Thavarajah, W., Silverman, A. S., Verosloff, M. S., Kelley-Loughnane, N., Jewett, M. C., and Lucks, J. B. (2020). Point-of-use detection of environmental fluoride *via* a cell-free riboswitch-based biosensor. *ACS Synth. Biol.* 9, 10–18. doi: 10.1021/acssynbio.9b00347

van Nies, P., Westerlaken, I., Blanken, D., Salas, M., Menciá, M., and Danelon, C. (2018). Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun.* 9:1583. doi: 10.1038/s41467-018-03926-1

Várnai, A., Tang, C., Bengtsson, O., Atterton, A., Mathiesen, G., and Eijsink, V. G. H. (2014). Expression of endoglucanases in *Pichia pastoris* under control of the GAP promoter. *Microb. Cell. Fact.* 13:57. doi: 10.1186/1475-2859-13-57

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

Wang, S., Rong, Y., Wang, Y., Kong, D., Wang, P. G., Chen, M., et al. (2020). Homogeneous production and characterization of

recombinant *N*-GlcNAc-protein in *Pichia pastoris*. *Microb. Cell. Fact.* 19:7. doi: 10.1186/s12934-020-1280-0

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. doi: 10.1021/acssynbio.8b00222

Wilding, K. M., Hunt, J. P., Wilkerson, J. W., Funk, P. J., Swensen, R. L., Carver, W. C., et al. (2019). Endotoxin-free *E. coli*-based cell-free protein synthesis: pre-expression endotoxin removal approaches for on-demand cancer therapeutic production. *Biotechnol. J.* 14:1800271. doi: 10.1002/biot.201800271

Wu, M., Shen, Q., Yang, Y., Zhang, S., Qu, W., Chen, J., et al. (2013). Disruption of YPS1 and PEP4 genes reduces proteolytic degradation of secreted HSA/PTH in *Pichia pastoris* GS115. *J. Ind. Microbiol. Biotechnol.* 40, 589–599. doi: 10.1007/s10295-013-1264-8

Yamamoto, T., Shimizu, Y., Ueda, T., and Shiro, Y. (2010). Mg^{2+} dependence of 70 S ribosomal protein flexibility revealed by hydrogen/deuterium exchange and mass spectrometry. *J. Biol. Chem.* 285, 5646–5652. doi: 10.1074/jbc.M109.081836

Yang, Z., and Zhang, Z. (2018). Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: a review. *Biotechnol. Adv.* 36, 182–195. doi: 10.1016/j.biotechadv.2017.11.002

Zemella, A., Thoring, L., Hoffmeister, C., and Kubick, S. (2015). Cell-free protein synthesis: pros and cons of prokaryotic and eukaryotic systems. *ChemBioChem* 16, 2420–2431. doi: 10.1002/cbic.201500340

Zheng, X., Zhang, Y., Zhang, X., Li, C., Liu, X., Lin, Y., et al. (2019). Fhl1p protein, a positive transcription factor in *Pichia pastoris*, enhances the expression of recombinant proteins. *Microb. Cell. Fact.* 18:207. doi: 10.1186/s12934-019-1256-0

Zhuang, L., Huang, S., Liu, W. Q., Karim, A. S., Jewett, M. C., and Li, J. (2020). Total *in vitro* biosynthesis of the nonribosomal macrolactone peptide valinomycin. *Metab. Eng.* 60, 37–44. doi: 10.1016/j.ymben.2020.03.009

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

BRIEF RESEARCH REPORT
published: 24 June 2020
doi: 10.3389/fbioe.2020.00571

Multi-Enzyme Assembly on T4 Phage Scaffold

Jinny L. Liu, Daniel Zabetakis, Joyce C. Breger, George P. Anderson and Ellen R. Goldman*

Center for Bio/Molecular Science and Engineering, Naval Research Laboratory, Washington, DC, United States

Over the past two decades, various scaffolds have been designed and synthesized to organize enzyme cascades spatially for enhanced enzyme activity based on the concepts of substrate channeling and enhanced stability. The most bio-compatible synthetic scaffolds known for enzyme immobilization are protein and DNA nanostructures. Herein, we examined the utility of the T4 phage capsid to serve as a naturally occurring protein scaffold for the immobilization of a three-enzyme cascade: Amylase, Maltase, and Glucokinase. Covalent constructs between each of the enzymes and the outer capsid protein Hoc were prepared through SpyTag–SpyCatcher pairing and assembled onto phage capsids *in vitro* with an estimated average of 90 copies per capsid. The capsid-immobilized Maltase has a fourfold higher initial rate relative to Maltase free in solution. Kinetic analysis also revealed that the immobilized three-enzyme cascade has an 18-fold higher converted number of NAD^+ to NADH relative to the mixtures in solution. Our results demonstrate that the T4 phage capsid can act as a naturally occurring scaffold with substantial potential to enhance enzyme activity by spatially organizing enzymes on the capsid Hoc.

Keywords: phage scaffold, Hoc, SpyTag-SpyCatcher, amylase, maltase, glucokinase

OPEN ACCESS

Edited by:
Simon J. Moore, University of Kent, United Kingdom

Reviewed by:
Hung-En Lai, Victoria University of Wellington, New Zealand
Shen-Long Tsai, National Taiwan University of Science and Technology, Taiwan

****Correspondence:***
Jinny L. Liu
jinny.liu@nrl.navy.mil

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

Received: *04 March 2020*
Accepted: *12 May 2020*
Published: *24 June 2020*

Citation:
Liu JL, Zabetakis D, Breger JC, Anderson GP and Goldman ER (2020) Multi-Enzyme Assembly on T4 Phage Scaffold. Front. Bioeng. Biotechnol. 8:571. doi: 10.3389/fbioe.2020.00571

INTRODUCTION

Catalytic properties of enzymes are greatly affected by their surrounding microenvironment, particularly enzymes retained in a small area either by limited surface or restricted volume (Kuchler et al., 2016). Specifically, high concentration of enzymes in a confined environment are more stable than those free in solution. As a result, there is a preference for a folded state over an unfolded state and their close proximity allows them to execute a series of biocatalytic events more efficiently through substrate channeling (Miles et al., 1999; Zhang, 2011). Based on these concepts, various synthetic scaffolds aimed to spatially organize enzymes were designed and used for immobilizing enzymes to enhance their activity (Conrado et al., 2008; Dueber et al., 2009; Fu et al., 2014). Artificial multi-enzyme scaffolds have been utilized both *in vivo* and *in vitro* to improve product production (Siu et al., 2015; Ellis et al., 2019). Among these synthetic scaffolds, protein arrays and DNA nanostructures are the most biocompatible and have the potential to form the basis of a powerful platform to enhance multi-enzyme catalysis for biotechnology applications (Klein et al., 2019; Lim et al., 2019). Our team is researching naturally occurring scaffolds possessing the ability to spatially organize enzymes.

One benefit that naturally occurring phage scaffolds possess is that they are monodisperse and can be produced economically from bacteria hosts. In addition, most of the icosahedral phage capsids are composed of arrays of hexamers formed by capsid proteins, which can serve as a

platform for immobilizing enzymes spatially. T4 capsids are composed of the major capsid protein, gp23, and minor capsid protein, gp24, along with two accessory proteins, Hoc (highly antigenic outer capsid) and Soc (small outer capsid), and the portal protein, gp20 (Rao and Black, 2010). The capsids are homogeneous in size and structure and can only assemble inside the host bacteria with the expression of phage chaperone proteins and proteases. During the early infection of phage in bacteria, gp23 and gp24 assemble into a prohead shell wrapping around the core structure, later removed by protease. The vacant proheads allow DNA packaged inside through a packaging machinery, followed by the attachment of tail and tail fibers. Without DNA packaging, tail and tail fibers do not attach to the capsids.

The biggest hurdle for using phage scaffolds for displaying protein is that most capsids are quite rigid and can only display short peptides or few numbers of large proteins *in vivo* (Cardinale et al., 2012; Patel et al., 2017). However, T4 phage capsids allow one to overcome this difficulty by providing the needed flexibility to enable the display of large proteins through fusion with either of two outer capsid proteins, Hoc and Soc, without disrupting the capsid structure (Ren and Black, 1998). Moreover, both Hoc and Soc are dispensable for T4 phage propagation, their absence having no impact on T4 production. Soc proteins are closely associated with capsid proteins and are assembled next to each other, while Hoc is located at the center of hexamers, separated by both capsid proteins and Soc, based on a cryo-electron microscopy (EM) model (Fokine et al., 2004). Another advantage is that both Hoc and Soc fusions can be assembled onto Hoc and Soc deletion phage capsids as a scaffold either *in vitro* or *in vivo* (Rao and Black, 2010).

Our previously established work on characterization of T4 phage capsid structure using atomic force microscopy (AFM) showed that purified capsids are intact and stable (Archer and Liu, 2009; Robertson and Liu, 2012), and based on this work, we developed a new strategy to display a multi-enzyme cascade on phage capsids through SpyTag(St)/SpyCatcher(SC) pairing (Reddington and Howarth, 2015). Amylase (Aml), Maltase (Mal), and Glucokinase (GK) were selected to assemble onto phage scaffolds for catalytic analysis. These three enzymes are part of a four-enzyme biocatalytic pathway, which converts maltoheptaose into NADH and 6-phosphogluconolactone and the bio-catalytic assays for analyzing the enzyme activity have been well-established (Klein et al., 2019). The assembly of Hoc fusions onto phage scaffolds was conducted *in vitro* for better characterization. Enzyme–SC–St–Hoc fusions were purified and incubated in controlled ratios with T4 capsid. As observed in other scaffold systems, when assembled on the T4 capsid, the spatially organized assembled enzymes exhibit enhanced enzyme catalytic activity. The immobilized enzyme fusions on the phage scaffold showed enhanced biocatalytic activity for the number of NAD^+ converted to NADH per second up to 18-fold higher than the enzyme fusions free in solution. We have successfully demonstrated T4 icosahedral phage as a naturally occurring scaffold adaptable for multi-enzyme cascade assembly.

MATERIALS AND METHODS

Construction of Enzyme–SC Fusions and St–Hoc Plasmids

All the primer sequences used for cloning are listed in **Table S1** and gene maps and sequences are described in the **Supplementary Information**. The enzyme–SC fusions were cloned into pET28, while St–Hoc was inserted into pACYCduet according to Anderson et al. (2010) and Goldman et al. (2017). Details are provided in the **Supplementary Information**.

Preparation of Fusion Proteins

Protein was produced using a protocol similar to one that had been previously described with several modifications (Walper et al., 2014). The protocol for protein production and purification is provided in the **Supplementary Information**. Protein concentrations were determined using UV-Visible spectroscopy and the molar extinction coefficient predicted from the protein sequence.

Preparation of Tailless ΔHoc T4 Phage Scaffolds

E. coli, CR63 (F^- *supD60 lamB63*), or B40I (F^-*supD*), containing tRNA suppressor E, was used for growing del Hoc T4 phage (ΔHoc T4), a gift from Dr. Lindsay W. Black, Professor of Biochemistry Department, UMB. The phage capsid preparation was performed according to Liu et al. (2014). Details are provided in the **Supplementary Information**. The concentration of phage capsids was estimated based on 1 mg = 10^{13} capsids by taking into account the copy numbers of capsid proteins, Hoc, Soc, and portal proteins in one capsid (Fokine et al., 2004; Liu et al., 2014). The protein concentration was determined by the absorbance 280 nm and Bradford assay (Bio-Rad, Hercules, CA).

Assembly of Hoc–Enzyme Fusions Onto ΔHoc T4 Phage Capsids

The three model enzyme–Hoc fusions (Aml–Hoc; Mal–Hoc; and GK–Hoc) were first mixed at various molar ratios, 1:1:1; 1:4:1; 1:16:1. Enzyme–Hoc mixtures were subsequently added to phage capsids at ratios of total enzyme to phage capsid of 60:1 or 40:1. The same amount of enzyme–Hoc fusions was also prepared without added phage capsids. The mixtures were incubated at 22°C for 2 h or overnight (16 h). The assembly buffer consisted of PBS supplemented with 10 mM $MgCl_2$. To assess assembly, after incubation, the reactions were loaded onto 1% agarose and passed through 100 V for 2 h in 1× Tris–Acetate–EDTA buffer. The scaffolds and proteins were then visualized with GelCode blue safe protein stain (Thermo Fisher Scientific).

Measurements of Enzymatic Activity

For assessing Maltase activity, the production of 4-Nitrophenol (absorbance at 405 nm) as a function of time was recorded for 3–4 h using a Tecan Infinite M1000 plate reader immediately after the addition of the substrate, 10 mM 4-Nitrophenyl-α-D-glucopyranoside to a mixture of Maltase or Mal-Hoc fusions in 250 mM HEPES (pH 7.4) buffer. For measuring the three-enzyme cascade activity, glucose-6-phopsphate dehydrogenase

was included in excess in the reaction mixture and the rate was measured as NAD^+ conversion to NADH (absorbance at 340 nm) as a function of time after the addition of Maltoheptaose. The reaction mixture included 5 mM ATP, 1 mM NAD^+, and 10 U of glucose-6-phosphate dehydrogenase (G8404-2KU) in 250 mM HEPES (pH 7.4). The reaction schemes are described in more detail in **Figure S1**. The initial rate was measured as the slope after conversion of absorbance to nM/s, and V_{max}, K_{cat}, and K_m were calculated using Michaelis-Menton kinetics. Measurements were performed in triplicate. Chemicals were purchased from Millipore Sigma (St. Louis, MO), unless otherwise mentioned. Error bars were standard deviation derived from either biological replicates or technical triplicates.

RESULTS

Demonstration of Purified Enzyme–Hoc Assembled on Phage Capsids

Wild-type (WT) T4 capsids are composed of mainly gp23, T4 capsid protein, which forms an array of hexamers, and two accessory outer capsid proteins (Hoc and Soc) as indicated in **Figure 1A** along with the minor capsid protein and portal protein (not shown in **Figure 1A**; Fokine et al., 2004). Although proteins can be displayed onto the capsids through both Hoc and Soc fusions, we choose Hoc for this work, as there is more intermolecular space for enzyme fusions immobilized onto phage capsids. Future studies could also examine Soc–enzyme fusions to evaluate what effects on the enzymatic activity are obtained when the intermolecular spacing of the fusion enzymes is lessened.

The enzyme–SC and St–Hoc were covalently attached *in vivo* to produce enzyme–SC–St–Hoc (a.k.a. enzyme–Hoc) fusions with a 6× His tag by co-expressing enzyme–SC and St–Hoc in the same *E. coli* host. The fusion was purified via Ni-Sepharose resins, and separated from un-fused enzyme–SC through size exclusion chromatography (Goldman et al., 2017). Larger enzyme–Hoc fusions eluted earlier than smaller enzyme–SC fusions from the size-exclusion column (**Figure S2**). This purified material migrated slower on SDS-PAGE gel since the enzyme–Hoc was 40 kDa larger than enzyme–SC (**Figure S3**). Tailless ΔHoc T4 capsids were produced from ΔHoc T4 phage with a yield of 1–3 mg/L in the presence of 9-aminoacridine, which blocks genomic DNA packaged into the capsids and thus results in no tail attachment to the capsids (Schaerli and Kellenberger, 1980). They show faster electrophoretic mobility than WT T4 capsids as indicated in **Figure 1C** (lanes 4 vs. 6). The addition of Mal–Hoc onto ΔHoc T4 capsids at various molar ratios slows down their mobility and the Mal–Hoc assembled T4 capsids at ratios over 60–1 (Mal–Hoc: capsid) all maintain the same mobility (**Figure 1B**), suggesting that the numbers of Mal–Hoc assembled on the phage capsids reached maximum between 60 and 120 (∼90 in average; Lanes 4 and 5 in **Figure 1B**). A darker band with the same position of free Mal–Hoc appears in the ratio of 120–1 (Lane 5 in **Figure 1B**), indicating that free Mal–Hoc exists following the binding reaction, while there is much less free Mal–Hoc in the ratio of 60–1 in **Figure 1B**. Likewise, the samples with the ratio of 60–1 for the total enzymes to phage capsids showed the same effect on the electrophoretic mobility as indicated in **Figure 1C**. Based on these results, we constructed the enzyme assembly onto phage scaffolds at ratios of 40–1 for Maltase alone, while 60–1 was used for the three enzyme mixtures

TABLE 1 | The initial rate for Mal–Hoc fusion with phage scaffold.

Δ Hoc T4 capsids (nM)	Mal–Hoc (nM)	Initial rate (nM/s)*
0	1.80	$(4.40 \pm 1.12) \times 10^2$
0.0450	1.80	$(1.73 \pm 0.16) \times 10^3$
0	0.90	$(4.66 \pm 0.50) \times 10^2$
0.0225	0.90	$(1.120 \pm 0.16) \times 10^3$
0	0.45	$(5.54 \pm 0.77) \times 10^2$
0.0113	0.45	$(8.66 \pm 1.09) \times 10^2$

**Error bars represent standard deviation (STDEV) derived from technical triplicates.*

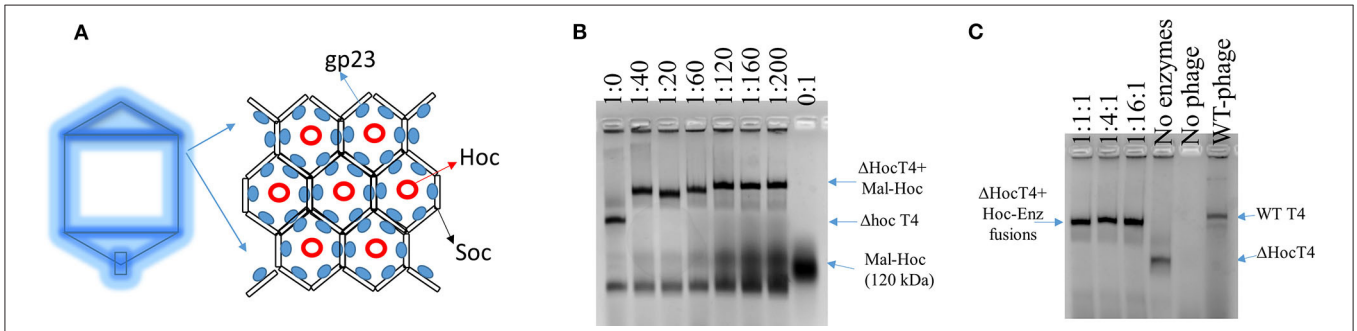


FIGURE 1 | Scheme of T4 capsid hexamers and the assembly of Hoc–enzyme fusions onto ΔHoc T4 phages. **(A)** T4 capsid protein, gp23, forms hexamers and Hoc is located at the center of the hexamers. **(B)** The agarose image of mixtures of phage scaffolds and Maltase-Hoc fusions at a series of ratios. Ratios are given above each lane as phage scaffold: Mal–Hoc. **(C)** The agarose image of mixtures of the phage scaffolds and three Hoc–enzyme fusions, Aml–Hoc, Mal–Hoc, and GK–Hoc, with the total sum concentration 60 nM, which have a 60:1 ratio for the total enzyme fusions to phage scaffolds. The stoichiometric ratios for Aml–Hoc, Mal–Hoc, and GK–Hoc are 1:1:1, 1:4:1, and 1:16:1 as indicated. The no enzyme lane contains only the ΔHoc phage, the no phage lane contains only Hoc–enzyme fusions (1:16:1), and wild-type (WT) phage is also shown. Protein bands were visualized by staining gels with GelCode blue safe protein dye.

TABLE 2 | Kinetics measurement for three enzyme–Hoc fusions vs. non-fusions.

Enzymes	#Ratio (Aml:Mal:GK)	V_{max} (nM/s)	K_m (nM)	*K_{cat} (1/s)	K_{cat}/K_m (1/M s)
Hoc–enzyme fusions	1:1:1	6.59 ± 1.56	(5.88 ± 3.31) × 10^4	0.10 ± 0.03	(1.77 ± 0.79) × 10^3
Enzyme alone	1:1:1	4.33 ± 1.21	(3.77 ± 1.59) × 10^5	0.87 ± 0.24	(2.30 ± 1.52) × 10^3

#The total concentration is 60 nM.

**The number of NAD^+ converted to NADH per second by the three-enzyme cascade. Error bars are STDEV derived from technical triplicates.*

TABLE 3 | Kinetic measurements for three enzyme–Hoc fusions with phage scaffolds.

ΔHocT4 capsids (nM)	#Ratio (A–Hoc:M–Hoc:GK–Hoc)	V_{max} (nM/s)	K_m (nM)	*K_{cat} (1/s)	K_{cat}/K_m (1/M s)
0	1:1:1	6.31 ± 0.12	(3.00 ± 0.71) × 10^5	0.32 ± 0.01	(1.07 ± 0.14) × 10^3
1.00	1:1:1	78.50 ± 4.9	(2.50 ± 0.20) × 10^6	3.93 ± 0.25	(1.57 ± 1.25) × 10^3
0	1:4:1	9.40 ± 2.04	(1.61 ± 0.59) × 10^5	0.24 ± 0.05	(1.63 ± 0.91) × 10^3
1.00	1:4:1	170.00 ± 32.50	(2.31 ± 0.23) × 10^6	4.26 ± 0.81	(1.83 ± 0.17) × 10^3

#The total concentration is 60 nm.

**The number of NAD^+ converted to NADH per second by the three-enzyme cascade. Error bars are STDEV derived from biological replicates.*

in **Tables 2**, **3** to ensure better basal K_{cat} from free enzyme–Hoc fusions for comparison with the assembled ones. Both ratios exhibited minimal background free enzyme–Hoc activity.

Higher Initial Rate for the Immobilized Mal–Hoc on Phage Capsids

Since Maltase is the rate-limiting step in the enzyme cascade, we measured the initial rate of Mal–Hoc (Klein et al., 2019). The enzyme activity of immobilized Mal–Hoc vs. free Mal–Hoc in solution was measured using 4-Nitrophenyl-α-D-glucopyranoside as a substrate and the yield of the end product of 4-Nitrophenol (4-NP) was monitored by measuring its absorbance at 405 nm. The assembly of Mal–Hoc and phage capsids was conducted at 22°C for 16 h before measuring Mal activity. After prolonged incubation, the immobilized Mal–Hoc on phage capsids still retains its activity and exhibit fourfold higher initial rate relative to the free Mal–Hoc in solution (**Figure S4** and **Table 1**).

Kinetic Analysis for Enzyme–Hoc Fusions vs. Unfused Enzymes

The kinetics for the three enzyme–Hoc fusions and the three unfused enzymes was measured at a mixing ratio of 1:1:1 of the three enzymes with the total concentration of 0.06 μM. The V_{max} was slightly higher for enzyme–Hoc, while K_m and K_{cat} were smaller than enzyme alone (**Table 2**). Although the K_{cat} was slower in enzyme–Hoc, a lower concentration of substrate was needed for enzyme–Hoc to reach 1/2 V_{max}, resulting in similar catalytic efficiency (K_{cat}/K_m) for unfused enzymes.

Kinetic Analysis for Immobilized Enzyme–Hoc vs. Enzyme–Hoc Free in Solution

The catalysis of a three-enzyme cascade immobilized vs. the mixture of three enzyme fusions free in solution was measured and compared. The immobilized enzymes on phage scaffolds consistently showed higher K_{cat} (**Table 3**) with more enzymatic efficiency (K_{cat}/K_m) for the conversion of NAD^+ to NADH. Specifically, assembled enzyme–Hoc fusions with the mixture of three enzymes at a ratio of 1:1:1 appears to have at least 12-fold higher numbers of conversion of NAD^+ to NADH relative to corresponding enzyme fusions free in solution (**Table 3**). Likewise, assembled enzyme–Hoc fusions at a mixing ratio of 1:4:1 without changing the total enzyme–Hoc concentration also exhibit an increase in K_{cat} of 18-fold relative to the corresponding ones off the scaffolds in solution. However, the sample with a 1:4:1 molar ratio of Aml:Mal:GK did not show better catalytic efficiency than the one with a 1:1:1 ratio of the three enzymes.

DISCUSSION

A Naturally Occurring Phage Scaffold for Assembly of Spatially Organized Enzymes

Reconstructed T4 capsid from cryo-electron microscopy (EM) images revealed that Hoc was dumbbell-shaped with a distance, 14 nm, apart from each other, while Soc was rod-shaped closely associated with each other with a few nanometers distance apart (Fokine et al., 2004). Therefore, the distance between two enzyme–Hoc fusions is within the range where substrate channeling can occur between two co-immobilized enzymes within a cascade (Fu et al., 2014; Lim et al., 2019). Consistent with this prediction, our results showed that the immobilized enzyme–Hoc fusions enhanced catalytic activity in comparison to enzymes free in bulk solution. The assembly of Hoc onto phage capsid is mainly driven by the interaction of its C-terminus with T4 capsid protein through electrostatic, Van der Waals, hydrogen bond forces (Sathaliyawala et al., 2010). According to the cryo-EM model, there are ∼155 positions for Hoc–enzyme fusions on each ΔHoc capsid; however, the decorated numbers on the capsids depend on the size or charges of the Hoc fusions (Fokine et al., 2004). The addition of enzyme–Hoc fusions significantly changes the electrophoretic mobility of the assembled phage scaffolds, shifting to the same

position as WT T4 (lanes 4 and 6 in **Figure 1C**). Our previous study showed that the loss of Hoc on the capsids (ΔHoc T4) resulted in a lower pI and more negative zeta potential, but no significant change in size (Robertson et al., 2012). Consistently, our results show that ΔHoc T4 migrated faster toward the positive electrode. Based on the estimate from the band density, ∼90 copies on average of enzyme–Hoc were decorated on the capsid; therefore, every immobilized enzyme should have at least two neighboring enzymes on the capsids. Having neighboring enzymes likely facilitates the catalysis via substrate channeling. Enhanced stability could also contribute to higher catalytic activity. Thus, we have demonstrated that T4 phage capsids provide a naturally occurring scaffold to immobilize enzymes for efficient biocatalysis.

Comparable Enzymatic Activity of Enzyme–Hoc Fusion vs. Enzyme Alone

St and SC are two partners of the same domain of *Streptococcus pyogenes* fibronectin-binding protein (FbaB) and the pair is covalently attached through spontaneous amide bond formation between Asp and Lys in the respective partner (Zakeri et al., 2012; Reddington and Howarth, 2015). We individually fused St onto the N-terminus of Hoc and SC to the C-terminus of enzyme to produce enzyme–SC–St–Hoc fusions *in vivo* for subsequent one-step purification. The advantage of St–SC pairing allows the separate expression of enzyme and Hoc in the host and avoids the direct expression of large Hoc fusions, which often have the propensity for improper folding and low yields during the production. By this method, we were able to obtain reasonable yields (∼5–10 mg/L) for the enzyme–Hoc fusions with the enzymatic activity comparable to enzymes alone.

Demonstration of Enhanced Activity for Immobilized Single Enzyme and the Three-Enzyme Cascade

The kinetics were measured after 2 h of assembly and the measurements were finished within 4 h to ensure that the majority of the enzyme fusions were still active. Although the specific molecular mechanism for the three-enzyme cascade is not clearly identified, we attribute the enhancement of the immobilized enzymes mainly to substrate channeling (Miles et al., 1999; Klein et al., 2019). Mal is the rate-limiting step of the three-enzyme cascade (Klein et al., 2019); therefore, we also specifically measured the initial rate for immobilized Mal–Hoc after 16 h of assembly. Our results show that immobilization on phage scaffolds stabilizes the Mal–Hoc after prolonged incubation (**Figure S4**). Therefore, the enhancement of the immobilized enzyme activity on phage scaffolds is the sum of substrate channeling and improved stability. Although the scaffolded enzyme–Hoc fusions have higher K_m with reduced binding affinity possibly due to the interference of enzyme active sites while attaching to the scaffolds, the greater benefit of substrate channeling compensates for the disadvantage and results in a higher conversion rate of NAD^+ to NADH.

In our experiments, the three enzyme–Hoc constructs were incorporated into ΔHoc T4 *in vitro*. This type of *in vitro* synthetic biology has been exploited to circumvent drawbacks to cell-based approaches, such as toxicity and the presence of competing pathways (Koch et al., 2018; Sperl and Sieber, 2018). In addition, *in vitro* assembly allows better control over the stoichiometry of the enzymes in the multi-enzyme-scaffolded complex. However, the T4 system could also be adapted for *in vivo* display of scaffolded enzymes.

CONCLUSION

T4 phage scaffolds are composed of prolated protein shells containing arrays of hexamers of capsid protein decorated with Hoc at the center of the hexamers. In this work, we prepared enzyme–Hoc fusions through St–SC pairing and successfully immobilized a three-enzyme cascade, Aml, Mal, and GK, onto naturally occurring T4 phage scaffolds. The immobilized enzymes exhibit enhanced catalytic activity up to 18-fold relative to enzymes free in solution. Thus, we have successfully demonstrated an icosahedral T4 phage scaffold as a new platform for enzyme assembly to enhance biocatalysis applicable to variety of biotechnology applications.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JL designed and conducted the experiments and wrote the manuscript. DZ, JB, and GA performed the experiments and edited the manuscript. EG designed the experiments and wrote the manuscript All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by office of Naval Research through NRL 6.1 base fund.

ACKNOWLEDGMENTS

We greatly thank Drs. Lindsay W. Black at University of Maryland at Baltimore for the generous gift of T4 phage and Igor Medintz's insightful suggestions. We also thank Drs. Scott A. Walper and Kendrick D. Turner for the gift of plasmids containing the WT maltase, amylase, and glucokinase sequences.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00571/full#supplementary-material

REFERENCES

Anderson, J., Dueber, J. E., Leguia, M., Wu, G. C., Goler, J. A., Arkin, A. P., et al. (2010). BglBricks: a flexible standard for biological part assembly. *J. Biol. Eng.* 4:1. doi: 10.1186/1754-1611-4-1

Archer, M. J., and Liu, J. L. (2009). Bacteriophage t4 nanoparticles as materials in sensor applications: variables that influence their organization and assembly on surfaces. *Sensors (Basel)* 9, 6298–6311. doi: 10.3390/s90806298

Cardinale, D., Carette, N., Michon, T. (2012). Virus scaffolds as enzyme nano-carriers. *Trends Biotech.* 30, 369–376. doi: 10.1016/j.tibtech.2012.04.001

Conrado, R. J., Varner, J. D., and DeLisa, M. P. (2008). Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Curr. Opin. Biotechnol.* 19, 492–499. doi: 10.1016/j.copbio.2008.07.006

Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., et al. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* 27, 753–759. doi: 10.1038/nbt.1557

Ellis, G. A., Klein, W. P., Lasarte-Aragonés, G., Thakur, M., Walper, S. A., and Medintz, I. L. (2019). Artificial multienzyme scaffolds: pursuing *in vitro* substrate channeling with an overview of current progress. *ACS Catal.* 9, 10812–10869. doi: 10.1021/acscatal.9b02413

Fokine, A., Chipman, P. R., Leiman, P. G., Mesyanzhinov, V. V., Rao, V. B., and Rossmann, M. G. (2004). Molecular architecture of the prolate head of bacteriophage T4. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6003–6008. doi: 10.1073/pnas.0400444101

Fu, J., Yang, Y. R., Johnson-Buck, A., Liu, M., Liu, Y., Walter, N. G., et al. (2014). Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nat. Nanotechnol.* 9, 531–536. doi: 10.1038/nnano.2014.100

Goldman, E. R., Broussard, A., Anderson, G. P., and Liu, J. L. (2017). Bglbrick strategy for the construction of single domain antibody fusions. *Heliyon* 3:e00474. doi: 10.1016/j.heliyon.2017.e00474

Klein, W. P., Thomsen, R. P., Turner, K. B., Walper, S. A., Vranish, J., Kjems, J., et al. (2019). Enhanced catalysis from multienzyme cascades assembled on a DNA origami triangle. *ACS Nano* 13, 13677–13689. doi: 10.1021/acsnano.9b05746

Koch, M., Faulon, J.-L., and Borkowski, O. (2018). Models for cell-free synthetic biology: make prototyping easier, better, and faster. *Front. Bioeng. Biotechnol.* 6:182. doi: 10.3389/fbioe.2018.00182

Kuchler, A., Yoshimoto, M., Luginbuhl, S., Mavelli, F., and Walde, P. (2016). Enzymatic reactions in confined environments. *Nat. Nanotechnol.* 11, 409–420. doi: 10.1038/nnano.2016.54

Lim, S., Jung, G. A., Glover, D. J., and Clark, D. S. (2019). Enhanced enzyme activity through scaffolding on customizable self-assembling protein filaments. *Small* 15:e1805558. doi: 10.1002/smll.201805558

Liu, J. L., Dixit, A. B., Robertson, K. L., Qiao, E., and Black, L. W. (2014). Viral nanoparticle-encapsidated enzyme and restructured DNA for cell delivery and gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13319–13324. doi: 10.1073/pnas.1321940111

Miles, E. W., Rhee, S., and Davies, D. R. (1999). The molecular basis of substrate channeling. *J. Biol. Chem.* 274, 12193–12196. doi: 10.1074/jbc.274.18.12193

Patel, A. N., Anne, A., Chovin, A., Demaille, C., Grelet, E., Michon, T., et al. (2017). Scaffolding of enzymes on virus nanoarrays: effects of confinement and virus organization on biocatalysis. *Small* 13:1603163. doi: 10.1002/smll.201603163

Rao, V. B., and Black, L. W. (2010). Structure and assembly of bacteriophage T4 head. *Virol. J.* 7:356. doi: 10.1186/1743-422X-7-356

Reddington, S. C., and Howarth, M. (2015). Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher. *Curr. Opin. Chem. Biol.* 29, 94–99. doi: 10.1016/j.cbpa.2015.10.002

Ren, Z., and Black, L. W. (1998). Phage T4 SOC and HOC display of biologically active, full-length proteins on the viral capsid. *Gene* 215, 439–444. doi: 10.1016/S0378-1119(98)00298-4

Robertson, K., Furukawa, Y., Underwood, A., Black, L., and Liu, J. L. (2012). Deletion of the Hoc and Soc capsid proteins affects the surface and cellular uptake properties of bacteriophage T4 derived nanoparticles. *Biochem. Biophys. Res. Commun.* 418, 537–540. doi: 10.1016/j.bbrc.2012.01.061

Robertson, K. L., and Liu, J. L. (2012). Engineered viral nanoparticles for flow cytometry and fluorescence microscopy applications. *Wiley Interdiscipl. Rev. Nanomed. Nanobiotechnol.* 4, 511–524. doi: 10.1002/wnan.1177

Sathaliyawala, T., Islam, M. Z., Li, Q., Fokine, A., Rossmann, M. G., and Rao, V. B. (2010). Functional analysis of the highly antigenic outer capsid protein, Hoc, a virus decoration protein from T4-like bacteriophages. *Mol. Microbiol.* 77, 444–455. doi: 10.1111/j.1365-2958.2010.07219.x

Schaerli, C., and Kellenberger (1980). Head maturation pathway of bacteriophages T4 and T2. V. Maturable epsilon-particle accumulating an acridine-treated bacteriophage T4-infected cells. *J. Virol.* 33, 830–844. doi: 10.1128/JVI.33.2.830-844.1980

Siu, K. H., Chen, R. P., Sun, Q., Chen, L., Tsai, S. L., and Chen, W. (2015). Synthetic scaffolds for pathway enhancement. *Curr. Opin. Biotechnol.* 36, 98–106. doi: 10.1016/j.copbio.2015.08.009

Sperl, J. M., and Sieber, V. (2018). Multienzyme cascade reactions—status and recent advances. *ACS Catal.* 8, 2385–2396. doi: 10.1021/acscatal.7b03440

Walper, S. A., Battle, S. R., Lee, P. A. B., Zabetakis, D., Turner, K. B., Buckley, P. E., et al. (2014). Thermostable single domain antibody-maltose binding protein fusion for *Bacillus anthracis* spore protein BclA detection. *Anal. Biochem.* 447, 64–73. doi: 10.1016/j.ab.2013.10.031

Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., et al. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. U.S.A.* 109, E690–697. doi: 10.1073/pnas.1115485109

Zhang, Y. H. (2011). Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* 29, 715–725. doi: 10.1016/j.biotechadv.2011.05.020

Disclaimer: Distribution a from Naval Research Laboratory: approved for public release, distribution is unlimited.

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

REVIEW
published: 22 July 2020
doi: 10.3389/fbioe.2020.00863

Emerging Methods for Efficient and Extensive Incorporation of Non-canonical Amino Acids Using Cell-Free Systems

*Yang Wu[1,2], Zhaoguan Wang[1,2], Xin Qiao[1,2], Jiaojiao Li[1,2], Xiangrong Shu[3] and Hao Qi[1,2]**

[1] School of Chemical Engineering and Technology, Tianjin University, Tianjin, China, [2] Key Laboratory of Systems Bioengineering, Ministry of Education, Tianjin University, Tianjin, China, [3] Department of Pharmacy, Tianjin Huanhu Hospital, Tianjin, China

Cell-free protein synthesis (CFPS) has emerged as a novel protein expression platform. Especially the incorporation of non-canonical amino acids (ncAAs) has led to the development of numerous flexible methods for efficient and extensive expression of artificial proteins. Approaches were developed to eliminate the endogenous competition for ncAAs and engineer translation factors, which significantly enhanced the incorporation efficiency. Furthermore, *in vitro* aminoacylation methods can be conveniently combined with cell-free systems, extensively expanding the available ncAAs with novel and unique moieties. In this review, we summarize the recent progresses on the efficient and extensive incorporation of ncAAs by different strategies based on the elimination of competition by endogenous factors, translation factors engineering and extensive incorporation of novel ncAAs coupled with *in vitro* aminoacylation methods in CFPS. We also aim to offer new ideas to researchers working on ncAA incorporation techniques in CFPS and applications in various emerging fields.

Keywords: cell-free synthetic biology, cell-free protein synthesis, non-canonical amino acids, competition elimination, *in vitro* aminoacylation

OPEN ACCESS

Edited by:
Yuan Lu,
Tsinghua University, China

Reviewed by:
Seok Hoon Hong,
Illinois Institute of Technology, United States
Dong-Myung Kim,
Chungnam National University, South Korea

***Correspondence:**
Hao Qi
haoq@tju.edu.cn

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

Received: *27 April 2020*
Accepted: *06 July 2020*
Published: *22 July 2020*

Citation:
Wu Y, Wang Z, Qiao X, Li J, Shu X and Qi H (2020) Emerging Methods for Efficient and Extensive Incorporation of Non-canonical Amino Acids Using Cell-Free Systems. Front. Bioeng. Biotechnol. 8:863. doi: 10.3389/fbioe.2020.00863

INTRODUCTION

Cell-free protein synthesis (CFPS) has emerged as an effective method for the production of recombinant proteins in specialty applications. By eliminating the cell membrane barrier and cell-viability constraint, CFPS offers several benefits over *in vivo* protein expression (Liu et al., 2019). Firstly, with the open nature of CFPS, almost any molecule can be manipulated precisely in the system for different research purposes, especially molecules whose incorporation is limited by inefficient transport across the cell membrane (Silverman et al., 2019). Secondly, by being able to disregard cell viability, toxic reagents and difficult to express proteins can be employed in CFPS and even some not biocompatible reaction conditions can be applied (Lu, 2017). Finally, without reproducible cells, biosafety can be guaranteed because artificial genes cannot pollute the environment through cells.

Basically, there are two main CFPS platforms: the PURE system (i.e., protein synthesis using purified recombinant elements), and the cell extract system. In the PURE system, all components of the transcription and translation apparatus are purified from cells individually and assembled

into a well-defined CFPS system. Although all components can be defined at precise concentration, the tedious purification steps make the PURE platform much more expensive than the cell extract system (Shimizu et al., 2001). Many efforts have been made recently to reduce the costs and labor, such as one-pot purification methods and purification from fewer fusion plasmids (Wang H.H. et al., 2012; Shepherd et al., 2017; Villarreal et al., 2018; Lavickova and Maerkl, 2019). However, partial component control and modularity may be lost in these approaches. The other system relies on non-defined cell extracts. The crude cell extract is separated by lysing cells, so it contains all the native intracellular translation components. Recombinant proteins are synthesized via cell extract based CFPS with the supplementation of additives, such as energy substrates, NTPs, T7 RNA polymerase, amino acids, and salts (Dopp et al., 2019; **Figure 1**). Due to the simple preparation, the cell extract platform is much cheaper and convenient. Additionally, with the help of ancillary translational factors in the cell extract, this platform also has higher protein yields (Karim and Jewett, 2016). Taken together, both CFPS systems are useful platforms for different applications.

Incorporating ncAAs into proteins is an emerging biological research area with fundamental science and engineering benefits. In fundamental science, lots of questions are being answered by ncAA techniques, such as labeling proteins by isotopic or fluorescent ncAAs, and immobilization of protein using ncAAs with special side chains (Narumi et al., 2018). Post-translational protein modifications (PTM) are difficult to study due to their rapidly shifting levels in the cell. With PTM-mimicking side-chains of ncAAs, high amounts of homologous PTM proteins can be synthesized for investigation (Park et al., 2011; Rogerson et al., 2015; Kightlinger et al., 2019). In engineering applications, a growing number of artificial protein applications are also emerging, including antibody-drug conjugates (Si et al., 2016), virus-like particle drug conjugates (Bundy et al., 2008), active protein polymers (Albayrak and Swartz, 2014), and screening of artificial enzymes (Ravikumar et al., 2015).

Over 230 ncAAs have been incorporated into proteins by *in vivo* or *in vitro* methods (Gfeller et al., 2013; Dumas et al., 2015). In living cells, an orthogonal amino-acyl tRNA synthetase/tRNA (aaRS/tRNA) pair is essential to precisely incorporate ncAAs into proteins. The orthogonality means that aaRS can only incorporate ncAAs at the specific tRNA and the tRNA can only be recognized by a corresponding aaRS (Hu et al., 2014). Recently, numerous ncAA aaRS/tRNA pairs were developed based on systems from archaea. For instance, tyrosine derivatives can be installed by *Methanococcus jannaschii* TyrRS/$tRNA^{Tyr}$ pair variants and lysine derivatives can be installed by variants of the *Methanosarcina mazei* or *Methanosarcina barkeri* PylRS/$tRNA^{Pyl}$ (Chin, 2017). However, due to great advantages over *in vivo* research, accelerated studies are concentrating on CFPS to incorporate ncAAs. Firstly, the concentration of ncAA and aaRS/tRNA could be conveniently improved for efficient incorporation without limitation by transport across the cell membrane. Secondly, the negative components, which influence efficient ncAA incorporation, can potentially be eliminated potentially in the open CFPS environment. Thirdly, the subsequent reaction can be performed *in situ*, avoiding tedious purification from cells (Italia et al., 2019). Finally, even toxic ncAAs, which cannot be employed in living cells, could be used in CFPS combined with *in vitro* aminoacylation methods (Kawakami et al., 2008; Haruna et al., 2014; Katoh et al., 2017a). Therefore, CFPS is a novel platform for ncAA incorporation with great potential.

However, ncAA incorporation in CFPS suffers from endogenous competition (**Figure 1**). In order to incorporate ncAAs at precise positions, a special codon should be reassigned. Both stop codons and sense codons have been reassigned in various studies (Zimmerman et al., 2014; Schinn et al., 2017; Stech et al., 2017). There is codon competition between ncAA complexes and endogenous biomolecules, which recognize the codon as a canonical amino acid or translation stop signal, resulting in a truncated or wrong protein instead of the ncAA protein (**Figure 1**). Fortunately, various strategies were developed to overcome competition at different levels, fully utilizing the advantages of CFPS (Lee K.-H. et al., 2016; Cui et al., 2018; Martin et al., 2018). We summarize and discuss these efficiency improvement efforts in this review. In addition to efficient ncAA incorporation, many novel ncAAs can only be used in CFPS using *in vitro* aminoacylation methods (Ohta et al., 2008; Katoh et al., 2017b; Katoh and Suga, 2018). These novel amino acids greatly expand the scope and flexibility of CFPS ncAA incorporation, which is also gaining increasing attention.

The purpose of this review is to give an overview of the emerging methods for improving the scope and efficiency of ncAA incorporation in CFPS. We focus on strategies to raise efficiency at the competition elimination and translation factors engineering. Moreover, the novel ncAAs and *in vitro* aminoacylation methods, that can only be used in CFPS and expand the range of available ncAAs, are given special attention. This review provides a reference for researchers to choose suitable ncAA incorporation techniques for CFPS and expand the research to more potential fields.

STRATEGIES FOR ELIMINATING COMPETITION

As shown in **Figure 1**, the competition for codons, that seriously affects the efficiency of ncAA incorporation, always exists in codon reassignment. Competitors are classified into two categories: peptide release factors and endogenous aminoacyl-tRNA, and methods for eliminating competition include genome engineering, protein elimination, tRNA manipulation, and amino acid replacement (**Figure 2**). Here, we discuss the challenges and opportunities of these current and emerging methods.

Genome Engineering

The amber codon (UAG), which is normally a translation termination signal recognized by peptide release factor 1 (RF1), is most commonly reassigned for ncAA incorporation. Deletion of the *prfA* gene encoding RF1 is a frequent strategy to improve the ncAA incorporation efficiency *in vivo* and *in vitro* (Johnson et al., 2011). Cell extracts from such strains have shown numerous

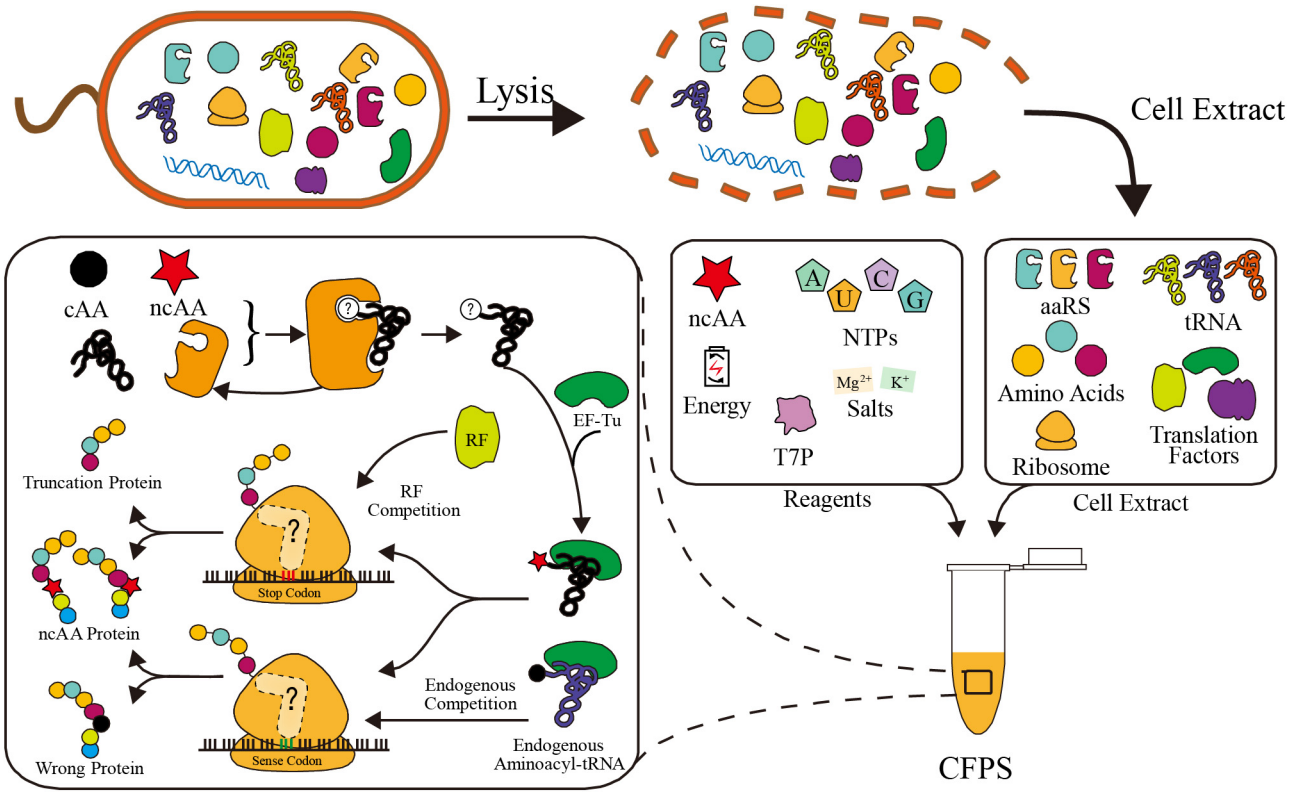


FIGURE 1 | Schematic of cell extract based CFPS system preparation and competitors in ncAA incorporation. In ribosome, the peptide release factor competes with ncAA aminoacyl-tRNA in stop codon reassignment. Endogenous aminoacyl-tRNAs compete with aminoacyl-tRNAs in sense codon reassignment. In aaRS, the canonical amino acid (cAA) may compete with ncAA in aminoacylation reaction.

advantages (Smolskaya et al., 2020). However, RF1 is essential for translation termination of over 300 genes in the genome, and the direct deletion of *prfA* could seriously affect cell growth before cell extract preparation, even leading to cell death (Chin, 2017). To address this problem, several detailed methods have been developed. These efforts and CFPS applications will be discussed in this section (**Figure 2A**).

The most direct method is the replacement of all amber codons with other stop codons, producing cells that do not need RF1, so that *prfA* can be knocked out without affecting cell viability. C321. ΔA is such an *E. coli* strain, whereby all of its 321 amber codons were replaced with ochre codons (UAA) (Lajoie et al., 2013). Without RF1 competition, ncAAs can be incorporated into proteins at high efficiency. For instance, phosphoserine was efficiently incorporated into human MEK1 kinase using CFPS based on a cell extract of C321.ΔA, in which robust production reached up to milligram quantities (Oza et al., 2015). By contrast, under RF1 competition in intracellular expression, only 25 μg of MEK1 containing phosphoserine was isolated from 1 liter of culture (Park et al., 2011). Recently, glycosylation enzymes were also successfully synthesized using this CFPS platform (Kightlinger et al., 2019). Furthermore, to improve the CFPS protein production capacity, great efforts were made to further improve C321. ΔA. Combined knockout of the genes *endA*, *gor*, *rne*, and *mazF* was shown to increase the yield of a model protein about fivefold, with up to 40 ncAAs (Martin et al., 2018). Subsequently, a C321. ΔA variant that can express T7 RNA polymerase from the chromosome was created and optimized. CFPS based on a cell extract of this strain yielded up to 70 μg/mL of a model protein with 40 ncAAs, demonstrating a highly productive one-pot CFPS system (Des Soye et al., 2019).

In addition to full replacement of amber codons, partial replacement has also been demonstrated, in which only the stop codons of some essential genes were replaced. The strains RFzero, B95.ΔA and rEc.E13.ΔA are representatives of this strategy. The RFzero strain did not contain any replaced stop codons in chromosome. Ingeniously, it was introduced the BCA7 plasmid harboring seven essential genes (*coaD*, *hda*, *hemA*, *mreC*, *murF*, *lolA*, and *lpxK*) with replaced stop codons for compensation. Subsequently, a chromosomal *prfA* deletion was carried out successfully, with only a minimal decrease of the cell growth rate (Mukai et al., 2010, 2011; Ohtake et al., 2012). Then, CFPS using a cell extract of the RFzero strain was verified to incorporate six ncAAs with high efficiency and productivity (Adachi et al., 2019). The strains B60.ΔA and B95.ΔA were developed from *E. coli* strain BL21(DE3) by Mukai et al. (2015) by respectively replacing 60 and 90 chromosomal amber codons before *prfA* deletion. Due to the rapid growth rates of B60.ΔA and B95.ΔA, a CFPS protocol

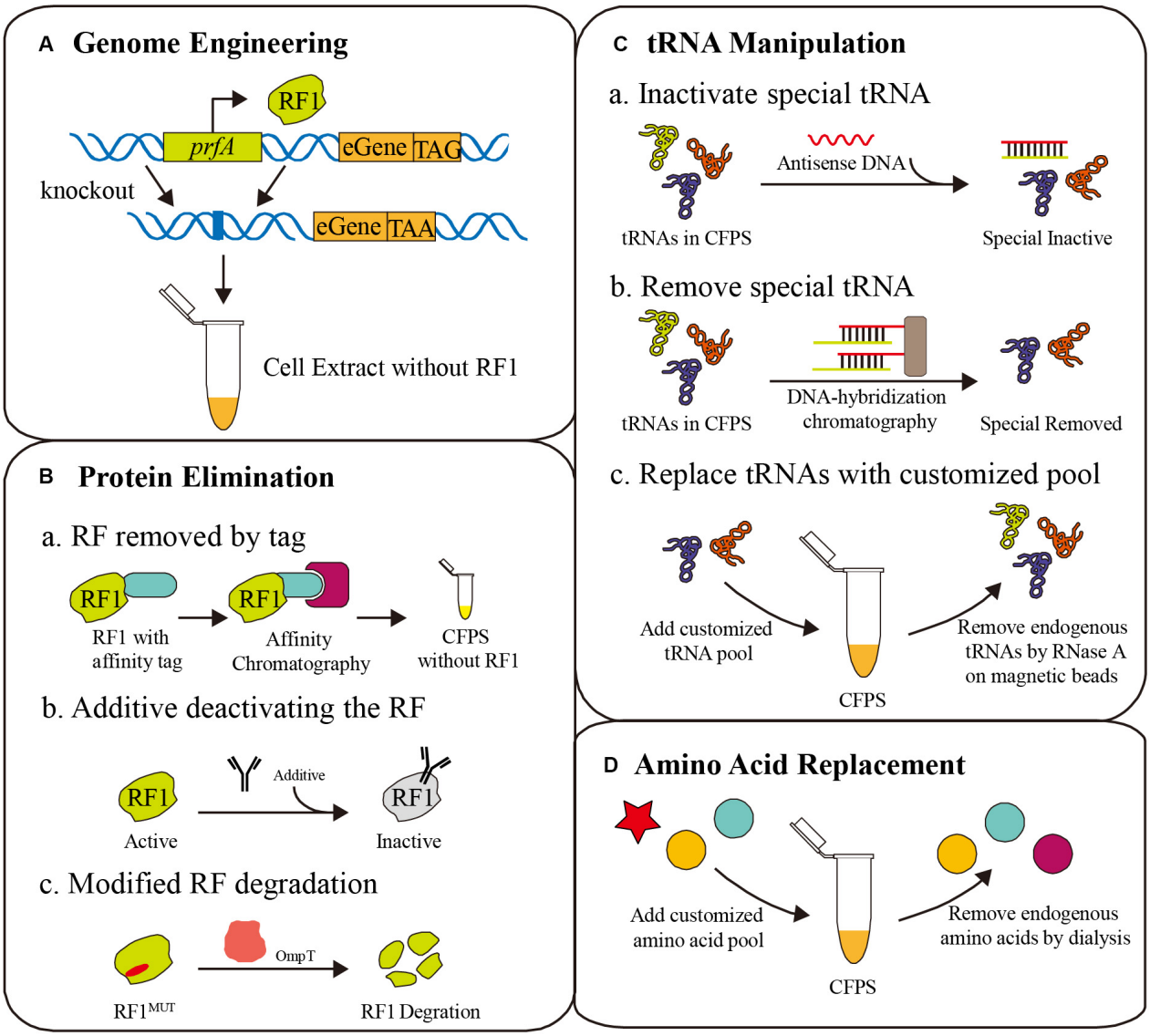


FIGURE 2 | Different strategies for eliminating competition to ncAAs. **(A)** Genome engineering for deleting the prfA gene [corresponding to release factor 1 (RF1)], combined with amber codon replacement in essential genes (eGene). **(B)** Protein elimination strategies for the removal of RF1, including affinity tags, deactivating additives and protease degradation. **(C)** tRNA manipulation to erase specific tRNAs for ncAA reassignment, such as antisense DNA additives, DNA-hybridization chromatography, and tRNA pool replacement. **(D)** Amino acid replacement with a customized amino acid pool containing ncAAs.

for highly efficient incorporations of ncAAs was also described in detailed (Seki et al., 2018). Before research on CFPS based on C321.ΔA, Jewett et al., developed the RF1 deletion strain rEc.E13.ΔA, which is homologous to C321.ΔA (Hong et al., 2014, 2015). In conclusion, by bypassing the time, labor and cost needed for full codon replacement, a partial amber codon replacement strategy is a rapid and simple method for constructing the desired strains, even if the growth rate may be decreased to a certain extent. This method has been applied to diverse *E. coli* strains for various purposes, including MG1655, HST108, and BL21 (DE3).

In addition to the established strains for CFPS, some genome engineering strains also have considerable potential to incorporate ncAAs in CFPS. Wang et al. found that *prfA* could be deleted unconditionally in *E. coli* B strain derivatives. Moreover, after reverting the release factor 2 (RF2) A246T mutation back to Ala, *prfA* could also be knocked out in *E. coli* K12 strain derivatives, demonstrating that RF1 is non-essential in *E. coli* (Johnson et al., 2011, 2012; Wang, 2017). Without any stop codon replacement, it is the most rapid and simple method for RF1-free strain construction.

Although genome engineering provides a completely RF1 elimination strategy without residuals, these above methods also inevitably affect the activity and yield of CFPS and consume time and laborious for genetic manipulation. Consequently, additional strategies were developed for CFPS with unique

advantages, which may be combined with genome engineering for future improvement.

Protein Elimination

Because the open reaction environment can be controlled directly, the CFPS system is more flexible than *in vivo* translation systems. Various competition eliminating strategies are only possible in the CFPS system with its special properties. An important strategy is protein elimination from the CFPS system. By manipulating protein levels, essential proteins can remain active to aid cell growth, but can be removed from the CFPS system to remove competition. For example, RF1 can be selectively removed using affinity tags, deactivating additives, or protease degradation (**Figure 2B**).

Selective removal of RF1 using an affinity tag has already been demonstrated by Otting et al. This approach requires the genetic fusion of a chitin-binding domain sequence to *prfA*, so that the wild-type RF1 is replaced by a variant with three continuous chitin-binding domains at the C-terminus, which maintained the peptide release activity required for cell growth. After cell extract preparation, the lysate was filtered through a chitin column to selectively remove the tagged RF1, resulting in an improved CFPS system with efficient ncAA incorporation at multiple sites (Loscha et al., 2012; Abdelkader et al., 2015; Jabar et al., 2017).

Similarly, the RF1 protein can also be eliminated by modifying the protein to make it sensitive to certain factors and eliminating it from the CFPS system. For example, a temperature-sensitive variant of RF1 had been engineered in *E. coli*, and the cell extract with a thermosensitive RF1 demonstrated a 11-fold increase of efficiency in ncAA incorporation following heat treatment (Short et al., 1999). However, longstanding heat shock may result in the loss of basic activity of the CFPS system, while too little heat can leave residual RF1. Recently, another strategy was developed with no need for heat shock. Yin et al. (2017) successfully inserted an outer membrane protease (OmpT) recognition site into the switch loop of RF1. OmpT protease only identifies extracellular target proteins. Thus, the engineered RF1 can maintain the cell growth rate before cell extract preparation, after which it is cleaved by OmpT during cell lysis, which led to an increase of ncAA incorporation in CFPS.

These protein elimination strategies rely on the modification of proteins by genetic engineering. However, certain additives can also deactivate RF1 in the CFPS system. Agafonov et al. (2005) used an anti-RF1 antibody to block RF1 activity in CFPS and improve the ncAA incorporation efficiency. Nonetheless, the antibody requires expression and purification steps, which increased the costs. Consequently, the easily prepared and adaptable RNA aptamers have also been investigated for RF1 inactivation (Szkaradkiewicz et al., 2002; Shafer et al., 2004). For instance, an RNA aptamer that can orthogonally deactivate RF1 in CFPS was selected from a 50 nt random sequence RNA pool, showing the ability to increase the efficiency of ncAA incorporation (Ogawa et al., 2005; Sando et al., 2007).

RF1 could be removed at either genome or protein level. By comparison, deletion of RF1 gene from genome completely get rid of RF1 bioactivity from cell. And RF1-deficient *E. coli* cells have demonstrated its potential in preparation of RF1-free cell extraction for incorporation of ncAA into protein. However, knockout of gene for RF1 generated side effects, such as slow cell growth. For instance, C321.ΔA has been evolved to improve cell viability by correcting mistakes of genetic manipulation, such as off-target modifications (Wannier et al., 2018). In contrast, RF1 elimination at protein level requires complicated procedures, such as chromatography, degradation or deactivation (Agafonov et al., 2005; Sando et al., 2007; Loscha et al., 2012; Yin et al., 2017). On the other hand, without stop codon replacements by multiple-sites genome editing, protein elimination strategy can be easily applied to other strains. However, due to the open nature of CFPS, more strategies can be flexibly designed.

Manipulation of tRNAs

When a ncAA is inserted into a protein at a sense codon, the endogenous aminoacyl-tRNA will compete with the ncAA aminoacyl-tRNA in the ribosome during translation, resulting in partially native protein. Directly manipulating tRNAs in the cell is unrealistic, but it is convenient in CFPS without a membrane barrier, freeing a large number of sense codons for ncAA incorporation by tRNA manipulation (Tajima et al., 2018). Considering the one-to-one correspondence between tRNAs and codons, artificial codon tables, and even minimum codon combinations can be achieved by tRNA manipulation in CFPS (Calles et al., 2019; Wang et al., 2020). Recently, several tRNA level methods were developed, including tRNA pool replacement, elimination using antisense oligonucleotides, and tRNA-specific enzymatic degradation (**Figure 2C**).

The most direct method is the replacement of all tRNAs with a tRNA pool customized according to research purposes. Due to the molecular weight of ∼25,000 with 70–90 nucleotides in tRNAs, straightforward dialysis is not capable of removing tRNA from cell extracts and some small molecular translation enzymes, such as initiation factor 1, can be lost during dialysis (Bhat et al., 2005). Nevertheless, several methods for the replacement of endogenous tRNAs have been developed. Recently, a near-complete endogenous tRNA eliminating method was reported Bundy et al. (2008) in which RNase A was attached to magnetic beads to remove native tRNAs from cell extract (Salehi et al., 2017). Additionally, semisynthetic tRNA complements, which contain the majority of 48 synthetic tRNAs, were shown to be functional in CFPS, extending the ability to customize additive tRNA pools for ncAA incorporation at multiple sense codons (Cui et al., 2015).

Considering the time-consuming steps required for native tRNA removal and custom tRNA synthesis in the tRNA pool replacement strategy, tRNA antisense oligonucleotides have emerged as a powerful tool for tRNA-specific elimination and inactivation in CFPS. Cui et al. (2017) removed the $tRNA^{Arg}_{CCU}$ by DNA-hybridization chromatography of tRNA-specific antisense oligonucleotides, displaying a universal method for sense codon reassignment. However, without additional chromatography steps, methylated anti-tRNA oligonucleotides were demonstrated to deactivate multiple selected tRNAs by tRNA–DNA hybridization, which was induced due to heating and annealing with the tRNA mix of CFPS. This approach is more efficient than chromatography for tRNA inactivation (Cui et al., 2018).

In addition to tRNA-specific antisense oligonucleotides, the $tRNA^{Arg}$-specific tRNase, colicin D, was also utilized for specific tRNA elimination. Colicin D specifically cleaves $tRNA^{Arg}$ at the anticodon loop, including $tRNA^{Arg}_{ACG}$, $tRNA^{Arg}_{GCG}$, $tRNA^{Arg}_{CCG}$, $tRNA^{Arg}_{UCG}$, $tRNA^{Arg}_{UCU}$, and $tRNA^{Arg}_{CCU}$, which can free six codons for reassignment (Tomita et al., 2000). Kang et al. utilized resin-bound colicin D to cleave all $tRNA^{Arg}$ in the cell-free system and supplemented synthetic $tRNA^{Arg}_{CCU}$ for arginine translation at a single codon, successfully demonstrating the recoding of four sense codons to ncAAs (Lee K.B. et al., 2016).

These tRNA manipulation strategies are able to free a large number of sense codons for efficient ncAA incorporation which are unattainable via stop codon suppression. Nonetheless, limitations still exist. For tRNA pool replacement, only 48 synthetic tRNAs were shown functional in CFPS, and the complete synthetic tRNA substitution pool was not achieved, which may hamper some specific artificial codon tables (Cui et al., 2015). For specific tRNA elimination, the $tRNA^{Arg}$-specific tRNase (colicin D) method was limited by tRNase deficiency and could not be adopted more widely, while the tRNA antisense oligonucleotides may be a potential powerful tool for researchers.

Amino Acid Replacement

Due to the relaxed amino acid specificity of aaRS, a few ncAAs that are structurally similar to canonical amino acids can be charged to native tRNAs by corresponding aaRS. Examples include selenomethionine (Kigawa et al., 2002), fluorinated tryptophan analogs (Mat et al., 2010), chlorinated tyrosine analogs (Singh-Blom et al., 2014), 4-^{18}F-fluoro-L-proline and 3,4-dihydroxy-l-phenylalanine (Harada et al., 2016). Thus, structurally similar ncAAs can be incorporated in place of the corresponding native amino acids. *In vivo*, a few nutrient deficient strains lacking the specific amino acid biosynthesis enzymes have been constructed and cultured with supplementation of ncAAs for artificial protein expression (Budisa et al., 2004; Montclare et al., 2009). However, ncAAs are toxic to the host strain because they will also be incorporated into essential endogenous proteins, resulting in a low growth rate and low protein yield. Moreover, residues of the corresponding canonical amino acid added during the preculture will also compete with the ncAA, leading to partially native protein.

Recently, CFPS was suggested to overcome these defects in amino acid analog incorporation. Even a detailed video protocol has been posted online (Worst et al., 2016). A cell extract of a nutrient deficient strain lacking the specific amino acid biosynthesis enzymes was prepared, followed by multiple diafiltration processes to remove all amino acids. The specific ncAA and 19 other canonical amino acids were added to the CFPS reaction for uncompetitive ncAA incorporation (Singh-Blom et al., 2014; **Figure 2D**). For instance, L-DOPA, a tyrosine substitute, was incorporated into protein with almost eight times higher efficiency than competitive systems (Lee K.-H. et al., 2016). The highly toxic arginine analog canavanine was also firstly incorporated into proteins in a tyrosine-replaced CFPS system (Worst et al., 2015).

This amino acid replacement CFPS method has been a cost-effective and simple alternative to ncAA incorporation methodologies without requiring engineered aaRS or tRNAs. However, owing to the loss of replaced canonical amino acids, various potential applications, which need the 20 canonical amino acids, are limited in this method. What's more, although the specific amino acid biosynthesis enzymes had been knock out in cell-free strains, which avoided other amino acids generating the omitted amino acid to make incorrect incorporations. The omitted amino acid may still be generated from metabolic process in cell extract, such as degradation of endogenous proteins. This problem may be solved by some protease-deficient strains in future studies.

Incorporation of Multiple Different ncAAs

In addition to the incorporation of single ncAAs, the incorporation of multiple different ncAAs into the same protein has been achieved *in vivo* and *in vitro*, widely expanding the ability to engineer artificial proteins. By repeating ncAA reassigned codons in template, it is accessible to incorporate a single type ncAA into multiple different locations without methodological modification (Johnson et al., 2011; Martin et al., 2018). On the countrary, efficiently incorporating multiple different ncAAs into a single protein requires method improvements and modifications. So this section will focus on the emerging methods of multiple different ncAAs incorporation into a single protein. However, *in vivo* methods mainly relied on stop codon repression of the cell protein expression machine, leading to an undistinguished product yield due to the competition of peptide release factors (Wan et al., 2010; Chatterjee et al., 2013a; Wang et al., 2014; Venkat et al., 2018). Only one release factor has been deleted, because the other peptide release factor must be retained for translational termination in cells. Consequently, the RF1-deficient strains, such as C321.ΔA, are currently the most appropriate host strain for reducing competition during the incorporation of multiple ncAAs (Zheng et al., 2018). Although up to three different ncAAs have been incorporated into the same protein in the C321.ΔA strain, the competition between ncAAs and peptide RF2 still had negative effects in cells (Italia et al., 2019).

Recently, cell extracts of RF1-deficient strains were also utilized in a CFPS system for efficient and modular artificial protein expression with multiple different ncAAs, but partial competition was still present, similar to the cells (Hong et al., 2014; Ozer et al., 2017; Chemla et al., 2019). Further, several total competition-removing CFPS methods were developed using different strategies. For instance, Cui et al. (2017) combined tRNA-specific affinity chromatography with stop codon repression in RF1-deficient strains for dual protein fluorescent labeling by incorporating AzF and BPFL-Cys into a single protein. This opens up the possibility of combining approaches at different levels to engineer proteins more freely.

TRANSLATION FACTORS ENGINEERING

Various factors play important role in protein translation as well, especially in ncAA incorporation. These translation factors are

also engineered to facilitate efficient ncAA incorporation. In this part, engineering achievement of elongation factor Tu (EF-Tu), ribosome and aaRS/tRNA pairs are summarized and discussed.

EF-Tu

EF-Tu is crucial in translation elongation phase, which is abundant in *E. coli* (Furano, 1975). EF-Tu binds aminoacyl-tRNA and delivery them to the ribosome for peptide elongation. However, EF-Tu may not efficiently bind to ncAA with macromolecular or negatively charged groups, leading to ineffective incorporation (Doi et al., 2007; Haruna et al., 2014). Thus far, many efforts had been made to improve EF-Tu in ncAA incorporation by engineering the amino acid binding domain.

With evolution of the binding domain in EF-Tu, ncAAs such as phosphoserine, phosphotyrosine, selenocysteine, pyrenylalanine, and *p*-azido-phenylalanine, were efficiently incorporated into peptide. Some ncAAs carrying negative charged phosphorylation side chain are hard to be recognized by wild type EF-Tu. Therefore, mutant EF-Tu (H67R, E216N, D217G, F219Y, T229S, and N274W) was developed to efficiently incorporated phosphoserine into peptide (Park et al., 2011) and milligram of protein with phosphoserine incorporation was obtained in cell-free system extracted from strain C321.ΔA (Oza et al., 2015). Similarly, the EF-Tu variant (E216V, D217G, and F219G) was also successfully selected for phosphotyrosine incorporation (Fan et al., 2016). And selenocysteine, another negatively charged amino acid, was also efficiently recognized by EF-Tu variant (H67Y, Q98Q, E216D, D217R, and N274R) (Haruna et al., 2014). Doi et al. (2007) developed EF-Tu mutants (E215A and D216A), in which the binding pocket of aminoacyl-tRNA was enlarged, to improve the binding of ncAAs carrying bulky side chain and this system successfully worked with ncAAs of L-1-pyrenylalanine, L-2-pyrenylalanine, and DL-2-anthraquinonylalanine as well. Incorporation of *p*-azido-phenylalanine was assisted by EF-Tu with mutant binding domain (S65A, D216A, and V274A) (Gan et al., 2017). Although the function of these EF-Tu variants has been proved *in vivo*, it is easy to applied them in CFPS by either direct addition of purified EF-Tu variants protein or co-expression with other endogenous wild-type translation factors in cell. For instance, protein with phosphoserine incorporation was synthesized to milligram in CFPS with engineered EF-Tu (Oza et al., 2015).

Ribosome

Besides the normal proteinogenic amino acids, wild-type ribosome is also able to accept some analogs of L-α-amino acids (Dedkova and Hecht, 2019). Nonetheless, ncAAs with modified chemical backbone such as partial D-amino acids were incompatible with the wild-type translation machinery (Melnikov et al., 2019). Modification of the peptidyl transferase center on ribosomal 50S subunit has high potential to enhance special ncAAs incorporation. For instance, the mutant 23s rRNA mutations (region of 2447–2451, 2457–2462, 2057–2063, and 2502–2507) allowed efficient peptidyl transfer of D-amino acids or β-amino acids, such as D-methionine, D-phenylalanine (Dedkova et al., 2003, 2006), β-alanine and β-phenylalanine (Maini et al., 2013; Melo Czekster et al., 2016). Interestingly, mutant 16S rRNA (A238U, G849U, G1175U, G1516U) was identified as well with improved activity for selenocysteine incorporation (Thyer et al., 2013). Moreover, several novel features have been evolved in ribosome to improve ncAA incorporation. Wang et al. (2007) developed an orthogonal ribosome with a decreased functional interaction with RF1 and improved amber codon suppression. Another orthogonal ribosome was developed for efficient quadruplet codon suppression (Neumann et al., 2010). Engineered ribosome may bring the risk of losing cell viability especially under overexpression (Dedkova et al., 2003, 2006). Therefore, specially designed orthogonal ribosome may be the good solution, by which ncAA incorporation could be separated from endogenous translation and minimize the influence on cell viability.

Evolution of aaRS/tRNA Pairs

Aminoacylation is the key procedure for orthogonality of tRNA and corresponding amino acids. Thus far, a few of specific aaRS/tRNA pairs were developed, such as the PylRS/tRNA pair for lysine, phenylalanine and pyrrolysine analogs from *Methanosarcina* spp. (Polycarpo et al., 2006; Katayama et al., 2012; Wang Y.S. et al., 2012), *Methanococcus jannaschii* and *acetivorans* TyrRS/tRNA pairs for tyrosine analogs (Ikeda-Boku et al., 2013), *Methanococcus maripaludis* SepRS/tRNA pair for phosphoserine (Rogerson et al., 2015), *Saccharomyces cerevisiae* PheRS/tRNA pair for phenylalanine and alanine analogs (Kwon and Lim, 2015), *Saccharomyces cerevisiae* TrpRS/tRNA pair for tryptophan and alanine analogs (Chatterjee et al., 2013b), *Desulfitobacterium hafniense* PylRS/tRNA for lysine analogs and *Pyrococcus horikoshii* ProRS/tRNA for proline analogs (Chatterjee et al., 2012).

However, the slow incorporation rate of nnAAs charged anminoacyl-tRNA leads to increased misincorporation of the endogenous aminoacyl-tRNA, inserting undesired canonical amino acid (O'Donoghue et al., 2012). Additionally, the site-specific incorporation requires accurate aminoacylation reaction between ncAA and its partner tRNA. Thus, directed evolution was applied to evolve efficient and accurate aaRS/tRNA pairs, including phage-assisted continuous evolution (PACE) (Esvelt et al., 2011), compartmentalized partnered replication (CPR) (Ellefson et al., 2014), parallel positive selection combined with deep sequencing and tRNA Extension (tREX) assisted with computationally identification (Zhang et al., 2017; Cervettini et al., 2020). And improved efficiency and specificity have been demonstrated for aaRS/tRNA pairs of *N*ε-Boc-L-lysine (BocK), *N*ε-acetyl- L-lysine (AcK) (Bryson et al., 2017), *p*-azido-L-phenylalanine (pAzF) (Amiram et al., 2015), 3-iodo-L-tyrosine (Oki et al., 2008). O-methyl-L-tyrosine (Cervettini et al., 2020), and so on. However, the degree of the orthogonality of these evolved aaRS/tRNA pairs is crucial for *in vivo* application. All of these evolved aaRS/tRNA pairs could also be applied in CFPS by either directly addition of purified protein or pre-expression in the cell. And, the activity of orthogonal aaRS/tRNA could be improved by accurately adjusting its relative concentration in cell-free system.

NOVEL ncAAs AND *IN VITRO* AMINOACYLATION METHODS FOR CFPS

Without limitations of cell viability and membrane barrier, CFPS allows the direct addition of specific aminoacyl-tRNA to participate in ribosomal translation, in which the aminoacyl-tRNA is pre-synthesized before the CFPS reaction. By decoupling aminoacylation and translation in CFPS, several limitations can be overcome. On the one hand, because the ncAA aminoacylation reaction is physically separated from endogenous reactions, the orthogonality of aaRS is unnecessary, so that time-consuming enzyme evolution can be avoided and more ncAAs can be utilized by *in vitro* aminoacylation (Urbanek et al., 2018). On the other hand, more chemical and enzymatic methods can be utilized to generate aminoacyl-tRNAs, breaking the limitation that only analogs of L-α-amino acids can be charged by aaRS (Ninomiya et al., 2003; Murakami et al., 2006). Through *in vitro* aminoacylation methods, orthogonality requirement of aminoacylation process can be avoided, in which tRNA can only be charged by the only amino acid *in vitro*. But after translation, the released tRNA may be recharged by endogenous aaRS if it is not orthogonal to endogenous system, possibly resulting in incorrect incorporation. Thus, orthogonality evolution of aminoacylation factors can be skipped to save cost, while orthogonal tRNA is needed to achieve accurate ncAAs incorporation. Recently, several novel ncAAs were successfully incorporated into protein by such *in vitro* aminoacylation methods, including D-α-amino acids (Katoh et al., 2017b), β-amino acids (Katoh and Suga, 2018), γ-amino acids (Ohshiro et al., 2011), N-alkylated-α-amino acids (Kawakami et al., 2013), N-acylated-α-amino acids (Kwiatkowski et al., 2014), α-hydroxy acids (Ohta et al., 2008), α-benzoic acids and even foldamers (Kawakami et al., 2016; Rogers et al., 2018; **Figure 3**).

These novel ncAAs widely expand the ability to engineer artificial proteins with novel features, which may increase the proteolytic stability, membrane permeability, and conformational rigidity of the peptide. For instance, the thioether macrocyclic peptides incorporated with D-amino acids could provide high serum stability due to structural difference and protease resistance (Bashiruddin and Suga, 2015). And β-amino acids were considered as the potential building blocks of peptide analogs for pharmaceutical uses with the improved stability against proteolysis (Maini et al., 2013). Cyclic peptide antagonists containing N-alkylated-α-amino acids also showed more membrane-permeable and stable to proteolysis (Chatterjee et al., 2008; Baeriswyl and Heinis, 2013). The foldamers may potentially be highly water-soluble and possess cell-penetrating properties (Gillies et al., 2007). However, more applications are limited by preparation method of peptide with novel ncAAs. In this part, novel amino acids and different *in vitro* aminoacylation methods for CFPS platforms will be summarized and discussed.

In vitro Aminoacylation by aaRS

Utilizing the relaxed amino acid specificity of some aaRS, a few analogs of L-α-amino acids can be directly charged to tRNAs by aaRS (Katoh and Suga, 2019; **Figure 4A**). When aminoacylation and translation are conducted in the same system, in order to avoid competition between amino acids in the same aaRS/tRNA pair, the corresponding canonical amino acid should be removed to safeguard the orthogonality of the translation machine. Via pre-aminoacylation with ncAAs by aaRS, not only any codon, including stop codons, can be reassigned to a ncAA with a defined aminoacyl-tRNA, but also more efficient ncAA incorporation can be realized by optimizing additive concentrations (Urbanek et al., 2020). For instance, the general method for the expression of isotopically labeled proteins relies on aaRS charging isotopically labeled amino acids. With [^{15}N,^{13}C]-glutamine pre-charged by yeast glutaminyl-tRNA synthetase, homogeneous protein was obtained for high-dimensional NMR spectroscopy (Urbanek et al., 2018).

Another method is the chemical modification of side chains in canonical aminoacyl-tRNAs after normal aminoacylation by aaRS, generating ncAA aminoacyl-tRNAs to participate in translation (**Figure 4A**). Several side chains were successfully modified in canonical aminoacyl-tRNAs, generating moieties such as phenyllactyl amino acids (Fahnestock and Rich, 1971), N-methyl amino acids (Merryman and Green, 2004), glycosyl amino acids (Fahmi et al., 2007), fluorescence labeled amino acids and BODIPY labeled amino acids (Iijima and Hohsaka, 2009).

Compared to in vivo methods relying on introducing orthogonal aaRS/tRNA pairs into cells, in vitro aminoacylation by aaRS physically separates the ncAA aminoacylation reaction from endogenous reactions, avoiding orthogonality requirement of aaRS in ncAA aminoacylation process. The cost and time-consuming aaRS evolution can be skipped. And available ncAAs are expanded without orthogonality. However, after the dehydration condensation reaction in ribosome, the released tRNA could be recharged with unwished amino acids if it is not

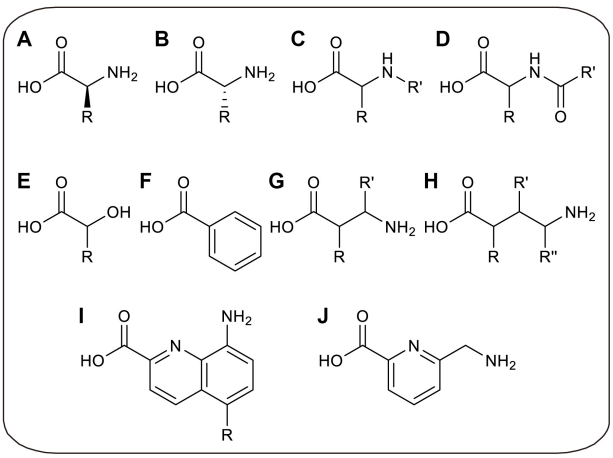


FIGURE 3 | Recently developed novel ncAAs only utilized in CFPS. **(A)** L-α-amino acids with non-canonical side chains, which can only be charged to tRNAs *in vitro*. **(B)** D-α-amino acids. **(C)** N-alkylated-α-amino acids. **(D)** N-acylated-α-amino acids. **(E)** α-hydroxy acids. **(F)** α-benzoic acid. **(G)** β-amino acids. **(H)** γ-amino acids. **(I)** Quinoline-based foldamers. **(J)** Pyridine-based foldamer.

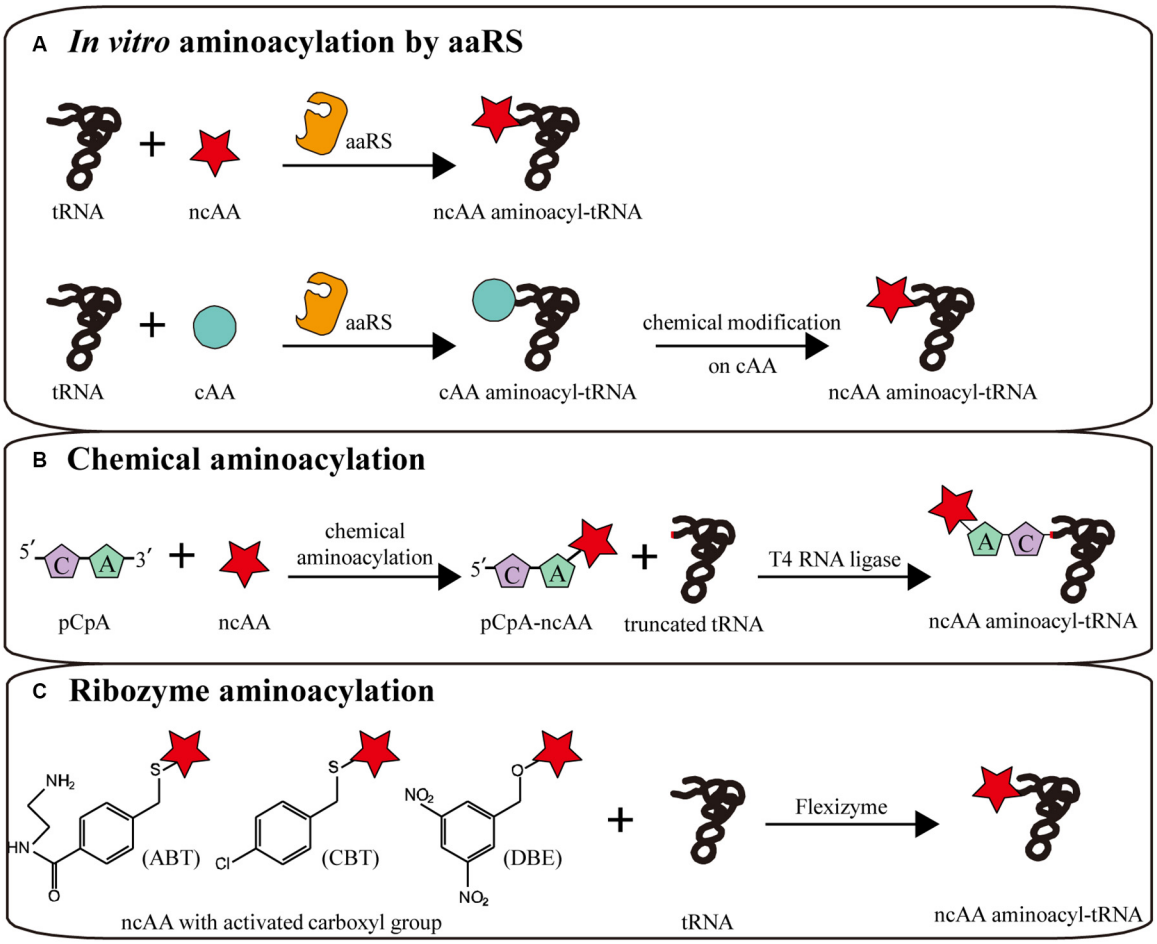


FIGURE 4 | *In vitro* aminoacylation methods combined with CFPS. **(A)** *In vitro* aminoacylation by aaRS. Without an orthogonality barrier, extracellular aaRS can load tRNAs with more kinds of ncAAs. Canonical amino acid (cAA) aminoacyl-tRNAs can be chemically modified to become non-canonical. **(B)** Chemical aminoacylation. The truncated tRNA can be ligated with chemically aminoacylated pCpA-ncAA by T4 RNA ligase, generating a complete ncAA aminoacyl-tRNA. **(C)** Ribozyme aminoacylation. Flexizymes can synthesize aminoacyl-tRNAs from ncAAs with different activated carboxyl groups.

orthogonal to endogenous aaRS, leading to misincorporations. For precise and efficient ncAA incorporation, the corresponding tRNA needs to be orthogonal to endogenous systems. What's more, the available specter of ncAAs is still largely limited to analogs of L-α-amino acids.

Chemical Aminoacylation

Chemical aminoacylation is a versatile classical strategy for aminoacyl-tRNA synthesis *in vitro* (Hecht et al., 1978; **Figure 4B**). Although several improvements have been made, the general process remained consistent (Heckler et al., 1984; Noren et al., 1989; Ninomiya et al., 2003; Lodder et al., 2005; Kwiatkowski et al., 2014). The amino acid with appropriate protecting groups is chemically charged to the 3′-OH of pCpA or pdCpA, which was chemical synthesized. Subsequently, the corresponding truncated tRNA, lacking two bases at the 3′-terminal, is ligated with aminoacyl-pCpA (or pdCpA) by T4 RNA ligase to generate the complete aminoacyl-tRNA. In theory, this method can be used to load any tRNA with any desired amino acid, such as N-acetyl amino acids (Yamanaka et al., 2004), glycosylated serine, and tyrosine (Fahmi et al., 2007), N-methyl and N-nitro arginine (Choudhury et al., 2007), citrulline (Breuillard et al., 2015), homoarginine, and isotopically labeled canonical amino acids (Peuker et al., 2016).

In theory, the greatest strength of this method is that any desired amino acid can be charged to any tRNA. However, the combination of chemical synthesis with enzymatic ligation is labor-intensive and it is challenging to produce homogeneous aminoacyl-tRNAs. Especially the primary by-product, self-ligated tRNA, is unavoidably generated by T4 RNA ligase, leading to inefficient synthesis, which may be avoided in future researches.

Ribozyme Aminoacylation

In addition to aaRS and chemical aminoacylation, Suga et al. evolved small ribozymes (44–46 nt) that synthesize aminoacyl-tRNAs. These so-called flexizymes can charge tRNAs with amino acids with an activated carboxyl group (Murakami et al., 2006; Goto et al., 2011). This activated carboxyl group and the

3′-terminal NCCA sequence of tRNA are crucial recognized sites of the flexizyme catalytic domain (Xiao et al., 2008). Therefore, in theory, the ncAA can be reassigned to any codon. Recently, some flexizyme variants were artificially evolved to accept amino acids with different activated carboxyl groups (Morimoto et al., 2011). The aFx, dFx, and eFx flexizymes are three widely used representatives, respectively recognizing amino acids carrying an activated amino-derivatized benzyl thioester (ABT), dinitrobenzyl ester (DBE) or chlorobenzyl ester (CBE) (Niwa et al., 2009; Passioura and Suga, 2013; **Figure 4C**). More recently, more than thirty activated groups were verified to be accessible to flexizyme aminoacylation, significantly expanding the range of available amino acids (Lee et al., 2019).

As an emerging method, flexizyme aminoacylation has been applied as a versatile strategy for the incorporation of a great many ncAAs. The only two limitations imposed on the amino acids are the feasibility to activate the carboxyl with a certain group and chemical stability during the aminoacylation process. In fact, different ncAAs have been utilized even with multiple incorporations, including D-α-amino acids (Katoh et al., 2017b), β-amino acids (Katoh and Suga, 2018), γ-amino acids (Ohshiro et al., 2011), α-hydroxy acids (Ohta et al., 2008), α-benzoic acids (Kawakami et al., 2016), N-alkyl-L-α-amino acids (Kawakami et al., 2013), N-methyl-L-α-amino acids (Kwiatkowski et al., 2014), and even foldamers (Rogers et al., 2018). However, the efficiency of flexizyme aminoacylation varies depending on the ncAAs with a wide range of 17–91%, which can only be optimized empirically, since there are no established design rules (Goto et al., 2011). Moreover, the ribosome incorporation capacity of ncAA aminoacyl-tRNAs also ranges from zero to high efficiency. For instance, only eight of nineteen D-α-amino acids showed highly efficient incorporation, while low efficiencies were observed in four D-α-amino acids and the other was incompatible with the ribosome (Fujino et al., 2013, 2016). Henceforth, more versatile aminoacylation and translation techniques are worth developing to satisfy the increasing demand for ncAA incorporation.

CONCLUSION AND PROSPECTS

Cell-free protein synthesis has emerged as a powerful platform for engineering the genetic code with its unique advantages allowing the efficient and extensive incorporation of ncAAs, which were summarized in this review. To improve the efficiency and yield of artificial proteins, several strategies have been presented to remove ncAA competitors at different levels, including genome engineering, elimination of peptide release factors, tRNA manipulation and amino acid replacement. The pros and cons of these strategies have been discussed in detail above. By combining methods at different levels, multiple codons, including sense and nonsense codons, could be simultaneously reassigned to different ncAAs without endogenous competitors, contributing to more flexible artificial protein synthesis (Cui et al., 2017). To enhance ncAA incorporation, several translation factors have been engineered for improvement, such as EF-Tu, ribosome and aaRS/tRNA pairs. At the same time, CFPS is compatible with *in vitro* aminoacylation methods, which can supply a wider range of novel and unique ncAAs for artificial protein research. Compared with aaRS and chemical aminoacylation methods, flexizymes have been developed into a highly versatile and efficient approach for aminoacyl-tRNA synthesis, with many recent studies confirming its expansion. By emphasizing these crucial techniques for ncAA incorporation via CFPS, this review is intended to provide a bridge between current technologies and future directions in this rapidly developing field.

Looking forward, these tools have the potential to be applied in various emerging fields, such as antibody-drug conjugates, protein labeling, peptide-based materials and so on (Albayrak and Swartz, 2014; Stech et al., 2017; Suga, 2018). Especially in enzyme activity improvement, several efforts have been made with significant influence. For instance, enzyme phosphorylation is a key modification for protein activity, which may occur on tyrosines, threonines and serines. Phosphotyrosine and phosphothreonine were used to activate vasodilator-stimulated phosphoprotein (VASP) and the nicotinic acetyl choline receptor (nAChR) (Rothman et al., 2005). With phosphoserine incorporation, several homogeneous phosphorylated enzymes have been expressed, such as human MEK1 kinase, human Stimulator of Interferon Genes (STING), and Bcl2-associated agonist of cell death (BAD) (Oza et al., 2015; Zhu et al., 2019). Furthermore, methylhistidine also showed important activating effects in the enzyme catalytic center of fluorescent proteins, heme enzymes, and metalloenzymes (Xiao et al., 2014; Gan et al., 2018; Ortmayer et al., 2020). Acetyllysine and methyllysine play a crucial role in the control of chromatin and epigenetic programs by histones, which has been demonstrated by site-specific incorporation with *in vitro* aminoacylation in CFPS (Sterner and Berger, 2000; Yanagisawa et al., 2014; Wakamori et al., 2015; Xiong et al., 2016). In addition, numerous unclear modifications of amino acids in the enzymatic active center need to be investigated in the future, such as hydroxylation, glycosylation, sulfation, and amidation. Overall, this review hopes to assist researchers in choosing suitable methods to incorporate ncAAs efficiently and easily using CFPS, hopefully inspiring breakthroughs of ncAA techniques and applications in the future.

AUTHOR CONTRIBUTIONS

YW gathered the information, drafted the original manuscript, and designed the figures. ZW contributed to literature collection. HQ conceived the presented idea and supervised the work. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key R&D Program of China (Grant No. 2019YFA0904103) and the National Science Foundation of China (Grant Nos. 21476167, 21778039, and 21621004).

REFERENCES

Abdelkader, E. H., Feintuch, A., Yao, X., Adams, L. A., Aurelio, L., Graham, B., et al. (2015). Protein conformation by EPR spectroscopy using gadolinium tags clicked to genetically encoded p-azido-L-phenylalanine. *Chem. Commun.* 51, 15898–15901. doi: 10.1039/c5cc07121f

Adachi, J., Katsura, K., Seki, E., Takemoto, C., Shirouzu, M., Terada, T., et al. (2019). Cell-free protein synthesis using S30 extracts from *Escherichia coli* RFzero strains for efficient incorporation of non-natural amino acids into proteins. *Int. J. Mol. Sci.* 20:492. doi: 10.3390/ijms20030492

Agafonov, D. E., Huang, Y., Grote, M., and Sprinzl, M. (2005). Efficient suppression of the amber codon in *E. coli* in vitro translation system. *FEBS Lett.* 579, 2156–2160. doi: 10.1016/j.febslet.2005.03.004

Albayrak, C., and Swartz, J. R. (2014). Direct polymerization of proteins. *ACS Synth. Biol.* 3, 353–362. doi: 10.1021/sb400116x

Amiram, M., Haimovich, A. D., Fan, C., Wang, Y. S., Aerni, H. R., Ntai, I., et al. (2015). Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat. Biotechnol.* 33, 1272–1279. doi: 10.1038/nbt.3372

Baeriswyl, V., and Heinis, C. (2013). Phage selection of cyclic peptide antagonists with increased stability toward intestinal proteases. *Protein Eng. Des. Sel.* 26, 81–89. doi: 10.1093/protein/gzs085

Bashiruddin, N. K., and Suga, H. (2015). Construction and screening of vast libraries of natural product-like macrocyclic peptides using in vitro display technologies. *Curr. Opin. Chem. Biol.* 24, 131–138. doi: 10.1016/j.cbpa.2014.11.011

Bhat, S. V., Nagasampagi, B. A., and Sivakumar, M. (2005). *Chemistry of Natural Products*. Berlin: Springer Science & Business Media.

Breuillard, C., Bonhomme, S., Couderc, R., Cynober, L., and De Bandt, J. P. (2015). In vitro anti-inflammatory effects of citrulline on peritoneal macrophages in Zucker diabetic fatty rats. *Br. J. Nutr.* 113, 120–124. doi: 10.1017/S0007114514002086

Bryson, D. I., Fan, C., Guo, L. T., Miller, C., Soll, D., and Liu, D. R. (2017). Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat. Chem. Biol.* 13, 1253–1260. doi: 10.1038/nchembio.2474

Budisa, N., Pal, P. P., Alefelder, S., Birle, P., Krywcun, T., Rubini, M., et al. (2004). Probing the role of tryptophans in aequorea victoria green fluorescent proteins with an expanded genetic code. *Biol. Chem.* 385, 191–202. doi: 10.1515/BC.2004.038

Bundy, B. C., Franciszkowicz, M. J., and Swartz, J. R. (2008). *Escherichia coli*-based cell-free synthesis of virus-like particles. *Biotechnol. Bioeng* 100, 28–37. doi: 10.1002/bit.21716

Calles, J., Justice, I., Brinkley, D., Garcia, A., and Endy, D. (2019). Fail-safe genetic codes designed to intrinsically contain engineered organisms. *Nucleic Acids Res.* 47, 10439–10451. doi: 10.1093/nar/gkz745

Cervettini, D., Tang, S., Fried, S. D., Willis, J. C. W., Funke, L. F. H., Colwell, L. J., et al. (2020). Rapid discovery and evolution of orthogonal aminoacyl-tRNA synthetase-tRNA pairs. *Nat. Biotechnol.* doi: 10.1038/s41587-020-0479-2

Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H., and Schultz, P. G. (2013a). A versatile platform for single- and multiple-unnatural amino acid mutagenesis in *Escherichia coli*. *Biochemistry* 52, 1828–1837. doi: 10.1021/bi4000244

Chatterjee, A., Xiao, H., Yang, P. Y., Soundararajan, G., and Schultz, P. G. (2013b). A tryptophanyl-tRNA synthetase/tRNA pair for unnatural amino acid mutagenesis in *E. coli*. *Angew. Chem. Int. Ed. Engl.* 52, 5106–5109. doi: 10.1002/anie.201301094

Chatterjee, A., Xiao, H., and Schultz, P. G. (2012). Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14841–14846. doi: 10.1073/pnas.1212454109

Chatterjee, J., Gilon, C., Hoffman, A., and Kessler, H. (2008). N-methylation of peptides: a new perspective in medicinal chemistry. *Acc. Chem. Res.* 41, 1331–1342. doi: 10.1021/ar8000603

Chemla, Y., Ozer, E., Shaferman, M., Zaad, B., Dandela, R., and Alfonta, L. (2019). Simplified methodology for a modular and genetically expanded protein synthesis in cell-free systems. *Synth. Syst. Biotechnol.* 4, 189–196. doi: 10.1016/j.synbio.2019.10.002

Chin, J. W. (2017). Expanding and reprogramming the genetic code. *Nature* 550, 53–60. doi: 10.1038/nature24031

Choudhury, A. K., Golovine, S. Y., Dedkova, L. M., and Hecht, S. M. (2007). Synthesis of proteins containing modified arginine residues. *Biochemistry* 46, 4066–4076. doi: 10.1021/bi062042r

Cui, Z., Mureev, S., Polinkovsky, M. E., Tnimov, Z., Guo, Z., Durek, T., et al. (2017). Combining sense and nonsense codon reassignment for site-selective protein modification with unnatural amino acids. *ACS Synth. Biol.* 6, 535–544. doi: 10.1021/acssynbio.6b00245

Cui, Z., Stein, V., Tnimov, Z., Mureev, S., and Alexandrov, K. (2015). Semisynthetic tRNA complement mediates in vitro protein synthesis. *J. Am. Chem. Soc.* 137, 4404–4413. doi: 10.1021/ja5131963

Cui, Z., Wu, Y., Mureev, S., and Alexandrov, K. (2018). Oligonucleotide-mediated tRNA sequestration enables one-pot sense codon reassignment in vitro. *Nucleic Acids Res.* 46, 6387–6400. doi: 10.1093/nar/gky365

Dedkova, L. M., Fahmi, N. E., Golovine, S. Y., and Hecht, S. M. (2003). Enhanced D-amino acid incorporation into protein by modified ribosomes. *J. Am. Chem. Soc.* 125, 6616–6617. doi: 10.1021/ja035141q

Dedkova, L. M., Fahmi, N. E., Golovine, S. Y., and Hecht, S. M. (2006). Construction of modified ribosomes for incorporation of D-amino acids into proteins. *Biochemistry* 45, 15541–15551. doi: 10.1021/bi060986a

Dedkova, L. M., and Hecht, S. M. (2019). Expanding the scope of protein synthesis using modified ribosomes. *J. Am. Chem. Soc.* 141, 6430–6447. doi: 10.1021/jacs.9b02109

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.e9–1754.e9. doi: 10.1016/j.chembiol.2019.10.008

Doi, Y., Ohtsuki, T., Shimizu, Y., Ueda, T., and Sisido, M. (2007). Elongation factor Tu mutants expand amino acid tolerance of protein biosynthesis system. *J. Am. Chem. Soc.* 129, 14458–14462. doi: 10.1021/ja075557u

Dopp, B. J. L., Tamiev, D. D., and Reuel, N. F. (2019). Cell-free supplement mixtures: elucidating the history and biochemical utility of additives used to support in vitro protein synthesis in *E. coli* extract. *Biotechnol. Adv.* 37, 246–258. doi: 10.1016/j.biotechadv.2018.12.006

Dumas, A., Lercher, L., Spicer, C. D., and Davis, B. G. (2015). Designing logical codon reassignment - Expanding the chemistry in biology. *Chem. Sci.* 6, 50–69. doi: 10.1039/c4sc01534g

Ellefson, J. W., Meyer, A. J., Hughes, R. A., Cannon, J. R., Brodbelt, J. S., and Ellington, A. D. (2014). Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat. Biotechnol.* 32, 97–101. doi: 10.1038/nbt.2714

Esvelt, K. M., Carlson, J. C., and Liu, D. R. (2011). A system for the continuous directed evolution of biomolecules. *Nature* 472, 499–503. doi: 10.1038/nature09929

Fahmi, N. E., Dedkova, L., Wang, B., Golovine, S., and Hecht, S. M. (2007). Site-specific incorporation of glycosylated serine and tyrosine derivatives into proteins. *J. Am. Chem. Soc.* 129, 3586–3597. doi: 10.1021/ja067466n

Fahnestock, S., and Rich, A. (1971). Ribosome-catalyzed polyester formation. *Science* 173, 340–343. doi: 10.1126/science.173.3994.340

Fan, C., Ip, K., and Soll, D. (2016). Expanding the genetic code of *Escherichia coli* with phosphotyrosine. *FEBS Lett.* 590, 3040–3047. doi: 10.1002/1873-3468.12333

Fujino, T., Goto, Y., Suga, H., and Murakami, H. (2013). Reevaluation of the D-amino acid compatibility with the elongation event in translation. *J. Am. Chem. Soc.* 135, 1830–1837. doi: 10.1021/ja309570x

Fujino, T., Goto, Y., Suga, H., and Murakami, H. (2016). Ribosomal synthesis of peptides with multiple β-amino acids. *J. Am. Chem. Soc.* 138, 1962–1969. doi: 10.1021/jacs.5b12482

Furano, A. V. (1975). Content of elongation factor Tu in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*72, 4780–4784. doi: 10.1073/pnas.72.12.4780

Gan, F., Liu, R., Wang, F., and Schultz, P. G. (2018). Functional replacement of histidine in proteins to generate noncanonical amino acid dependent organisms. *J. Am. Chem. Soc.* 140, 3829–3832. doi: 10.1021/jacs.7b13452

Gan, R., Perez, J. G., Carlson, E. D., Ntai, I., Isaacs, F. J., Kelleher, N. L., et al. (2017). Translation system engineering in *Escherichia coli* enhances non-canonical

amino acid incorporation into proteins. *Biotechnol. Bioeng* 114, 1074–1086. doi: 10.1002/bit.26239

Gfeller, D., Michielin, O., and Zoete, V. (2013). SwissSidechain: a molecular and structural database of non-natural sidechains. *Nucleic Acids Res.* 41, D327–D332. doi: 10.1093/nar/gks991

Gillies, E. R., Deiss, F., Staedel, C., Schmitter, J. M., and Huc, I. (2007). Development and biological assessment of fully water-soluble helical aromatic amide foldamers. *Angew. Chem. Int. Ed.* 46, 4081–4084. doi: 10.1002/anie.200700301

Goto, Y., Katoh, T., and Suga, H. (2011). Flexizymes for genetic code reprogramming. *Nat. Protoc.* 6, 779–790. doi: 10.1038/nprot.2011.331

Harada, R., Furumoto, S., Yoshikawa, T., Ishikawa, Y., Shibuya, K., Okamura, N., et al. (2016). Synthesis and characterization of (1)(8)F-Interleukin-8 using a cell-free translation system and 4-(1)(8)F-Fluoro-L-Proline. *J. Nucl. Med.* 57, 634–639. doi: 10.2967/jnumed.115.162602

Haruna, K., Alkazemi, M. H., Liu, Y., Soll, D., and Englert, M. (2014). Engineering the elongation factor Tu for efficient selenoprotein synthesis. *Nucleic Acids Res.* 42, 9976–9983. doi: 10.1093/nar/gku691

Hecht, S. M., Alford, B. L., Kuroda, Y., and Kitano, S. (1978). "Chemical aminoacylation" of tRNA's. *J. Biol. Chem.* 253, 4517–4520.

Heckler, T. G., Chang, L. H., Zama, Y., Naka, T., Chorghade, M. S., and Hecht, S. M. (1984). T4 RNA ligase mediated preparation of novel "chemically misacylated" tRNAPheS. *Biochemistry* 23, 1468–1473. doi: 10.1021/bi00302a020

Hong, S. H., Kwon, Y. C., Martin, R. W., Des Soye, B. J., de Paz, A. M., Swonger, K. N., et al. (2015). Improving cell-free protein synthesis through genome engineering of *Escherichia coli* lacking release factor 1. *Chembiochem* 16, 844–853. doi: 10.1002/cbic.201402708

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. doi: 10.1021/sb400140t

Hu, C., Chan, S. I., Sawyer, E. B., Yu, Y., and Wang, J. (2014). Metalloprotein design using genetic code expansion. *Chem. Soc. Rev.* 43, 6498–6510. doi: 10.1039/c4cs00018h

Iijima, I., and Hohsaka, T. (2009). Position-specific incorporation of fluorescent non-natural amino acids into maltose-binding protein for detection of ligand binding by FRET and fluorescence quenching. *Chembiochem* 10, 999–1006. doi: 10.1002/cbic.200800703

Ikeda-Boku, A., Ohno, S., Hibino, Y., Yokogawa, T., Hayashi, N., and Nishikawa, K. (2013). A simple system for expression of proteins containing 3-azidotyrosine at a pre-determined site in *Escherichia coli*. *J. Biochem.* 153, 317–326. doi: 10.1093/jb/mvs153

Italia, J. S., Addy, P. S., Erickson, S. B., Peeler, J. C., Weerapana, E., and Chatterjee, A. (2019). Mutually orthogonal nonsense-suppression systems and conjugation chemistries for precise protein labeling at up to three distinct sites. *J. Am. Chem. Soc.* 141, 6204–6212. doi: 10.1021/jacs.8b12954

Jabar, S., Adams, L. A., Wang, Y., Aurelio, L., Graham, B., and Otting, G. (2017). Chemical tagging with tert-Butyl and trimethylsilyl groups for measuring intermolecular nuclear overhauser effects in a large protein-ligand complex. *Chemistry* 23, 13033–13036. doi: 10.1002/chem.201703531

Johnson, D. B., Wang, C., Xu, J., Schultz, M. D., Schmitz, R. J., Ecker, J. R., et al. (2012). Release factor one is nonessential in *Escherichia coli*. *ACS Chem. Biol.* 7, 1337–1344. doi: 10.1021/cb300229q

Johnson, D. B., Xu, J., Shen, Z., Takimoto, J. K., Schultz, M. D., Schmitz, R. J., et al. (2011). RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat. Chem. Biol.* 7, 779–786. doi: 10.1038/nchembio.657

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. doi: 10.1016/j.ymben.2016.03.002

Katayama, H., Nozawa, K., Nureki, O., Nakahara, Y., and Hojo, H. (2012). Pyrrolysine analogs as substrates for bacterial pyrrolysyl-tRNA synthetase in vitro and in vivo. *Biosci. Biotechnol. Biochem.* 76, 205–208. doi: 10.1271/bbb.110653

Katoh, T., Iwane, Y., and Suga, H. (2017a). Logical engineering of D-arm and T-stem of tRNA that enhances d-amino acid incorporation. *Nucleic Acids Res.* 45, 12601–12610. doi: 10.1093/nar/gkx1129

Katoh, T., Tajima, K., and Suga, H. (2017b). Consecutive elongation of D-Amino acids in translation. *Cell Chem. Biol.* 24, 46–54. doi: 10.1016/j.chembiol.2016.11.012

Katoh, T., and Suga, H. (2018). Ribosomal incorporation of consecutive beta-Amino acids. *J. Am. Chem. Soc.* 140, 12159–12167. doi: 10.1021/jacs.8b07247

Katoh, T., and Suga, H. (2019). Engineering translation components improve incorporation of exotic amino acids. *Int. J. Mol. Sci.* 20:522. doi: 10.3390/ijms20030522

Kawakami, T., Ishizawa, T., and Murakami, H. (2013). Extensive reprogramming of the genetic code for genetically encoded synthesis of highly N-alkylated polycyclic peptidomimetics. *J. Am. Chem. Soc.* 135, 12297–12304. doi: 10.1021/ja405044k

Kawakami, T., Murakami, H., and Suga, H. (2008). Messenger RNA-programmed incorporation of multiple N-methyl-amino acids into linear and cyclic peptides. *Chem. Biol.* 15, 32–42. doi: 10.1016/j.chembiol.2007.12.008

Kawakami, T., Ogawa, K., Hatta, T., Goshima, N., and Natsume, T. (2016). Directed evolution of a cyclized peptoid-peptide chimera against a cell-free expressed protein and proteomic profiling of the interacting proteins to create a protein-protein interaction inhibitor. *ACS Chem. Biol.* 11, 1569–1577. doi: 10.1021/acschembio.5b01014

Kigawa, T., Yamaguchi-Nunokawa, E., Kodama, K., Matsuda, T., Yabuki, T., Matsuda, N., et al. (2002). Selenomethionine incorporation into a protein by cell-free synthesis. *J. Struct. Funct. Genomics* 2, 29–35. doi: 10.1023/a:1013203532303

Kightlinger, W., Duncker, K. E., Ramesh, A., Thames, A. H., Natarajan, A., Stark, J. C., et al. (2019). A cell-free biosynthesis platform for modular construction of protein glycosylation pathways. *Nat. Commun.* 10:5404. doi: 10.1038/s41467-019-12024-9

Kwiatkowski, M., Wang, J., and Forster, A. C. (2014). Facile synthesis of N-acyl-aminoacyl-pCpA for preparation of mischarged fully ribo tRNA. *Bioconjug Chem.* 25, 2086–2091. doi: 10.1021/bc500441b

Kwon, I., and Lim, S. I. (2015). Tailoring the substrate specificity of yeast phenylalanyl-tRNA synthetase toward a phenylalanine analog using multiple-site-specific incorporation. *ACS Synth. Biol.* 4, 634–643. doi: 10.1021/sb500309r

Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H. R., Haimovich, A. D., Kuznetsov, G., et al. (2013). Genomically recoded organisms expand biological functions. *Science* 342, 357–360. doi: 10.1126/science.1241459

Lavickova, B., and Maerkl, S. J. (2019). A simple, robust, and low-cost method to produce the PURE cell-free system. *ACS Synth. Biol.* 8, 455–462. doi: 10.1021/acssynbio.8b00427

Lee, J., Schwieter, K. E., Watkins, A. M., Kim, D. S., Yu, H., Schwarz, K. J., et al. (2019). Expanding the limits of the second genetic code with ribozymes. *Nat. Commun.* 10:5097. doi: 10.1038/s41467-019-12916-w

Lee, K. B., Hou, C. Y., Kim, C. E., Kim, D. M., Suga, H., and Kang, T. J. (2016). Genetic code expansion by degeneracy reprogramming of arginyl codons. *Chembiochem* 17, 1198–1201. doi: 10.1002/cbic.201600111

Lee, K.-H., Catherine, C., and Kim, D.-M. (2016). Enhanced production of unnatural amino acid-containing proteins in a cell-free protein synthesis system. *J. Ind. Eng. Chem.* 37, 90–94. doi: 10.1016/j.jiec.2016.03.008

Liu, W.-Q., Zhang, L., Chen, M., and Li, J. (2019). Cell-free protein synthesis: recent advances in bacterial extract sources and expanded applications. *Biochem. Eng. J.* 141, 182–189. doi: 10.1016/j.bej.2018.10.023

Lodder, M., Wang, B., and Hecht, S. M. (2005). The N-pentenoyl protecting group for aminoacyl-tRNAs. *Methods* 36, 245–251. doi: 10.1016/j.ymeth.2005.04.002

Loscha, K. V., Herlt, A. J., Qi, R., Huber, T., Ozawa, K., and Otting, G. (2012). Multiple-site labeling of proteins with unnatural amino acids. *Angew. Chem. Int. Ed. Engl.* 51, 2243–2246. doi: 10.1002/anie.201108275

Lu, Y. (2017). Cell-free synthetic biology: engineering in an open world. *Synth. Syst. Biotechnol.* 2, 23–27. doi: 10.1016/j.synbio.2017.02.003

Maini, R., Nguyen, D. T., Chen, S., Dedkova, L. M., Chowdhury, S. R., Alcala-Torano, R., et al. (2013). Incorporation of beta-amino acids into dihydrofolate reductase by ribosomes having modifications in the peptidyltransferase center. *Bioorg. Med. Chem.* 21, 1088–1096. doi: 10.1016/j.bmc.2013.01.002

Martin, R. W., Des Soye, B. J., Kwon, Y. C., Kay, J., Davis, R. G., Thomas, P. M., et al. (2018). Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat. Commun.* 9:1203. doi: 10.1038/s41467-018-03469-5

Mat, W. K., Xue, H., and Wong, J. T. (2010). Genetic code mutations: the breaking of a three billion year invariance. *PLoS One* 5:e12206. doi: 10.1371/journal.pone.0012206

Melnikov, S. V., Khabibullina, N. F., Mairhofer, E., Vargas-Rodriguez, O., Reynolds, N. M., Micura, R., et al. (2019). Mechanistic insights into the slow peptide bond formation with D-amino acids in the ribosomal active site. *Nucleic Acids Res.* 47, 2089–2100. doi: 10.1093/nar/gky1211

Melo Czekster, C., Robertson, W. E., Walker, A. S., Soll, D., and Schepartz, A. (2016). In vivo biosynthesis of a beta-amino acid-containing protein. *J. Am. Chem. Soc.* 138, 5194–5197. doi: 10.1021/jacs.6b01023

Merryman, C., and Green, R. (2004). Transformation of aminoacyl tRNAs for the in vitro selection of "drug-like" molecules. *Chem. Biol.* 11, 575–582. doi: 10.1016/j.chembiol.2004.03.009

Montclare, J. K., Son, S., Clark, G. A., Kumar, K., and Tirrell, D. A. (2009). Biosynthesis and stability of coiled-coil peptides containing (2S,4R)-5,5,5-trifluoroleucine and (2S,4S)-5,5,5-trifluoroleucine. *Chembiochem* 10, 84–86. doi: 10.1002/cbic.200800164

Morimoto, J., Hayashi, Y., Iwasaki, K., and Suga, H. (2011). Flexizymes: their evolutionary history and the origin of catalytic function. *Acc. Chem. Res.* 44, 1359–1368. doi: 10.1021/ar2000953

Mukai, T., Hayashi, A., Iraha, F., Sato, A., Ohtake, K., Yokoyama, S., et al. (2010). Codon reassignment in the *Escherichia coli* genetic code. *Nucleic Acids Res.* 38, 8188–8195. doi: 10.1093/nar/gkq707

Mukai, T., Hoshi, H., Ohtake, K., Takahashi, M., Yamaguchi, A., Hayashi, A., et al. (2015). Highly reproductive *Escherichia coli* cells with no specific assignment to the UAG codon. *Sci. Rep.* 5:9699. doi: 10.1038/srep09699

Mukai, T., Yanagisawa, T., Ohtake, K., Wakamori, M., Adachi, J., Hino, N., et al. (2011). Genetic-code evolution for protein synthesis with non-natural amino acids. *Biochem. Biophys. Res. Commun.* 411, 757–761. doi: 10.1016/j.bbrc.2011.07.020

Murakami, H., Ohta, A., Ashigai, H., and Suga, H. (2006). A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods* 3, 357–359. doi: 10.1038/nmeth877

Narumi, R., Masuda, K., Tomonaga, T., Adachi, J., Ueda, H. R., and Shimizu, Y. (2018). Cell-free synthesis of stable isotope-labeled internal standards for targeted quantitative proteomics. *Synth. Syst. Biotechnol.* 3, 97–104. doi: 10.1016/j.synbio.2018.02.004

Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J. W. (2010). Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464, 441–444. doi: 10.1038/nature08817

Ninomiya, K., Kurita, T., Hohsaka, T., and Sisido, M. (2003). Facile aminoacylation of pdCpA dinucleotide with a nonnatural amino acid in cationic micelle. *Chem. Commun.* 7, 2242–2243. doi: 10.1039/b306011j

Niwa, N., Yamagishi, Y., Murakami, H., and Suga, H. (2009). A flexizyme that selectively charges amino acids activated by a water-friendly leaving group. *Bioorg. Med. Chem. Lett.* 19, 3892–3894. doi: 10.1016/j.bmcl.2009.03.114

Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989). A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* 244, 182–188. doi: 10.1126/science.2649980

O'Donoghue, P., Prat, L., Heinemann, I. U., Ling, J., Odoi, K., Liu, W. R., et al. (2012). Near-cognate suppression of amber, opal and quadruplet codons competes with aminoacyl-tRNAPyl for genetic code expansion. *FEBS Lett.* 586, 3931–3937. doi: 10.1016/j.febslet.2012.09.033

Ogawa, A., Nishi, T., Sando, S., and Aoyama, Y. (2005). In vitro selection of RNA aptamers for the *Escherichia coli* release factor 1. *Nucleic Acids Symp. Ser.* 2005, 269–270. doi: 10.1093/nass/49.1.269

Ohshiro, Y., Nakajima, E., Goto, Y., Fuse, S., Takahashi, T., Doi, T., et al. (2011). Ribosomal synthesis of backbone-macrocyclic peptides containing gamma-amino acids. *Chembiochem* 12, 1183–1187. doi: 10.1002/cbic.201100104

Ohta, A., Murakami, H., and Suga, H. (2008). Polymerization of alpha-hydroxy acids by ribosomes. *Chembiochem* 9, 2773–2778. doi: 10.1002/cbic.200800439

Ohtake, K., Sato, A., Mukai, T., Hino, N., Yokoyama, S., and Sakamoto, K. (2012). Efficient decoding of the UAG triplet as a full-fledged sense codon enhances the growth of a prfA-deficient strain of *Escherichia coli*. *J. Bacteriol.* 194, 2606–2613. doi: 10.1128/JB.00195-12

Oki, K., Sakamoto, K., Kobayashi, T., Sasaki, H. M., and Yokoyama, S. (2008). Transplantation of a tyrosine editing domain into a tyrosyl-tRNA synthetase variant enhances its specificity for a tyrosine analog. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13298–13303. doi: 10.1073/pnas.0803531105

Ortmayer, M., Fisher, K., Basran, J., Wolde-Michael, E. M., Heyes, D. J., Levy, C., et al. (2020). Rewiring the "Push-Pull" catalytic machinery of a heme enzyme using an expanded genetic code. *ACS Catalysis* 10, 2735–2746. doi: 10.1021/acscatal.9b05129

Oza, J. P., Aerni, H. R., Pirman, N. L., Barber, K. W., Ter Haar, C. M., Rogulina, S., et al. (2015). Robust production of recombinant phosphoproteins using cell-free protein synthesis. *Nat. Commun.* 6:8168. doi: 10.1038/ncomms9168

Ozer, E., Chemla, Y., Schlesinger, O., Aviram, H. Y., Riven, I., Haran, G., et al. (2017). In vitro suppression of two different stop codons. *Biotechnol. Bioeng.* 114, 1065–1073. doi: 10.1002/bit.26226

Park, H. S., Hohn, M. J., Umehara, T., Guo, L. T., Osborne, E. M., Benner, J., et al. (2011). Expanding the genetic code of *Escherichia coli* with phosphoserine. *Science* 333, 1151–1154. doi: 10.1126/science.1207203

Passioura, T., and Suga, H. (2013). Flexizyme-mediated genetic reprogramming as a tool for noncanonical peptide synthesis and drug discovery. *Chemistry* 19, 6530–6536. doi: 10.1002/chem.201300247

Peuker, S., Andersson, H., Gustavsson, E., Maiti, K. S., Kania, R., Karim, A., et al. (2016). Efficient isotope editing of proteins for site-directed vibrational spectroscopy. *J. Am. Chem. Soc.* 138, 2312–2318. doi: 10.1021/jacs.5b12680

Polycarpo, C. R., Herring, S., Berube, A., Wood, J. L., Soll, D., and Ambrogelly, A. (2006). Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase. *FEBS Lett.* 580, 6695–6700. doi: 10.1016/j.febslet.2006.11.028

Ravikumar, Y., Nadarajan, S. P., Yoo, T. H., Lee, C. S., and Yun, H. (2015). Unnatural amino acid mutagenesis-based enzyme engineering. *Trends Biotechnol.* 33, 462–470. doi: 10.1016/j.tibtech.2015.05.002

Rogers, J. M., Kwon, S., Dawson, S. J., Mandal, P. K., Suga, H., and Huc, I. (2018). Ribosomal synthesis and folding of peptide-helical aromatic foldamer hybrids. *Nat. Chem.* 10, 405–412. doi: 10.1038/s41557-018-0007-x

Rogerson, D. T., Sachdeva, A., Wang, K., Haq, T., Kazlauskaite, A., Hancock, S. M., et al. (2015). Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol.* 11, 496–503. doi: 10.1038/nchembio.1823

Rothman, D. M., Petersson, E. J., Vazquez, M. E., Brandt, G. S., Dougherty, D. A., and Imperiali, B. (2005). Caged phosphoproteins. *J. Am. Chem. Soc.* 127, 846–847. doi: 10.1021/ja043875c

Salehi, A. S. M., Smith, M. T., Schinn, S. M., Hunt, J. M., Muhlestein, C., Diray-Arce, J., et al. (2017). Efficient tRNA degradation and quantification in *Escherichia coli* cell extract using RNase-coated magnetic beads: a key step toward codon emancipation. *Biotechnol. Prog.* 33, 1401–1407. doi: 10.1002/btpr.2511

Sando, S., Ogawa, A., Nishi, T., Hayami, M., and Aoyama, Y. (2007). In vitro selection of RNA aptamer against *Escherichia coli* release factor 1. *Bioorg. Med. Chem. Lett.* 17, 1216–1220. doi: 10.1016/j.bmcl.2006.12.013

Schinn, S. M., Bradley, W., Groesbeck, A., Wu, J. C., Broadbent, A., and Bundy, B. C. (2017). Rapid in vitro screening for the location-dependent effects of unnatural amino acids on protein expression and activity. *Biotechnol. Bioeng.* 114, 2412–2417. doi: 10.1002/bit.26305

Seki, E., Yanagisawa, T., and Yokoyama, S. (2018). Cell-free protein synthesis for multiple site-specific incorporation of noncanonical amino acids using cell extracts from rf-1 deletion *E. coli* Strains. *Methods Mol. Biol.* 1728, 49–65. doi: 10.1007/978-1-4939-7574-7_3

Shafer, A. M., Kálai, T., Liu, S. Q. B., Hideg, K., and Voss, J. C. (2004). Site-specific insertion of spin-labeled L-amino acids in *Xenopus* oocytes. *Biochemistry* 43, 8470–8482. doi: 10.1021/bi035542i

Shepherd, T. R., Du, L., Liljeruhm, J., Samudyata, Wang, J., and Sjödin, M. O. D. (2017). De novo design and synthesis of a 30-cistron translation-factor module. *Nucleic Acids Res.* 45, 10895–10905. doi: 10.1093/nar/gkx753

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., et al. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755. doi: 10.1038/90802

Short, G. F. III, Golovine, S. Y., and Hecht, S. M. (1999). Effects of release factor 1 on in vitro protein translation and the elaboration of proteins containing unnatural amino acids. *Biochemistry* 38, 8808–8819. doi: 10.1021/bi990281r

Si, L. L., Xu, H., Zhou, X. Y., Zhang, Z. W., Tian, Z. Y., Wang, Y., et al. (2016). Generation of influenza A viruses as live but replication-incompetent virus vaccines. *Science* 354, 1170–1173. doi: 10.1126/science.aah5869

Silverman, A. D., Karim, A. S., and Jewett, M. C. (2019). Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* 21, 151–170. doi: 10.1038/s41576-019-0186-3

Singh-Blom, A., Hughes, R. A., and Ellington, A. D. (2014). An amino acid depleted cell-free protein synthesis system for the incorporation of non-canonical amino acid analogs into proteins. *J. Biotechnol.* 178, 12–22. doi: 10.1016/j.jbiotec.2014.02.009

Smolskaya, S., Logashina, Y. A., and Andreev, Y. A. (2020). *Escherichia coli* Extract-based cell-free expression system as an alternative for difficult-to-obtain protein biosynthesis. *Int. J. Mol. Sci.* 21:928. doi: 10.3390/ijms21030928

Stech, M., Nikolaeva, O., Thoring, L., Stocklein, W. F. M., Wustenhagen, D. A., Hust, M., et al. (2017). Cell-free synthesis of functional antibodies using a coupled in vitro transcription-translation system based on CHO cell lysates. *Sci. Rep.* 7:12030. doi: 10.1038/s41598-017-12364-w

Sterner, D. E., and Berger, S. L. (2000). Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.* 64, 435–459. doi: 10.1128/mmbr.64.2.435-459.2000

Suga, H. (2018). Max-Bergmann award lecture:A RaPID way to discover bioactive nonstandard peptides assisted by the flexizyme and FIT systems. *J. Pept. Sci.* 24:e3055. doi: 10.1002/psc.3055

Szkaradkiewicz, K., Nanninga, M., Nesper-Brock, M., Gerrits, M., Erdmann, V. A., and Sprinzl, M. (2002). RNA aptamers directed against release factor 1 from *Thermus thermophilus*. *FEBS Lett.* 514, 90–95. doi: 10.1016/s0014-5793(02)02308-6

Tajima, K., Katoh, T., and Suga, H. (2018). Genetic code expansion via integration of redundant amino acid assignment by finely tuning tRNA pools. *Curr. Opin. Chem. Biol.* 46, 212–218. doi: 10.1016/j.cbpa.2018.07.010

Thyer, R., Filipovska, A., and Rackham, O. (2013). Engineered rRNA enhances the efficiency of selenocysteine incorporation during translation. *J. Am. Chem. Soc.* 135, 2–5. doi: 10.1021/ja3069177

Tomita, K., Ogawa, T., Uozumi, T., Watanabe, K., and Masaki, H. (2000). A cytotoxic ribonuclease which specifically cleaves four isoaccepting arginine tRNAs at their anticodon loops. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8278–8283. doi: 10.1073/pnas.140213797

Urbanek, A., Elena-Real, C. A., Popovic, M., Morato, A., Fournet, A., Allemand, F., et al. (2020). Site-specific isotopic labeling (SSIL): access to high-resolution structural and dynamic information in low-complexity proteins. *Chembiochem* 21, 769–775. doi: 10.1002/cbic.201900583

Urbanek, A., Morató, A., Allemand, F., Delaforge, E., Fournet, A., Popovic, M., et al. (2018). A general strategy to access structural information at atomic resolution in polyglutamine homorepeats. *Angew. Chem.* 130, 3660–3663. doi: 10.1002/ange.201711530

Venkat, S., Sturges, J., Stahman, A., Gregory, C., Gan, Q., and Fan, C. (2018). Genetically incorporating two distinct post-translational modifications into one protein simultaneously. *ACS Synth. Biol.* 7, 689–695. doi: 10.1021/acssynbio.7b00408

Villarreal, F., Contreras-Llano, L. E., Chavez, M., Ding, Y., Fan, J., Pan, T., et al. (2018). Synthetic microbial consortia enable rapid assembly of pure translation machinery. *Nat. Chem. Biol.* 14, 29–35. doi: 10.1038/nchembio.2514

Wakamori, M., Fujii, Y., Suka, N., Shirouzu, M., Sakamoto, K., Umehara, T., et al. (2015). Intra- and inter-nucleosomal interactions of the histone H4 tail revealed with a human nucleosome core particle with genetically-incorporated H4 tetra-acetylation. *Sci. Rep.* 5:17204. doi: 10.1038/srep17204

Wan, W., Huang, Y., Wang, Z., Russell, W. K., Pai, P. J., Russell, D. H., et al. (2010). A facile system for genetic incorporation of two different noncanonical amino acids into one protein in *Escherichia coli*. *Angew. Chem. Int. Ed. Engl.* 49, 3211–3214. doi: 10.1002/anie.201000465

Wang, H. H., Huang, P. Y., Xu, G., Haas, W., Marblestone, A., Li, J., et al. (2012). Multiplexed in vivo His-tagging of enzyme pathways for in vitro single-pot multienzyme catalysis. *ACS Synth. Biol.* 1, 43–52. doi: 10.1021/sb3000029

Wang, K., Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2007). Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat. Biotechnol.* 25, 770–777. doi: 10.1038/nbt1314

Wang, K., Sachdeva, A., Cox, D. J., Wilf, N. M., Lang, K., Wallace, S., et al. (2014). Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* 6, 393–403. doi: 10.1038/nchem.1919

Wang, L. (2017). Engineering the genetic code in cells and animals: biological considerations and impacts. *Acc. Chem. Res.* 50, 2767–2775. doi: 10.1021/acs.accounts.7b00376

Wang, T., Liang, C., An, Y., Xiao, S., Xu, H., Zheng, M., et al. (2020). Engineering the translational machinery for biotechnology applications. *Mol. Biotechnol.* 62, 219–227. doi: 10.1007/s12033-020-00246-y

Wang, Y. S., Fang, X., Wallace, A. L., Wu, B., and Liu, W. R. (2012). A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum. *J. Am. Chem. Soc.* 134, 2950–2953. doi: 10.1021/ja211972x

Wannier, T. M., Kunjapur, A. M., Rice, D. P., McDonald, M. J., Desai, M. M., and Church, G. M. (2018). Adaptive evolution of genomically recoded *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 115, 3090–3095. doi: 10.1073/pnas.1715530115

Worst, E. G., Exner, M. P., De Simone, A., Schenkelberger, M., Noireaux, V., Budisa, N., et al. (2015). Cell-free expression with the toxic amino acid canavanine. *Bioorg. Med. Chem. Lett.* 25, 3658–3660. doi: 10.1016/j.bmcl.2015.06.045

Worst, E. G., Exner, M. P., De Simone, A., Schenkelberger, M., Noireaux, V., Budisa, N., et al. (2016). Residue-specific incorporation of noncanonical amino acids into model proteins using an *Escherichia coli* Cell-free Transcription-translation System. *J. Vis. Exp.* 1:54273. doi: 10.3791/54273

Xiao, H., Murakami, H., Suga, H., and Ferre-D'Amare, A. R. (2008). Structural basis of specific tRNA aminoacylation by a small in vitro selected ribozyme. *Nature* 454, 358–361. doi: 10.1038/nature07033

Xiao, H., Peters, F. B., Yang, P. Y., Reed, S., Chittuluru, J. R., and Schultz, P. G. (2014). Genetic incorporation of histidine derivatives using an engineered pyrrolysyl-tRNA synthetase. *ACS Chem. Biol.* 9, 1092–1096. doi: 10.1021/cb500032c

Xiong, H., Reynolds, N. M., Fan, C., Englert, M., Hoyer, D., Miller, S. J., et al. (2016). Dual genetic encoding of acetyl-lysine and non-deacetylatable thioacetyl-lysine mediated by flexizyme. *Angew. Chem. Int. Ed. Engl.* 55, 4083–4086. doi: 10.1002/anie.201511750

Yamanaka, K., Nakata, H., Hohsaka, T., and Sisido, M. (2004). Efficient synthesis of nonnatural mutants in *Escherichia coli* S30 in vitro protein synthesizing system. *J. Biosci. Bioeng.* 97, 395–399. doi: 10.1016/s1389-1723(04)70225-x

Yanagisawa, T., Takahashi, M., Mukai, T., Sato, S., Wakamori, M., Shirouzu, M., et al. (2014). Multiple site-specific installations ofNε-Monomethyl-L-Lysine into histone proteins by cell-based and cell-free protein synthesis. *ChemBioChem* 15, 1830–1838. doi: 10.1002/cbic.201402291

Yin, G., Stephenson, H. T., Yang, J., Li, X., Armstrong, S. M., Heibeck, T. H., et al. (2017). RF1 attenuation enables efficient non-natural amino acid incorporation for production of homogeneous antibody drug conjugates. *Sci. Rep.* 7:3026. doi: 10.1038/s41598-017-03192-z

Zhang, M. S., Brunner, S. F., Huguenin-Dezot, N., Liang, A. D., Schmied, W. H., Rogerson, D. T., et al. (2017). Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nat. Methods* 14, 729–736. doi: 10.1038/nmeth.4302

Zheng, Y., Gilgenast, M. J., Hauc, S., and Chatterjee, A. (2018). capturing post-translational modification-triggered protein-protein interactions using dual noncanonical amino acid mutagenesis. *ACS Chem. Biol.* 13, 1137–1141. doi: 10.1021/acschembio.8b00021

Zhu, P., Gafken, P. R., Mehl, R. A., and Cooley, R. B. (2019). A highly versatile expression system for the production of multiply phosphorylated proteins. *ACS Chem. Biol.* 14, 1564–1572. doi: 10.1021/acschembio.9b00307

Zimmerman, E. S., Heibeck, T. H., Gill, A., Li, X., Murray, C. J., and Madlansacay, M. R. (2014). Production of site-specific antibody-drug conjugates using optimized non-natural amino acids in a cell-free expression system. *Bioconjug Chem.* 25, 351–361. doi: 10.1021/bc400490z

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

ORIGINAL RESEARCH
published: 06 August 2020
doi: 10.3389/fbioe.2020.00854

Modular Enzymatic Cascade Synthesis of Nucleotides Using a (d)ATP Regeneration System

*Maryke Fehlau[1,2], Felix Kaspar[1,2], Katja F. Hellendahl[1], Julia Schollmeyer[1,2], Peter Neubauer[1] and Anke Wagner[1,2]**

[1] Chair of Bioprocess Engineering, Institute of Biotechnology, Faculty III Process Sciences, Technische Universität Berlin, Berlin, Germany, [2] BioNukleo GmbH, Berlin, Germany

Nucleoside-5'-triphosphates (NTPs) and their analogs are building blocks of DNA and are important compounds in both pharmaceutical and molecular biology applications. Currently, commercially available base or sugar modified NTPs are mainly synthesized chemically. Since the chemical production of NTPs is time-consuming and generally inefficient, alternative approaches are under development. Here we present a simple, efficient and generalizable enzymatic synthesis method for the conversion of nucleosides to NTPs. Our one-pot method is modular, applicable to a wide range of natural and modified nucleotide products and accesses NTPs directly from cheap nucleoside precursors. Nucleoside kinases, nucleoside monophosphate (NMP) kinases and a nucleoside diphosphate (NDP) kinase were applied as biocatalysts. Enzymes with different substrate specificities were combined to produce derivatives of adenosine and cytidine triphosphate with conversions of 4 to 26%. The implementation of a (deoxy)ATP recycling system resulted in a significant increase in the conversion to all NTP products, furnishing 4 different NTPs in quantitative conversion. Natural (deoxy)NTPs were synthesized with 60 to >99% conversion and sugar- and base-modified NTPs were produced with 69 to >99% and 27 to 75% conversion, respectively. The presented method is suitable for the efficient synthesis of a wide range of natural and modified NTPs in a sustainable one-pot process.

Keywords: enzymatic cascade synthesis, nucleoside-5′-triphosphate, one-pot multi-enzyme reaction, nucleotide analog, nucleotide kinase, nucleoside kinase, modular, ATP regeneration system

OPEN ACCESS

Edited by:
Simon J. Moore, University of Kent, United Kingdom

Reviewed by:
Lothar Elling, RWTH Aachen University, Germany
Zongbao K. Zhao, Dalian Institute of Chemical Physics (CAS), China

***Correspondence:**
Anke Wagner
anke.wagner@tu-berlin.de

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

Received: *03 May 2020*
Accepted: *02 July 2020*
Published: *06 August 2020*

Citation:
Fehlau M, Kaspar F, Hellendahl KF, Schollmeyer J, Neubauer P and Wagner A (2020) Modular Enzymatic Cascade Synthesis of Nucleotides Using a (d)ATP Regeneration System. Front. Bioeng. Biotechnol. 8:854. doi: 10.3389/fbioe.2020.00854

INTRODUCTION

Modified nucleotides are important small molecules in molecular biology and pharmaceutical applications. Natural and modified nucleoside-5'-triphosphates (NTPs) are valuable building blocks for PCR, fluorescent *in situ* hybridization (FISH), aptamer production as well as for next generation sequencing (Prober et al., 1987; Zaccolo et al., 1996; Giller et al., 2003; Ni et al., 2017). Nucleotides with an azide or alkyne function enable the post-synthetic modification of oligonucleotides via click-chemistry (Gramlich et al., 2008).

Furthermore, the application of nucleotide prodrugs is of increasing interest (Pradere et al., 2014) as most of the known nucleoside analog drugs are only active as the respective nucleoside

diphosphate (NDP) or NTP and activation *in vivo* is often insufficient (Deville-Bonne et al., 2010). In a number of approaches including the application of sofosbuvir or remdesivir, protected nucleoside monophosphates (NMPs) were administered to overcome the first activation step *in vivo* (Pradere et al., 2014; Ko et al., 2020). Additionally, methods have been developed for the production of NDP or NTP prodrugs with an increased biological availability of the respective nucleoside drug as shown for sofosbuvir or remdesivir (Pradere et al., 2014).

To date, nucleotides are primarily prepared by chemical methods such as the Yoshikawa protocol or the Ludwig-Eckstein method (Burgess and Cook, 2000). A common disadvantage of these multistep synthesis reactions is a limited regioselectivity leading to the formation of different phosphorylation products, as well as the limited control over the exclusive formation of the triphosphate, as higher phosphorylation products can be generated. The overall process is laborious, involves toxic reagents and often achieves only moderate product yields. Furthermore, some nucleotide products harboring sensitive functional groups are not accessible in these approaches as they do not withstand the harsh reaction conditions and generate undesired side products (Lee and Momparler, 1976; Burgess and Cook, 2000).

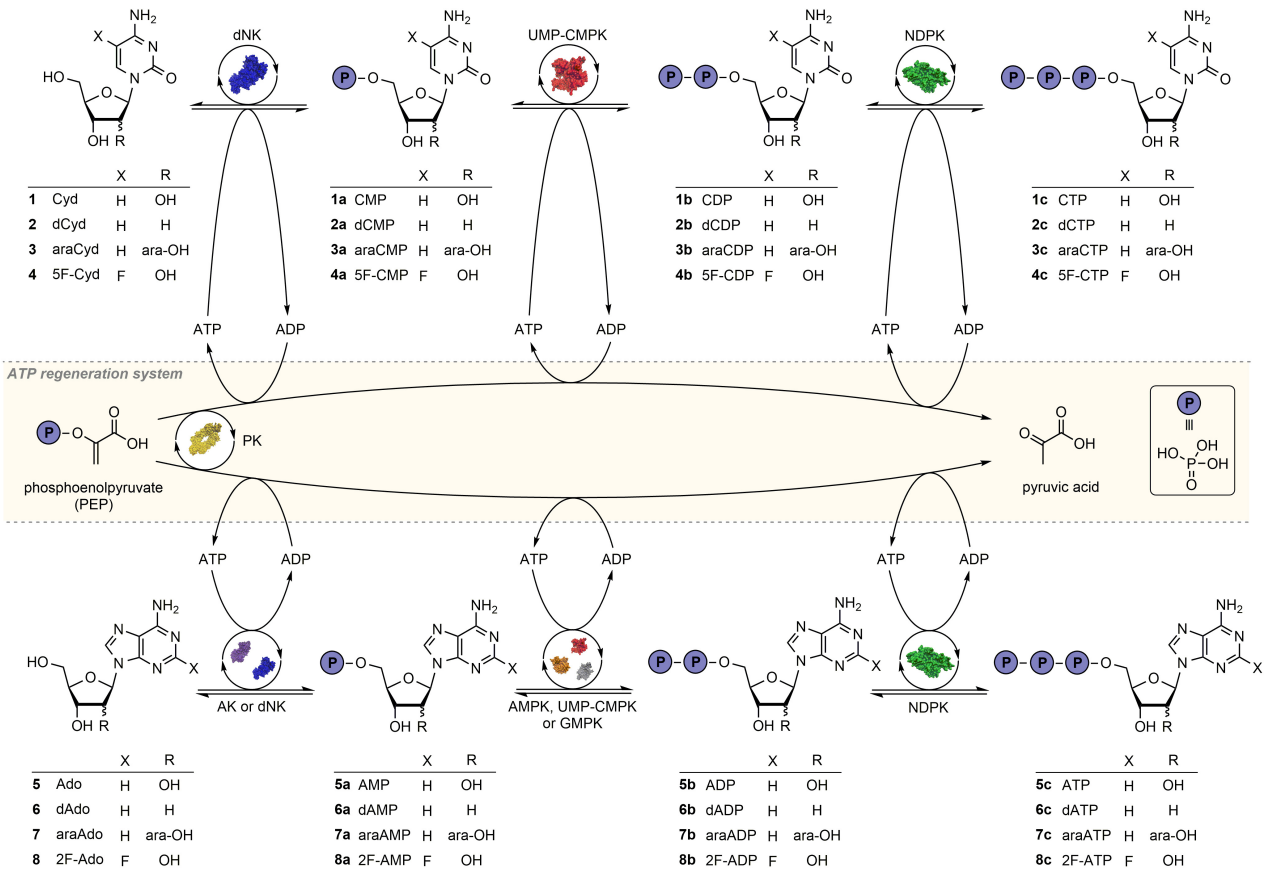
Biocatalytic routes exploiting enzymes from the nucleotide metabolism promise improved regioselectivity and exquisite control over the phosphorylation product. Driven by these advantages, several attempts have been made to replace chemical synthesis routes for NTP preparation. Natural (deoxy)NTPs have been produced by isolation from animal extracts (Berger, 1956), phosphorylation of RNA/DNA degradation products (Haynie and Whitesides, 1990) or by enzymatic cascade reactions (Ding et al., 2019). While the described methods are suitable for the synthesis of natural NTPs they cannot be easily transferred to the synthesis of NTP analogs bearing sugar or base modifications. Therefore, specific enzymatic synthesis routes were only rarely developed for few NTP analogs (Lee and Momparler, 1976; Da Costa et al., 2007; Hennig et al., 2007).

The application of nucleosides as substrates for the synthesis of nucleotides together with purified biocatalysts offers significant advantages over the approaches described above. Natural nucleosides are generally cheaply available as commercial reagents and modified nucleosides can be accessed enzymatically via known methods in high yield (Kamel et al., 2019; Kaspar et al., 2020; Yehia et al., 2020). The use of purified enzymes minimizes side reactions and simplifies reaction workup. Despite these advantages, however, there are only few examples in the literature of NTP synthesis from nucleosides with purified enzymes. As an example Baughn et al. (1978) used purified adenosine kinase, adenosine monophosphate kinase, acetate kinase and acetyl phosphate to convert adenosine to adenosine triphosphate (ATP) in a one-pot approach with 80% conversion. However, small amounts of nucleotide products ATP, adenosine diphosphate (ADP) and adenosine monophosphate (AMP) were needed to start the reaction. To our knowledge, only a single one-pot approach for the synthesis of a modified NTP from a nucleoside has been described by Hennig et al. (2007) who produced 5-fluorocytidine triphosphate (5F-CTP, **4c**). The compound was prepared with 78% isolated yield using a uridine kinase, nucleoside monophosphate kinase, pyruvate kinase as well as an enolase and 3-phosphoglycerate mutase. ATP and 3-phosphoglycerate were used as phosphate donors.

Phosphate donor recycling systems promise increased yields in these reactions by shifting the reaction equilibrium toward the desired product NTP (Abu and Woodley, 2015). These recycling systems are coupled enzymatic reactions which constantly (re)generate NTPs from the respective NDPs and a secondary phosphate donor. Indeed, the incorporation in nucleotide synthesis reactions has been described to enable higher conversions via an equilibrium shift (Wu et al., 2003). The recovery of (d)ATP from (d)ADP lowers substrate and product inhibition effects by decreasing the concentration of the side product (d)ADP. At the same time, the concentration of (d)ATP is kept high to facilitate the kinase reactions. Further advantages of phosphate donor recycling systems include simplified product purification and cost savings by decreasing the amount of (d)ATP required for high conversions.

For the enzymatic regeneration of (d)NTPs acetate kinase, pyruvate kinase and polyphosphate kinase are most commonly used (Endo and Koizumi, 2001; Andexer and Richter, 2015). Those enzymes possess high substrate promiscuity and are able to (re)generate most natural and some modified (d)NTPs from the respective (d)NDPs (Ishige et al., 2001; Gao et al., 2008; Zou et al., 2013). For example, pyruvate kinase has been applied for both the conversion of 5-fluorinated NDPs to 5-fluorinated NTPs as well as the regeneration of the phosphate donor ATP for other coupled enzymatic reactions (Hennig et al., 2007). Regeneration systems also offer the opportunity to reduce the concentration of non-ATP phosphate donors as shown for the enzymatic synthesis of natural dNMPs from nucleosides using the comparably expensive phosphate donor guanosine triphosphate (GTP). Due to a constant regeneration from the secondary phosphate donor acetyl phosphate only small amounts of GTP (2.5 mol%) were needed (Zou et al., 2013). Inorganic polyphosphate is the cheapest available secondary phosphate donor. However, energy-rich phosphate donors such as phosphoenolpyruvate (PEP) offer a significant thermodynamic advantage over other possible phosphate donors (Andexer and Richter, 2015). This allows ATP (re)generation in nearly quantitative fashion and prompted us to employ this system for this study.

Inspired by the initial success of Hennig and colleagues, we herein report a pilot study to develop a generalizable method for the synthesis of natural and modified NTPs from nucleosides as cheaply available precursors (**Scheme 1**). Our modular one-pot four-enzyme cascade with a PEP-based phosphate donor recycling system allows the simple and efficient production of a wide range of NTPs, as demonstrated by examples bearing pyrimidine and purine bases as well as several different sugars. Nucleoside, NMP and NDP kinases were applied as biocatalysts using a standardized protocol. The application of different enzyme combinations allowed for a rapid adjustment to substrates of interest. Furthermore, the impact of a phosphate donor regeneration system on NTP yields was quantified for the first time. Significantly higher conversions were achieved for natural (4 to 9 times) and modified NTP products (4 to 6 times)



SCHEME 1 | Reaction sequence for the phosphorylation of natural and modified nucleosides in a multi-enzyme cascade reaction with an ATP regeneration system based on pyruvate kinase and phosphoenolpyruvate. Depending on the substrate the enzymes *Drosophila melanogaster* deoxynucleoside kinase (dNK), adenosine kinase (AK), UMP-CMP kinase (UMP-CMPK), guanylate kinase (GMPK), adenylate kinase (AMPK), nucleoside diphosphate kinase (NDPK) and pyruvate kinase (PK) were applied. The protein structures shown belong to the PDB entries 2vqs (dNK), 3uq6 (AK), 2anc (GMPK), 1e4v (AMPK), 2cmk (UMP-CMPK), 2hur (NDPK) and 1pky (PK).

using (d)ATP regeneration in a one-pot system, providing 4 NTPs in quantitative conversion. The present study lays the groundwork for future high-yielding syntheses of NTPs from nucleosides.

MATERIALS AND METHODS

General Information

All chemicals and solvents were of analytical grade or higher and purchased, if not stated otherwise, from Sigma-Aldrich (Steinheim, Germany), Carl Roth (Karlsruhe, Germany), TCI Deutschland (Eschborn, Germany), Carbosynth (Berkshire, United Kingdom) or VWR (Darmstadt, Germany).

The following natural and modified nucleoside substrates were used: Natural pyrimidine and purine nucleoside substrates were cytidine (Cyd, **1**), deoxycytidine (dCyd, **2**), adenosine (Ado, **5**) and deoxyadenosine (dAdo, **6**). Sugar-modified nucleosides were arabinofuranosylcytosine (araCyd, **3**) known as the anti-leukemia drug Cytarabine and arabinofuranosyladenine (araAdo, **7**), a nucleoside antibiotic isolated from Streptomyces antibioticus. Base-modified nucleoside substrates were 5-fluorocytidine (5F-Cyd, **4**) and 2-fluoroadenosine (2F-Ado, **8**). Halogenated nucleotides like 5F-CTP (**4c**) are interesting building blocks for studies of secondary structures of DNA or RNA using NMR analysis (Hennig et al., 2007). 2F-Ado (**8**) is known as Fludarabine and its monophosphorylated form **8a** is used as an anti-cancer chemotherapy drug.

Wild-type nucleoside and nucleotide kinases were obtained from BioNukleo GmbH (Berlin, Germany) except for wide-spectrum deoxynucleoside kinase from *Drosophila melanogaster* (*Dm*dNK). The expression vector of *Dm*dNK was kindly provided by Prof. Birgitte Munch-Petersen (Roskilde University). According to the manufacturer the kinases possess the following substrate specificities: adenosine kinase (AK, NK14), guanylate kinase (GMPK, NMPK21) and adenylate kinase (AMPK, NMPK23) convert purine nucleoside/nucleotide substrates, while uridine monophosphate-cytidine monophosphate kinase (UMP-CMPK, NMPK22) and nucleoside diphosphate kinase (NDPK, NDPK32) accept both purine and pyrimidine nucleoside/nucleotide substrates. All enzymes obtained from BioNukleo were provided as stock solutions (0.1 to 1 mg/mL)

and aliquots stored at −0.20°C until use. Pyruvate kinase (PK, P9136) was obtained from Sigma Aldrich as lyophilized powder, dissolved in 70 mM Tris–HCl pH 7.6 (1.74 mg/mL) and stored in aliquots at −0.20°C. All enzymes are active at 37°C and combinable in the same reaction buffer (70 mM Tris–HCl pH 7.6, 5 mM $MgCl_2$).

Expression of *Dm*dNK

Recombinant *Dm*dNK was expressed and purified as described before (Munch-Petersen et al., 2000). Briefly, the enzyme was produced as a GST-fusion protein using *Escherichia coli* BL21. Following glutathione sepharose purification the GST tag was cleaved of using thrombin and the enzyme was stored at −20°C with 8% glycerol, 1% Triton X-100 and 1 mM DTT.

Enzymatic Cascade Reaction

Enzymatic cascade reactions were performed in a total volume of 50 µL with 70 mM Tris–HCl pH 7.6 (measured at 25°C), 5 mM $MgCl_2$, 1 mM nucleoside substrate and 3.6 mM ATP. 2'-Deoxyadenosine triphosphate (dATP) was used as phosphate donor for reactions with adenosine as substrate to enable substrate and product differentiation. Reactions were started by adding 0.016-0.02 mg/mL (ratio 1:1:1; **Supplementary Table SI**) of each enzyme. Concentrations were chosen based on preliminary experiments. Reactions were incubated at 37°C in a thermocycler with a heatable lid. An equal volume of methanol (50 µL) was added after 19 h to stop the reaction. After centrifugation at 21.500 × *g* and 4°C for 15 min (Himac CT15RE, Hitachi, Tokyo, Japan) 75 µL of the quenched reaction mixture were diluted with 25 µL water and analyzed by high performance liquid chromatography (HPLC) as described below.

Enzymatic Cascade Reaction With Phosphate Donor Recycling System

The phosphate donor regeneration system applied in this study consisted of a regeneration kinase (pyruvate kinase, PK) and a phosphate donor (phosphoenolpyruvate, PEP). Both ATP and dATP were accepted as substrates and were regenerated from the respective (deoxy)nucleoside diphosphates at 37°C. The enzymatic cascade reactions with phosphate donor recycling system were performed in a total volume of 50 µL with 70 mM Tris–HCl pH 7.6, 5 mM $MgCl_2$, 1 mM nucleoside substrate, 3.6 mM ATP or dATP, 5 mM PEP and 0.17 mg/mL PK. Reactions were started by adding 0.016-0.02 mg/mL of each of the enzymes (ratio 1:1:1, **Supplementary Table SI**) and were incubated at 37°C for 19 h. The reactions were stopped by adding equal volumes of methanol (50 µL) followed by centrifugation at 21.500 × *g* and 4°C for 15 min. After centrifugation, 75 µL of the samples were mixed with 25 µL water and analyzed by HPLC.

Time Course of the Enzymatic Cascade Reaction With and Without Regeneration System

To analyze the enzymatic reactions over time, reaction volumes were scaled up to 0.7 mL with and 1 mL without ATP regeneration system. Enzymatic reactions were performed as described above, but incubated in a thermoblock without lid-heating. The reaction tubes were incubated at 300 rpm for 33 h. Regular samples of 50 µL were taken and mixed with equal volumes of methanol followed by centrifugation at 21.500 × *g* and 4°C for 15 min. After centrifugation, 75 µL of the samples were mixed with 25 µL water and analyzed by HPLC.

High Pressure Liquid Chromatography (HPLC)

Samples from nucleoside/nucleotide kinase reactions were analyzed by HPLC-DAD (Agilent 1200 series) with a detection wavelength of 260 nm using a Kinetex Evo column (C18, 100 A, 250 × 4.6 mm, Phenomenex, Aschaffenburg, Germany). The method was adapted from literature (Ryll and Wagner, 1991). The flow rate was set to 1 mL/min at 34°C and the gradient consisted of A (0.1 M KH_2PO_4/K_2HPO_4, 8 mM tetrabutylammonium bisulfate, pH ca. 5.4) and B (70% A, 30% methanol): 0 min – 80% A, 4 min – 80% A, 14 min – 40% A, 26 min – 38% A, 26.5 min – 80% A, and 29 min – 80% A. Reactions with compound 2F-Ado (**8**) as substrate were analyzed using a prolonged gradient: 0 min – 80% A, 4 min – 80% A, 14 min – 40% A, 35 min – 36.5% A, 35.5 min – 80% A, and 38 min – 80% A. Natural nucleosides and nucleotides as well as nucleoside analogs were identified using authentic standards. Nucleotide analog peaks were assigned based on analogy of retention times and characteristic UV absorption spectra. Typical retention times [min] were: Cyd (**1**) – 2.8, CMP (**1a**) – 3.3, CDP (**1b**) – 5.5, CTP (**1c**) – 12.9, dCyd (**2**) – 2.6, dCMP (**2a**) – 3.4, dCDP (**2b**) – 6.3, dCTP (**2c**) – 14.5, araCyd (**3**) – 3, araCMP (**3a**) – 3.5, araCDP (**3b**) – 6.2, araCTP (**3c**) – 13.7, 5F-Cyd (**4**) – 2.8, 5F-CMP (**4a**) – 3.2, 5F-CDP (**4b**) – 5.4, 5F-CTP (**4c**) – 12.8, Ado (**5**) – 5.9, AMP (**5a**) – 7.3, ADP (**5b**) – 14, ATP (**5c**) – 20.8, dAdo (**6**) – 6.4, dAMP (**6a**) – 10.1, dADP (**6b**) – 17.3, dATP (**6c**) – 26.3, araAdo (**7**) – 5.5, araAMP (**7a**) - 8, araADP(**7b**) – 15.5, araATP (**7c**) – 23, 2F-Ado (**8**) - 9, 2F-AMP (**8a**) – 12.8, 2F-ADP (**8b**) – 19.3, 2F-ATP (**8c**) – 29.9 (**Supplementary Figures 2, 3**).

Conversion was calculated as:

$$Conversion\ (X)[\%] = 100 \times \frac{P_X}{P_{total}} \quad (1)$$

where P_X is the peak area of compound X and P_{total} is the sum of all peak areas of the substrate and product(s) in the reaction. Consumption of the cofactor (d)ATP was not considered in the calculation.

RESULTS

Time Course of Enzymatic Cascade Reactions Using Nucleoside and Nucleotide Kinases

We aimed to use a modular enzymatic cascade system to synthesize a range of natural as well as sugar- and base-modified NTPs using the respective nucleosides as substrates. Specifically, starting from nucleosides **1**-**8** we aimed to access cytidine-5′-triphosphate (**1c**) and its 2'-deoxy (**2c**), arabinofuranosyl (**3c**) and

5-fluoro analogs (**4c**) as well as adenosine-5′-triphosphate (**5c**) and its 2′-deoxy (**6c**), arabinofuranosyl (**7c**) and 2-fluoro analogs (**8c**) through a one-pot enzyme cascade reaction (**Scheme 1**). Based on literature data (Serra et al., 2014) we anticipated that a slight excess of ATP would enable access to the NTP products and performed initial experiments with 1.2 equivalents of ATP per product phosphate (3.6 eq. per nucleoside).

In a first set of experiments we sought to confirm if the envisioned enzyme cascade could deliver the expected NTP products and to identify suitable reaction times for the synthesis of these nucleotides. To this end the time course of the conversions of cytidine (Cyd, **1**) and arabinosyl cytosine (araCyd, **3**) (**Figure 1A**) were analyzed over a period of 33 h as exemplary transformations. In both reactions the enzymes *Dm*dNK, UMP-CMPK and NDPK were used as catalysts with ATP serving as the phosphate donor. Enzymes were either chosen based on substrate specificities given by the supplier (UMP-CMPK, NDPK) or due to a known wide substrate spectrum (*Dm*dNK) (Serra et al., 2014).

HPLC analysis of the reaction samples revealed the formation of the desired products, confirming the successful implementation of the enzymatic cascade. With the applied enzyme combination nucleoside analog **3** was faster converted than the natural counterpart **1**. In both reactions NTP products were already detectable at the first time point of sampling (1 h for **1** and 0.5 h for **3**) and maximum concentration was reached after 2 h (25% araCTP, **3c**) and 6 h (28% CTP, **1c**) (**Figures 1B,C**). Although the amount of the NTP products **1c** and **3c** did not change further at that point, conversion of the nucleosides to the NDPs **1b** and **3b** continued until 4 and 19 h after reaction initiation. An equilibrium of the reaction was reached after 19 and 4 h for **1** and **3**, respectively, confirmed by datapoints after 24 and 33 h showing no change in reactant concentrations. At the end of the reaction at 33 h, the final nucleotide ratios (nucleoside:NMP:NDP:NTP) were 0:31:47:22% for **1** and 0:30:53:18% for **3**, respectively. These experiments showed that the selected enzymes indeed delivered the desired phosphorylated products directly from the nucleoside through cascade catalysis. Thus, stability and performance of these biocatalysis under the selected conditions encouraged further exploration of this system. Furthermore, since the equilibrium was reached after 19 h even for the rather sluggish reaction of these enzymes with **1**, we selected this time for the following reactions.

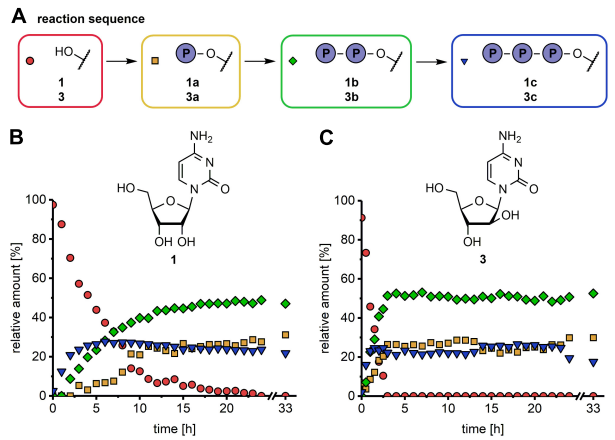


FIGURE 1 | Time course for the synthesis of natural (**B**, 1a–1c) and modified nucleotides (**C**, 3a–3c) in an one-pot multi-enzyme cascade reaction **(A)**. Experimental conditions: 1 mM substrate, 3.6 mM phosphate donor, 70 mM Tris–HCl pH 7.6, 5 mM $MgCl_2$, 0.016-0.02 mg/mL of each enzyme, 37°C. The formation of NMP (yellow square), NDP (green diamond) and NTP (blue triangle) was analyzed over 33 h using the natural nucleoside cytidine (**1**, **B**) and sugar-modified nucleoside analog arabinosylcytosine (**3**, **C**) as substrates (red circle). The enzyme combination *Dm*dNK/UMP-CMPK/NDPK was applied. The first datapoint was taken 30 s after reaction initiation.

One-Pot Enzymatic Cascade Reaction to Produce Natural and Modified Nucleoside Triphosphates

To investigate if this one-pot cascade system is a generalizable method for the synthesis of natural and modified NTPs a wider spectrum of substrates was tested (**Scheme 1** and **Figure 2A**). Therefore, base- and sugar-modified derivatives of both cytidine (**2**-**4**) and adenosine (**6**-**8**) were subjected to enzymatic phosphorylation to access the respective NTPs in one pot. ATP was used as phosphate donor, except for reactions with adenosine as substrate, where dATP was applied as phosphate donor to enable differentiation between substrates and products. All reactions were run for 19 h at 37°C before analysis by HPLC. In initial experiments to produce pyrimidine NTPs **1c**-**4c** *Dm*dNK, UMP-CMPK and NDPK were used as biocatalysts while reactions toward purine nucleotides **5c**-**8c** reactions were performed using AK, GMPK, and NDPK.

Reactions with pyrimidine substrates yielded 11 to 26% of the respective NTP (**Figure 2B**). In reactions with **1**-**3** nearly no nucleoside substrate was left after 19 h reaction time. In contrast, 28% residual **4** was detected at the end of the reaction suggesting a lower activity of first cascade enzyme *Dm*dNK toward the base-modified cytidine analog. In reactions with **1** and **3** comparable amounts of NTP (24 to 26%), NDP (46 to 49%) and NMP (25 to 26%) were formed. With **2** as the substrate only 11% of the corresponding triphosphate **2c** were formed. Conversions to the NMP **2a** and the NDP **2b** of 43% and 46% were observed, respectively. These observations suggest that the second cascade enzyme UMP-CMPK and/or the final enzyme NDPK prefer ribo- and arabino-sugar moieties over deoxyribose.

Contrary to the initial success of the reactions with pyrimidines **1**-**4**, we failed to detect a conversion of purine nucleosides **6**-**8** with the applied enzyme system (**Table 1**). Only Ado (**5**) was efficiently transformed to the corresponding NTP **5c** in 24% conversion.

Therefore, we aimed to improve the product yields for these compounds by employing different enzyme combinations. The exchange of the first cascade enzyme AK with the more promiscuous *Dm*dNK resulted in over 90% conversion of the 2'-deoxyribosyl nucleoside **6** and the arabinosyl nucleoside **7** to the corresponding NMP (**Table 1**). Additionally, the formation of the fluorinated species **8** to the NMP **8a** was improved

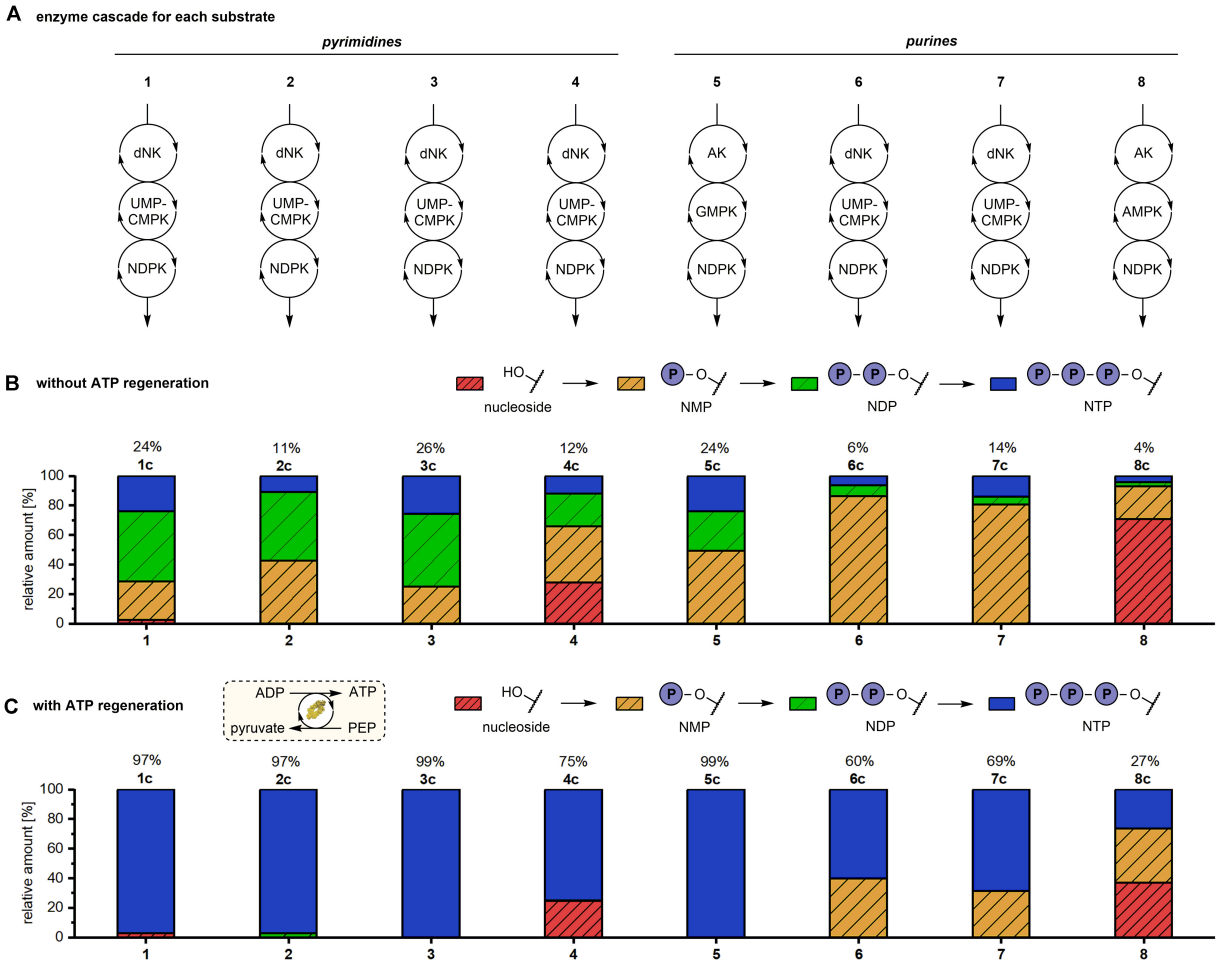


FIGURE 2 | Production of natural and modified NTPs in an one-pot multi-enzyme cascade reaction with and without phosphate donor regeneration system. Experimental conditions: 1 mM substrate, 3.6 mM phosphate donor, 70 mM Tris–HCl pH 7.6, 5 mM $MgCl_2$, 0.016-0.02 mg/mL each enzyme, 37°C, 19 h. Reactions with phosphate donor regeneration **(C)** additionally contained 5 mM phosphoenolpyruvate and 0.17 mg/ml pyruvate kinase. **(A)** Eight natural and modified nucleosides were used as substrates for the multi-enzyme cascade reactions. Applied enzymes: dNK = *Drosophila melanogaster* deoxynucleoside kinase, AK = adenosine kinase, UMP-CMPK = UMP-CMP kinase, GMPK = guanylate kinase, AMPK = adenylate kinase, NDPK = nucleoside diphosphate kinase. **(B,C)** Percentages of Nucleoside, NMP, NDP and NTP without **(B)** and with **(C)** phosphate donor regeneration system. Without ATP regeneration system dCTP conversion could not be determined by HPLC and was therefore calculated based on the dCMP and dCDP conversions.

with this strategy, increasing the conversion from 4 to 22% (**Table 1**). However, the near absence of NDP products in all reactions indicated that conversion from the respective NMP was a limiting factor. Therefore, the second cascade enzyme GMPK was replaced with the more promiscuous UMP-CMPK. This substitution enabled smooth conversion of NMPs **6a** and **7a** to the respective NDPs and NTPs, yielding **6c** and **7c** in 6% and 14% conversion, respectively. Nonetheless, UMP-CMPK showed no activity on the fluorinated **8a** in the one-pot cascade reaction, which delivered only the monophosphorylated product and no detectable formation of the NDP and NTP. Lastly, another purine-specific nucleotide kinase, namely AMPK, when incorporated in the cascade reaction yielded 5% conversion of the fluorinated purine nucleoside to its triphosphate **8c** (**Table 1**).

Taken together, these experiments demonstrate the general feasibility to access several NTPs from their corresponding nucleosides in a one-pot cascade reaction by adjusting the applied enzymes. However, the conversion to the NTPs **1c**-**8c** did not reach satisfactory levels (≤26% in all cases). This spurred us to investigate if a PEP-based (d)ATP recycling system could be employed to improve conversions.

One-Pot Enzymatic Cascade Reaction to Produce Natural and Modified Nucleoside Triphosphates Using a Phosphate Donor Recycling System

Next, the reactions described so far were repeated under the same conditions with a (d)ATP recycling system based on PK and PEP to investigate the impact of cofactor recycling on the one-pot enzymatic cascade reactions. A ratio of substrate to PEP to (d)ATP of 1:5:3.6 was applied, with the aim to avoid the

TABLE 1 | Phosphorylation of adenosine derivatives in a one-pot kinase cascade reaction using different enzyme combinations.

Substrate	Kinase 1		Kinase 2			Kinase 3	Relative amount [%]*			
	*Dm*dNK	AK	GMPK	UMP-CMPK	AMPK	NDPK	Nucleoside	NMP	NDP	NTP
Ado (**5**)		**x**	**x**			**x**	**0**	**49**	**27**	**24**
dAdo (**6**)		x	x			x	100	0	0	0
	x		x			x	0	93	4	3
	x			**x**		**x**	**0**	**86**	**8**	**6**
araAdo (**7**)		x	x			x	100	0	0	0
	x		x			x	0	100	0	0
	x			**x**		**x**	**0**	**80**	**5**	**14**
2F-Ado (**8**)		x	x			x	96	4	0	0
	x		x			x	76	24	0	0
	x			x		x	77	23	0	0
		x			**x**	**x**	**71**	**22**	**3**	**5**

**Experimental conditions: 1 mM substrate, 3.6 mM phosphate donor, 70 mM Tris HCl pH 7.6, 5 mM $MgCl_2$, 0.016-0.02 mg/mL of each enzyme, 37°C, 19 h. The enzyme combinations resulting in the largest NTP formation are depicted in bold. Applied enzymes: DmdNK = Drosophila melanogaster deoxynucleoside kinase, AK = adenosine kinase, UMP-CMPK = UMP-CMP kinase, GMPK = guanylate kinase, AMPK = adenylate kinase, NDPK = nucleoside diphosphate kinase.*

formation of (d)ADP. In experiments with and without (d)ATP regeneration the same substrate to (d)ATP ratio was used to enable direct a comparison of the substrate conversion in both reaction set ups. In literature it was shown before that different ATP to substrate ratios influenced product formation (Serra et al., 2014). All reactions were run for 19 h with 5 equivalents of PEP per nucleoside.

The incorporation of the coupled enzymatic (d)ATP recycling reaction had a strikingly positive effect on the NTP yield for all tested substrates, as determined by HPLC (**Figure 2C**). Natural nucleosides **1**, **2**, **3** and **5** were nearly completely converted to the respective (d)NTPs (≥97%), which is four to nine times higher than the conversions observed without (d)ATP recycling (6 to 24%).

Pyrimidine nucleoside analogs **3** and **4** were converted to their corresponding NTPs in >99% and 75%, respectively. Nonetheless, 25% residual **4** was detected in the reaction mixture after 19 h reaction time. Without ATP recycling approximately the same amount of residual nucleoside substrate was detected while only 12% NTP were formed (**Figure 2B**), indicating that conversion of **4** to the monophosphate **4a** clearly represents a bottleneck in the cascade.

HPLC analysis further revealed that the NTP yield of modified purine nucleotides **7c** and **8c** was improved by factors of 5 and 6, respectively. The conversion to **7c** was increased from 14% to 69% (**Figures 2B,C**). By including the ATP recycling system, the conversion to **8c** increased from 4 to 27% while the percentage of the remaining nucleoside substrate **8** decreased from 71 to 37%. This observation highlights that the inadequate conversion of these nucleosides without cofactor recycling was not due to a kinetically limited step (as seen for **4**), but due to thermodynamic limitations.

The incorporation of the phosphate donor regeneration system further led to reduced reaction times. The time courses of conversion of **1** and **3** to **1c** and **3c** with the ATP regeneration system showed a 4 to 6 times faster product formation compared to our initial experiments (**Supplementary Figure S1**). The maximum amount of both NTPs was approximately 25% without ATP regeneration after 6 and 2 h. HPLC analysis indicated that those NTP yields were already reached after 1 h (**1c**) and 30 min (**3c**) with ATP regeneration. NTP formation continued after that and reached a maximum of >97% after 18 h (**1c**) and 4 h (**3c**) when including the ATP regeneration system.

DISCUSSION

The one-pot enzyme cascade reaction described in this study allows for the efficient synthesis of natural and modified NTPs from nucleosides as substrates. A standardized protocol was used which includes a fixed reaction buffer, enzyme amount and reaction time. This system succeeded in delivering all desired NTPs in various conversions up to 26%. The integration of a phosphate donor regeneration system further improved product yields and reduced reaction times. Highest NTP formation was reached with this system using nucleosides **1**, **2**, **3** and **5** (all >97%).

The synthesis of NTPs from nucleosides offers the advantage that nucleosides are readily available and inexpensive substrates. Furthermore, the use of purified enzymes promises minimization of side reactions and an easier purification process. The feasibility of this approach for the synthesis of modified NTPs has already been successfully demonstrated, albeit only for selected examples. For example, Hennig et al. synthesized 5F-CTP (**4c**) in a one-pot approach with near complete conversion (Hennig et al., 2007). We were also able to show that high conversions can be achieved with nucleoside **4**, since the corresponding triphosphate **4c** was produced with 75% yield after integration of an ATP regeneration system. Nonetheless, a modular enzyme system for the synthesis of multiple NTPs has not been reported to date. We addressed this gap by assaying a number of kinases for their potential to furnish NTPs directly from nucleosides through cascade catalysis.

A crucial factor for the development of a generalizable method is the identification of suitable enzymes. This became evident in the synthesis of ATP derivatives in this study. Significantly higher product yields were observed with *Dm*dNK and UMP-CMPK compared to AK and GMPK as the latter were only suitable enzymes for the phosphorylation of adenosine (**5**). This observation is in good accordance with the available literature, showing that AK and GMPK performed suboptimally with deoxyribose- and arabinose-derivatives (Van Rompay et al., 2000, 2003), while *Dm*dNK and UMP-CMPK were described to accept a wide variety of sugar-modified substrates (Liou et al., 2002; Serra et al., 2014). Interestingly, the time course experiments performed with Cyd (**1**) and araCyd (**3**) even revealed a higher reaction rate of the applied enzymes (*Dm*dNK, UMP-CMPK, NDPK) toward the sugar-modified pyrimidine nucleoside analog **3** compared to the natural counterpart **1**. We ascribed this to *Dm*dNK's preference for deoxyribose- and arabinose-sugar moieties over ribose-containing nucleosides (Munch-Petersen et al., 2000) as well as the fact that CMP (**1a**) but not araCMP (**3a**) leads to significant substrate inhibition of UMP-CMPK (Pasti et al., 2003). In summary, wide-spectrum nucleoside kinases proved most suitable for the synthesis of sugar-modified nucleotides.

In contrast, for the synthesis of base-modified nucleotides like **4c** the use of a pyrimidine specific nucleoside kinase as the first cascade enzyme might be advantageous over the wide substrate-spectrum *Dm*dNK. Mammalian enzymes like uridine-cytidine kinase (UCK) and deoxycytidine kinase (dCK) were described to phosphorylate a variety of base-modified nucleoside analogs like **4**, 5-fluororuridine, 2'-deoxy-5-methylcytidine and 2-thiocytidine (Van Rompay et al., 2003; Hazra et al., 2010). Thus, UCK might be superior over dCK and the *Dm*dNK employed in this study for the phosphorylation of **4** because of a higher activity of UCK toward ribonucleoside substrates (Van Rompay et al., 2003). To identify suitable enzymes for the application in one-pot cascade reactions automated high-throughput assays might be a suitable tool. As shown for nucleoside kinases high-throughput assays allow for the fast and accurate activity screening using a large number of enzymes (Hellendahl/Fehlau et al., in preparation) and we expect future improvements in this area that enable efficient phosphorylation of previously challenging substrates.

Another approach to overcome limitations in the phosphorylation of base-modified nucleosides is to start syntheses from natural sugars. Nucleotide **8c** was produced in near quantitative yield applying ribokinase and phosphoribosyl pyrophosphate (PRPP) synthetase for sugar activation followed by adenine phosphoribosyl transferase, adenylate kinase and creatine kinase (Scott et al., 2004). Others have followed a similar approach to produce 8-azaguanosine triphosphate and 5-fluorouridine triphosphate with high yields (Da Costa et al., 2007; Hennig et al., 2007). Although these enzymes show good activity toward base-modified substrates, they, however, are not applicable for cytosine derivatives (Scheit and Linke, 1982; Hennig et al., 2007) and show limited usefulness for the preparation of sugar modified NTPs (Esipov et al., 2016).

Cofactor recycling systems are commonly used in industrial biotransformation processes where costly cofactors like NAD/H would lead to uneconomic production routes (Weckbecker et al., 2010). As an additional benefit reaction equilibria are shifted to favor product formation through cofactor (re)generation and the reduction of product inhibition (Abu and Woodley, 2015). Although (d)ATP recycling systems have been applied for the synthesis of natural and modified nucleotides a comparison to reactions without these has never been shown (Hennig et al., 2007; Zou et al., 2013) and the benefit of these strategies has remained elusive. In this study, a strong beneficial effect was demonstrated for all tested compounds. NTP product yields increased by a factor of 4 (for **3c**) to 9 (for **6c**) using a PEP-based phosphate donor regeneration system. This observation provides clear evidence for the usefulness of employing ATP recycling systems for NTP synthesis since half of the nucleosides in our study were quantitatively converted to the respective triphosphate.

It remains to be investigated to which extent the applied regeneration enzyme takes part in the last cascade step converting NDPs and NTPs. Pyruvate kinase, acetate kinase and polyphosphate kinase possess wide substrate spectra (Ishige et al., 2001; Gao et al., 2008; Zou et al., 2013) and may be envisioned to serve a twofold purpose. In addition to regenerating the phosphate donor, PK might replace the NDPK in some reactions. For example, the parallel use of a pyruvate kinase for NTP formation and phosphate donor regeneration was demonstrated by Hennig and coworkers in the cascade synthesis of 5F-UTP and -CTP (Hennig et al., 2007). Furthermore, in order to achieve economic (d)ATP recycling in the established cascade system future studies could focus on optimizing the ratio of phosphate donor to substrate by applying substoichiometric (d)ATP concentrations as has been successfully shown before for selected single and multi-enzyme syntheses (Hennig et al., 2007; Zou et al., 2013).

CONCLUSION

We established a modular enzymatic cascade for the synthesis of NTPs from nucleosides in one pot. With (d)ATP as the phosphate donor all desired NTPs were accessed in low yield. The application of a (d)ATP regeneration system resulted in a shift of the reaction equilibrium toward the desired product NTP. Our approach, for the first time, allows for the efficient production of both natural and modified NTPs in high conversion with a standardized protocol after identifying a suitable enzyme combination. Thus, the present study lays the foundation for future high-yielding biocatalytic syntheses of NTPs from nucleosides.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

MF performed experiments and data analysis and wrote the manuscript. FK revised and illustrated the manuscript together with AW. KH and JS assisted with experiments. PN and AW coordinated the project and were involved in study design, data analysis, and revision and submission of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Personnel and material resources were funded by the "Exist Transfer of Research" program of the Federal Ministry for Economic Affairs and Energy (Germany). KH was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – grant number 392246628.

ACKNOWLEDGMENTS

We thank Prof. Munch-Petersen for supplying *Drosophila melanogaster* deoxynucleoside kinase plasmid (*Dm*dNK). We thank the Open Access Publishing funds of TU Berlin for the support of this publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00854/full#supplementary-material

REFERENCES

Abu, R., and Woodley, J. M. (2015). Application of enzyme coupling reactions to shift thermodynamically limited biocatalytic reactions. *ChemCatChem* 7, 3094–3105. doi: 10.1002/cctc.201500603

Andexer, J. N., and Richter, M. (2015). Emerging enzymes for ATP regeneration in biocatalytic processes. *ChemBioChem* 16, 380–386. doi: 10.1002/cbic.201402550

Baughn, R. L., Adalsteinsson, O., and Whitesides, G. M. (1978). Large-scale enzyme-catalyzed synthesis of ATP from adenosine and acetyl phosphate. Regeneration of ATP from AMP. *J. Am. Chem. Soc.* 100, 304–306. doi: 10.1021/ja00469a063

Berger, L. (1956). Crystallization of the sodium salt of adenosine triphosphate. *Biochim. Biophys. Acta* 20, 23–26. doi: 10.1016/0006-3002(56)90257-90258

Burgess, K., and Cook, D. (2000). Syntheses of nucleoside triphosphates. *Chem. Rev.* 100, 2047–2059. doi: 10.1021/cr990045m

Da Costa, C. P., Fedor, M. J., and Scott, L. G. (2007). 8-Azaguanine reporter of purine ionization states in structured RNAs. *J. Am. Chem. Soc.* 129, 3426–3432. doi: 10.1021/ja067699e

Deville-Bonne, D., El Amri, C., Meyer, P., Chen, Y., Agrofoglio, L. A., and Janin, J. (2010). Human and viral nucleoside/nucleotide kinases involved in antiviral drug activation: structural and catalytic properties. *Antiviral Res.* 86, 101–120. doi: 10.1016/j.antiviral.2010.02.001

Ding, Y., Ou, L., and Ding, Q. (2019). Enzymatic synthesis of nucleoside triphosphates and deoxynucleoside triphosphates by surface-displayed kinases. *Appl. Biochem. Biotechnol.* 190, 1271–1288. doi: 10.1007/s12010-019-03138-3133

Endo, T., and Koizumi, S. (2001). Microbial conversion with cofactor regeneration using genetically engineered bacteria. *Adv. Synth. Catal.* 343, 521–526. doi: 10.1002/1615-4169(200108)343:6/7<521::aid-adsc521>3.0.co;2-5

Esipov, R. S., Abramchik, Y. A., Fateev, I. V., Konstantinova, I. D., Kostromina, M. A., Muravyova, T. I., et al. (2016). A cascade of thermophilic enzymes as an approach to the synthesis of modified nucleotides. *Acta Nat.* 8, 82–90. doi: 10.32607/20758251-2016-8-4-82-90

Gao, S., Bao, J., Gu, X., Xin, X., Chen, C., and Ryu, D. D. Y. (2008). Substrate promiscuity of pyruvate kinase on various deoxynucleoside diphosphates for synthesis of deoxynucleoside triphosphates. *Enzyme Microb. Technol.* 43, 455–459. doi: 10.1016/j.enzmictec.2008.06.004

Giller, G., Tasara, T., Angerer, B., Mühlegger, K., Amacker, M., and Winter, H. (2003). Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. I. Chemical synthesis of various reporter group-labeled 2'-deoxyribonucleoside-5'-triphosphates. *Nucleic Acids Res.* 31, 2630–2635. doi: 10.1093/nar/gkg370

Gramlich, P. M. E., Wirges, C. T., Manetto, A., and Carell, T. (2008). Postsynthetic DNA modification through the copper-catalyzed azide-alkyne cycloaddition reaction. *Angew. Chemie Int. Ed.* 47, 8350–8358. doi: 10.1002/anie.200802077

Haynie, S., and Whitesides, G. (1990). Preparation of a mixture of nucleoside triphosphates suitable for use in synthesis of phosphate sugars from ribonucleic acid using nuclease P1, a mixture of nucleoside monophosphokinases and acetate kinase. *Appl. Biochem. Biotechnol.* 23, 205–220. doi: 10.1007/bf02942055

Hazra, S., Ort, S., Konrad, M., and Lavie, A. (2010). Structural and kinetic characterization of human deoxycytidine kinase variants able to phosphorylate 5-substituted deoxycytidine and thymidine analogues. *Biochemistry* 49, 6784–6790. doi: 10.1021/bi100839e

Hennig, M., Scott, L. G., Sperling, E., Bermel, W., and Williamson, J. R. (2007). Synthesis of 5-fluoropyrimidine nucleotides as sensitive NMR probes of RNA structure. *J. Am. Chem. Soc.* 129, 14911–14921. doi: 10.1021/ja073825i

Ishige, K., Hamamoto, T., Shiba, T., and Noguchi, T. (2001). Novel method for enzymatic synthesis of CMP-NeuAc. *Biosci. Biotechnol. Biochem.* 65, 1736–1740. doi: 10.1271/bbb.65.1736

Kamel, S., Yehia, H., Neubauer, P., and Wagner, A. (2019). "Enzymatic synthesis of nucleoside analogues by nucleoside phosphorylases," in *Enzymatic and Chemical Synthesis of Nucleic Acid Derivatives*, ed. J. Fernandez-Lucas (Hoboken, NJ: Wiley-VCH), 1–28. doi: 10.1002/9783527812103.ch1

Kaspar, F., Giessmann, R. T., Hellendahl, K. F., Neubauer, P., Wagner, A., and Gimpel, M. (2020). General principles for yield optimization of nucleoside phosphorylase-catalyzed transglycosylations. *ChemBioChem.* 21, 1428–1432. doi: 10.1002/cbic.201900740

Ko, W. C., Rolain, J. M., Lee, N. Y., Chen, P. L., Huang, C. T., Lee, P. I., et al. (2020). Arguments in favour of remdesivir for treating SARS-CoV-2 infections. *Int. J. Antimicrob. Agents.* 55:105933. doi: 10.1016/j.ijantimicag.2020.105933

Lee, T. T., and Momparler, R. L. (1976). Enzymatic Synthesis of 5-Azacytidine 5'-Triphosphate from 5-Azacytidine. *Anal. Biochem.* 71, 60–67. doi: 10.1016/0003-2697(76)90011-7

Liou, J. Y., Dutschman, G. E., Lam, W., Jiang, Z., and Cheng, Y. C. (2002). Characterization of human UMP/CMP kinase and its phosphorylation of D- and L-form deoxycytidine analogue monophosphates. *Cancer Res.* 62, 1624–1631.

Munch-Petersen, B., Knecht, W., Lenz, C., Søndergaard, L., and Piškur, J. (2000). Functional expression of a multisubstrate deoxyribonucleoside kinase from Drosophila melanogaster and its C-terminal deletion mutants. *J. Biol. Chem.* 275, 6673–6679. doi: 10.1074/jbc.275.9.6673

Ni, S., Yao, H., Wang, L., Lu, J., Jiang, F., Lu, A., et al. (2017). Chemical modifications of nucleic acid aptamers for therapeutic purposes. *Int. J. Mol. Sci.* 18:1683. doi: 10.3390/ijms18081683

Pasti, C., Gallois-Montbrun, S., Munier-Lehmann, H., Veron, M., Gilles, A. M., and Deville-Bonne, D. (2003). Reaction of human UMP-CMP kinase with natural and analog substrates. *Eur. J. Biochem.* 270, 1784–1790. doi: 10.1046/j.1432-1033.2003.03537.x

Pradere, U., Garnier-Amblard, E. C., Coats, S. J., Amblard, F., and Schinazi, R. F. (2014). Synthesis of nucleoside phosphate and phosphonate prodrugs. *Chem. Rev.* 114, 9154–9218. doi: 10.1021/cr5002035

Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. L. E. S. W., Zagursky, R. J., et al. (1987). A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 238, 336–341. doi: 10.1126/science.2443975

Ryll, T., and Wagner, R. (1991). Improved ion-pair high-performance liquid chromatographic method for the quantification of a wide variety of nucleotides and sugar-nucleotides in animal cells. *J. Chromatogr.* 570, 77–88. doi: 10.1016/0378-4347(91)80202-N

Scheit, K. H., and Linke, H. (1982). Substrate specificity of CTP synthetase from *Escherichia coli*. *Eur. J. Biochem.* 126, 57–60. doi: 10.1111/j.1432-1033.1982.tb06745.x

Scott, L. G., Geierstanger, B. H., Williamson, J. R., and Hennig, M. (2004). Enzymatic synthesis and 19 F NMR studies of 2-fluoroadenine-substituted RNA. *J. Am. Chem. Soc.* 126, 11776–11777. doi: 10.1021/ja047556x

Serra, I., Conti, S., Piškur, J., Clausen, A. R., Munch-Petersen, B., Terreni, M., et al. (2014). Immobilized Drosophila melanogaster deoxyribonucleoside kinase (DmdNK) as a high performing biocatalyst for the synthesis of purine arabinonucleotides. *Adv. Synth. Catal.* 356, 563–570. doi: 10.1002/adsc.201300649

Van Rompay, A. R., Johansson, M., and Karlsson, A. (2000). Phosphorylation of nucleosides and nucleoside analogs by mammalian nucleoside monophosphate kinases. *Pharmacol. Ther.* 87, 189–198. doi: 10.1016/S0163-7258(00)00048-6

Van Rompay, A. R., Johansson, M., and Karlsson, A. (2003). Substrate specificity and phosphorylation of antiviral and anticancer nucleoside analogues by human deoxyribonucleoside kinases and ribonucleoside kinases. *Pharmacol. Ther.* 100, 119–139. doi: 10.1016/j.pharmthera.2003.07.001

Weckbecker, A., Gröger, H., and Hummel, W. (2010). Regeneration of nicotinamide coenzymes: principles and applications for the synthesis of chiral compounds. *Biosyst. Eng.* 120, 195–242. doi: 10.1007/10_2009_55

Wu, W., Bergstrom, D. E., and Davisson, V. J. (2003). A combination chemical and enzymatic approach for the preparation of azole carboxamide nucleoside triphosphate. *J. Org. Chem.* 68, 3860–3865. doi: 10.1021/jo020745i

Yehia, H., Westarp, S., Röhrs, V., Kaspar, F., Giessmann, R. T., Klare, H. F. T., et al. (2020). Efficient biocatalytic synthesis of dihalogenated purine nucleoside analogues applying thermodynamic calculations. *Molecules* 25:934. doi: 10.3390/molecules25040934

Zaccolo, M., Williams, D. M., Brown, D. M., and Gherardi, E. (1996). An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* 255, 589–603. doi: 10.1006/jmbi.1996.0049

Zou, Z., Ding, Q., Ou, L., and Yan, B. (2013). Efficient production of deoxynucleoside-5'-monophosphates using deoxynucleoside kinase coupled with a GTP-regeneration system. *Appl. Microbiol. Biotechnol.* 97, 9389–9395. doi: 10.1007/s00253-013-5173-6

Conflict of Interest: BioNukleo GmbH is a spin-out of the Chair of Bioprocess Engineering at the TU Berlin with AW as CEO, FK and JS as employees, and MF and PN as board members.

The remaining author declares 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

CORRECTION
published: 26 October 2020
doi: 10.3389/fbioe.2020.606584

Corrigendum: Modular Enzymatic Cascade Synthesis of Nucleotides Using a (d)ATP Regeneration System

*Maryke Fehlau[1,2], Felix Kaspar[1,2], Katja F. Hellendahl[1], Julia Schollmeyer[1,2], Peter Neubauer[1] and Anke Wagner[1,2]**

[1] Chair of Bioprocess Engineering, Institute of Biotechnology, Faculty III Process Sciences, Technische Universität Berlin, Berlin, Germany, [2] BioNukleo GmbH, Berlin, Germany

Keywords: enzymatic cascade synthesis, nucleoside-5′-triphosphate, one-pot multi-enzyme reaction, nucleotide analog, nucleotide kinase, nucleoside kinase, modular, ATP regeneration system

A Corrigendum on

Modular Enzymatic Cascade Synthesis of Nucleotides Using a (d)ATP Regeneration System
by Fehlau, M., Kaspar, F., Hellendahl, K. F., Schollmeyer, J., Neubauer, P., and Wagner, A. (2020). Front. Bioeng. Biotechnol. 8:854. doi: 10.3389/fbioe.2020.00854

In the original article, there was an error. We did not receive the expression plasmid of *Dm*dNK from Daniela Ubiali (which is written 2x in the manuscript), but from Prof. Munch-Petersen.

A correction has been made to **Materials and Methods, General Information**, paragraph 3.

The corrected paragraph appears below:

"Wild-type nucleoside and nucleotide kinases were obtained from BioNukleo GmbH (Berlin, Germany) except for wide-spectrum deoxynucleoside kinase from *Drosophila melanogaster* (*Dm*dNK). The expression vector of *Dm*dNK was kindly provided by Prof. Birgitte Munch-Petersen (Roskilde University). According to the manufacturer the kinases possess the following substrate specificities: adenosine kinase (AK, NK14), guanylate kinase (GMPK, NMPK21) and adenylate kinase (AMPK, NMPK23) convert purine nucleoside/nucleotide substrates, while uridine monophosphate-cytidine monophosphate kinase (UMP-CMPK, NMPK22) and nucleoside diphosphate kinase (NDPK, NDPK32) accept both purine and pyrimidine nucleoside/nucleotide substrates. All enzymes obtained from BioNukleo were provided as stock solutions (0.1 to 1 mg/mL) and aliquots stored at −20°C until use. Pyruvate kinase (PK, P9136) was obtained from Sigma Aldrich as lyophilized powder, dissolved in 70 mM Tris–HCl pH 7.6 (1.74 mg/mL) and stored in aliquots at −20°C. All enzymes are active at 37°C and combinable in the same reaction buffer (70 mM Tris–HCl pH 7.6, 5 mM $MgCl_2$)."

A correction has been made to **Acknowledgments.**

The corrected paragraph appears below:

"We thank Prof. Munch-Petersen for supplying *Drosophila melanogaster* deoxynucleoside kinase plasmid (*Dm*dNK). We thank the Open Access Publishing funds of TU Berlin for the support of this publication."

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

OPEN ACCESS

Approved by:
Frontiers Editorial Office, Frontiers Media SA, Switzerland

***Correspondence:**
Anke Wagner
anke.wagner@tu-berlin.de

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

Received: *15 September 2020*
Accepted: *22 September 2020*
Published: *26 October 2020*

Citation:
Fehlau M, Kaspar F, Hellendahl KF, Schollmeyer J, Neubauer P and Wagner A (2020) Corrigendum: Modular Enzymatic Cascade Synthesis of Nucleotides Using a (d)ATP Regeneration System. Front. Bioeng. Biotechnol. 8:606584. doi: 10.3389/fbioe.2020.606584

frontiers in Bioengineering and Biotechnology

METHODS
published: 19 August 2020
doi: 10.3389/fbioe.2020.01000

Functional Analysis of Aquaporin Water Permeability Using an *Escherichia coli*-Based Cell-Free Protein Synthesis System

*Ke Yue, Jihong Jiang, Peng Zhang and Lei Kai**

The Key Laboratory of Biotechnology for Medicinal Plants of Jiangsu Province, School of Life Sciences, Jiangsu Normal University, Xuzhou, China

Aquaporins are essential water channel proteins found in all kingdoms of life. Although the water permeability of aquaporins has been well characterized, sample preparation for aquaporin water permeability assays remains challenging and time-consuming. Besides the difficulty in overexpressing membrane proteins in a cell-based expression system, the unique requirement for homogeneity in aquaporin proteoliposome sample preparations for water transport assays further increases the complexity. In this study, a complementary Cell-free Protein Synthesis (CFPS) method is described in detail, providing three different strategies for the preparation of aquaporin proteoliposome samples. Aquaporin can be produced either as a pellet fraction and then resolubilized, or co-translationally as a detergent-soluble fraction. Furthermore, aquaporin can be directly incorporated into liposomes, which was included in the CFPS reactions. Although proteoliposomes tend to fuse during the incubation of the CFPS reactions, an additional treatment of the fused samples with detergent, followed by a detergent removal step, can re-form homogenously sized proteoliposomes suitable for functional analysis. Using this method, we successfully characterized aquaporins from both prokaryotic and eukaryotic organisms. In particular, in the presence of liposomes, the developed CFPS expression system is a fast and convenient method for sample preparation for the functional analysis of aquaporins.

Keywords: aquaporin, cell-free protein synthesis, water permeability, detergent, proteoliposomes

OPEN ACCESS

Edited by:
Yuan Lu,
Tsinghua University, China

Reviewed by:
Giuseppe Calamita,
University of Bari Aldo Moro, Italy
Eric Beitz,
University of Kiel, Germany

***Correspondence:**
Lei Kai
lkai@jsnu.edu.cn

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

Received: *09 May 2020*
Accepted: *31 July 2020*
Published: *19 August 2020*

Citation:
Yue K, Jiang J, Zhang P and Kai L (2020) Functional Analysis of Aquaporin Water Permeability Using an Escherichia coli-Based Cell-Free Protein Synthesis System. Front. Bioeng. Biotechnol. 8:1000. doi: 10.3389/fbioe.2020.01000

INTRODUCTION

Aquaporins are integral membrane proteins that facilitate the rapid and selective movement of water and neutral low-molecular-mass solutes across biological membranes along osmotic gradients. Although the function and structure of aquaporins have been extensively characterized (Preston et al., 1992; Agre et al., 1993; Verkman and Mitra, 2000; Tornroth-Horsefield et al., 2006), sample preparation for the functional analysis of aquaporins is still challenging and time-consuming (Yue et al., 2019). Three main methods are commonly used to determine the water permeability of aquaporins: (i) The *Xenopus laevis* oocyte system. First, the cDNA encoding target aquaporins is injected and overexpressed in the native oocyte membrane. Water permeability is then calculated as the volume change of the oocytes under an osmotic shock, which is recorded *via* a light microscope (Preston et al., 1992; Yang and Verkman, 1997). (ii) The yeast protoplast

system (Pettersson et al., 2006). Here, the water permeability of protoplasts overexpressing target aquaporins is measured using stopped-flow spectrophotometry (Bertl and Kaldenhoff, 2007). (iii) The liposome system, where overexpressed target aquaporins are reconstituted into artificial liposomes, and the permeability of the resulting proteoliposomes is determined according to the changed intensity of scattered light at a fixed angle *via* stopped-flow spectrophotometry under osmotic shock (Ye and Verkman, 1989; Zeidel et al., 1992). The first two methods do not require isolation and purification of aquaporins; however, they suffer from the influence of endogenous integral membrane proteins (Ho et al., 2009). Although the liposome system offers a precise measurement of the permeability of specific aquaporins, the sample preparation process is still laborious and challenging, especially when aquaporins are overexpressed *in vivo* (Yue et al., 2019). In contrast, the Cell-free Protein Synthesis (CFPS) system, which is devoid of living cells and cell membrane barriers, has the unique advantage of being an open system (Henrich et al., 2015; Rues et al., 2016), allowing the introduction of various additives in a co-translational manner (Schwarz et al., 2007, 2008). In particular, hydrophobic reagents such as detergents and lipids can be introduced directly into the CFPS system to promote the correct folding of newly expressed membrane proteins (Junge et al., 2011; Roos et al., 2013).

During the last few years, we have developed a set of protocols for the production and functional characterization of aquaporins based on the CFPS system (Kai et al., 2010; Kai and Kaldenhoff, 2014; Yue et al., 2019). Here, we first detail a number of detergents suitable for soluble aquaporin expression. Secondly, we provide a detailed purification strategy, including detergent exchange on a column for the downstream reconstitution process. Furthermore, based on the work of Hovijitra et al. (2009), we describe a detailed protocol for the direct insertion of cell-free (CF) expressed aquaporins into liposomes using a modified procedure that is suitable for the continuous exchange cell-free (CECF) expression mode (Kai and Kaldenhoff, 2014; Yue et al., 2019). In this study, we have summarized our previous methods and provided a systematic protocol for the expression and functional characterization of aquaporins based on an *Escherichia coli* CFPS system, including template design, aquaporin expression and purification, proteoliposome preparation, and an aquaporin water permeability assay using stopped-flow spectrophotometry.

MATERIALS AND EQUIPMENT

Materials

Materials for the CFPS Reaction

1. 50 × Complete® Protease Inhibitor Cocktail tablets (Roche Diagnostics), 1 tablet/mL of Milli-Q water.
2. Amino acid mixtures containing 8 mM each of the 20 natural amino acids (weigh in all the compounds and dissolve the powders in Milli-Q water; the stock remains turbid).
3. RCWMDE mix containing 16.7 mM each amino acid (the stock remains turbid).
4. 1 M acetyl phosphate lithium potassium salt (AcP) (Sigma–Aldrich), adjusted to pH 7.0 with KOH.
5. 1 M phospho(enol)pyruvic acid monopotassium salt (PEP) (Sigma–Aldrich), adjusted to pH 7.0 with KOH.
6. An NTP mixture containing 90 mM ATP, 60 mM CTP, 60 mM GTP, and 60 mM UTP (Sigma–Aldrich), adjusted to pH 7.0 with NaOH.
7. Pyruvate kinase (Roche Diagnostics), 10 mg/mL.
8. RiboLock® RNAse Inhibitor (Fermentas), 40 U/μL.
9. Total *E. coli* tRNA (Roche Diagnostics), 40 mg/mL.
10. Folinic acid (as calcium salt), 10 mg/mL (Sigma–Aldrich).
11. Polyethylene glycol 8000 (PEG 8000) (Sigma–Aldrich), 40% (*w*/*v*).
12. 4 M potassium acetate (KOAc).
13. 2.4 M Hepes/20 mM EDTA, pH 8.0 adjusted with KOH.
14. 500 mM 1,4-dithiothreitol (DTT).
15. *Escherichia coli* S30 extract, stored frozen at −80°C (see section "Preparation of S30 Extract and T7RNAP").
16. T7 RNA polymerase (T7RNAP), stored frozen at −80°C (see section "Preparation of S30 Extract and T7RNAP").
17. Template DNA (plasmid DNA or linear PCR products), 200–500 ng/μL [see section "CFPS Reactions (Overnight)"].
18. Reaction container: analytical and preparative scale reaction container [see **Figure 1** and section "Protein Purification (3 h)"]; D-tube containers, 12–14 kDa MWCO (Merck Biosciences); Slide-A-Lyzer, 10 kDa MWCO (Pierce); dialysis membrane tubing, 12–14 kDa MWCO (Spectra/Por®2).
19. Plasmid and PCR product purification kits (Qiagen, Macherey and Nagel).

Materials for S30 Extract and T7RNAP Preparation

1. 40 × S30-A/B buffer: 400 mM Tris-acetate pH 8.2; 560 mM $Mg(OAc)_2$; 2.4 M KCl.

Supplement 1 × S30-A buffer with 6 mM β-mercaptoethanol. Supplement 1 × S30-B buffer with 1 mM DTT and 1 mM phenylmethylsulfonyl fluoride (PMSF).

2. 40 × S30-C buffer: 400 mM Tris-acetate pH 8.2; 560 mM $Mg(OAc)_2$; 2.4 M KOAc.

Supplement 1 × S30-C buffer with 0.5 mM DTT.

3. 2 × YTPG medium: 22 mM KH_2PO_4, 40 mM K2HPO4, 100 mM glucose, tryptone 16 g/L, yeast extract 10 g/L, NaCl 5 g/L.
4. LB medium: Peptone 10 g/L, yeast extract 5 g/L, NaCl 5 g/L.
5. Buffer T7RNAP-A: 30 mM Tris-HCl pH 8.0, 50 mM NaCl, 10 mM EDTA, 10 mM β-mercaptoethanol, 5% glycerol.
6. Buffer T7RNAP-B: 30 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM EDTA, 10 mM β-mercaptoethanol, 5% glycerol.
7. Buffer T7RNAP-C: 30 mM Tris-HCl pH 8.0, 1 M NaCl, 1 mM EDTA, 10 mM β-mercaptoethanol, 5% glycerol.
8. Buffer T7RNAP-D: 10 mM K_2HPO_4/KH_2PO_4 pH 8.0, 10 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 5% glycerol.
9. 20% streptomycin sulfate.

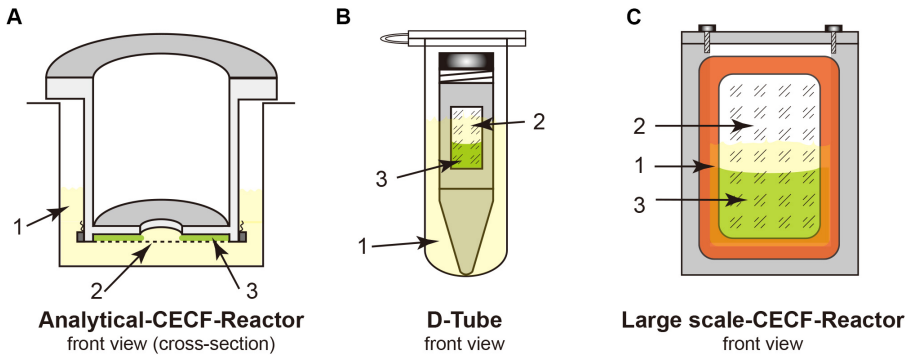


FIGURE 1 | Cell-free Protein Synthesis (CFPS) reactors for the continuous exchange cell-free (CECF) configuration. Panels **(A,C)** were custom made from Plexiglas. **(B)** Commercial D-Tube dialyzer either in a 2-mL Eppendorf tube or in 15–50-mL Falcon tubes, depending on the size of the D-Tube dialyzer. 1, light yellow indicates the outside feeding mixture (FM); 2, dialysis membrane; 3, light green indicates the reaction mixture (RM).

10. *Escherichia coli* strain for extract preparation: A19 (*E. coli* Genetic Stock Center, New Haven, CT, United States).
11. BL21(DE3) Star transformed with pAR1219 for T7RNAP preparation (Li et al., 1999).

Detergents and Lipids

1. Detergents: Triton X-100 (Merck Biosciences); Brij®35, Brij®58, Brij®S20, digitonin, and tyloxapol (Sigma; LMPG {1-myristoyl-2-hydroxy-sn-glycero-3-(phospho-rac-[1-glycerol])} and LPPG {1-palmitoyl-2-hydroxy-sn-glycero-3-(phospho-rac-[1-glycerol])} (Avanti Polar Lipids); Fos-Choline-12 (dodecylphosphocholine) and Fos- Choline-16 (n-hexadecylphosphocholine) (Affymetrix® Anatrace); and DDM (*N*-dodecyl-β-D-maltoside) and β-OG (octyl-β-D-glucopyranoside) (AppliChem).
2. Lipids: *E. coli* polar lipids (Avanti Polar Lipids) and L-α-phosphatidylcholine from soybean (Sigma–Aldrich).

Materials for Purification

1. Buffer P-B (binding buffer): 20 mM Tris-HCl pH 7.8, 150 mM NaCl, 20 mM imidazole, 0.05% DDM.
2. Buffer P-W (washing buffer): 20 mM Tris-HCl pH 7.8, 150 mM NaCl, 80 mM imidazole, 0.05% DDM.
3. Buffer P-E (elution buffer): 20 mM Tris-HCl pH 7.8, 150 mM NaCl, 300 mM imidazole, 0.05% DDM.

Materials for the Water Permeability Assay

1. 10 × Buffer R (reconstitution): 1 M MOPS buffer, pH 7.5.
2. Hyperosmotic solution: Buffer R + 400 mM sucrose.
3. Whatman polycarbonate membrane filter (200 nm) and filter supports (Florham Park, NJ, United States).
4. SM-2 beads (Bio-Rad, München, Germany).

Immunoblotting Assay

1. Penta·His Antibody, BSA-free (QIAGEN, Hilden, Germany).
2. Horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG (Sigma–Aldrich, Taufkirchen, Germany).
3. Polyvinylidene difluoride (PVDF) membrane (Merck Millipore, Darmstadt, Germany).

Equipment

1. Fermenter for 5–10 L of culture volume (B. Braun Biotech).
2. French press or other high-pressure, cell-disruption equipment.
3. Photometer.
4. Centrifuges and a set of rotors (Sorvall or Kontron).
5. Dialysis tubing, 30 kDa MWCO (Spectrum).
6. Mini extruder (Avanti Polar Lipids).
7. Ultrasonic water bath.
8. Thermo-shaker for incubation.
9. Chromatographic system (GE Healthcare).
10. Q-Sepharose column.
11. Ultrafiltration devices, 30 kDa MWCO.
12. Ultracentrifuge and suitable rotors (Beckman Coulter).
13. Rotary evaporator with vacuum control.
14. Stopped-flow spectrophotometer (SFM 300, Bio-Logic SAS, Claix, France).
15. ZetaPlus particle sizing software 2.27 for dynamic light scattering particle size analysis.
16. Gel electrophoresis system and blotting system.
17. Gel imaging system.

METHODS

Preparation of S30 Extract and T7RNAP

Preparation of S30 Extract (2.5 Days)

S30 is the key component of CFPS reactions, providing all the enzymes needed for translation and energy regeneration. A number of detailed protocols have been published elsewhere (Kai et al., 2012, 2015). Here, we briefly described the major steps:

1. Cell fermentation (using 2 × YTPG medium) at 37°C with vigorous aeration until the cells reach the mid-log growth phase, followed by fast chilling to 18°C.

2. Cell washing (using S30-A buffer) followed by cell lysis (in 110% S30-B buffer) using a French press (one passage above 1,000 psi is sufficient to disrupt the cells) or similar high-pressure cell homogenizer.
3. Centrifugation at 30,000 × *g* (S30) to clarify the lysate.
4. Runoff steps with high salt and heat (42°C) to remove endogenous mRNA and undesired proteins.
5. Overnight dialysis with one exchange against S30-C buffer and another S30 centrifugation step.
6. Aliquot the resulting supernatant, freeze with liquid nitrogen, and store at −80°C until use. Approximately 60 mL of S30 extract can be obtained from 10 L of fermented *E. coli* cells.

T7RNAP Preparation (1 Day)

Overexpression of T7RNAP is performed through transforming *E. coli* BL21 (DE3) Star cells with the pAR1219 plasmid (Li et al., 1999). One step of anion exchange chromatography followed by a concentration step is sufficient to obtain a highly functional enzyme. The resulting T7RNAP should show a prominent band around the 90 kDa region with Coomassie Blue staining following SDS–PAGE. The average yield is approximately 20,000–40,000 units per liter of culture.

1. Inoculate a fresh, overnight culture of newly transformed BL21(DE3) Star cells containing the pAR1219 plasmid into fresh LB medium at a 1:100 ratio. Let the cells grow at 37°C with shaking and induce with 1 mM IPTG when the OD600 reaches 0.6–0.8. Incubate for a further 5 h and harvest by centrifugation at 8,000 × *g* for 15 min at 4°C.
2. Resuspend the cell pellet in 30 mL of T7RNAP-A buffer and disrupt the cells by passaging once through a French press at 1,000 psi. Remove the cell debris by centrifugation at 20,000 × *g* for 30 min at 4°C. The ensuing steps should be performed at 4°C.
3. Adjust the supernatant to a final concentration of 4% streptomycin sulfate by the stepwise addition of a 20% stock solution with gently mixing to remove the released DNA. Incubate on ice for 5 min and centrifuge at 20,000 × *g* for 30 min at 4°C.
4. Perform anion exchange purification by loading the supernatant onto a 40-mL Q-Sepharose column equilibrated with T7RNAP-B buffer and wash the column extensively with T7RNAP-B buffer.
5. Elute the T7RNAP with a 50–500 mM NaCl gradient using T7RNAP-C buffer for 10 column volumes at a flow rate of 3–4 mL/min. Collect the fractions and analyze aliquots by SDS–PAGE (overexpression of T7RNAP should be indicated by the presence of a prominent band at approximately 90 kDa; however, significant impurities might still exist).
6. Pool the T7RNAP-containing fractions (combine only peak fractions) and dialyze against T7RNAP-D buffer overnight. Adjust to a final concentration of 10% glycerol and concentrate the resulting T7RNAP-containing fraction to a total protein concentration of 3–4 mg/mL by ultrafiltration (T7RNAP begins to precipitate at higher concentrations). Adjust to a final concentration of 50% glycerol, aliquot, and store at −80°C.

DNA Template Preparation (0.5–1 Day)

Clone the aquaporin gene of interest between the T7 promoter and T7 terminator. Common vectors such as pET (Merck Biosciences) or pIVEX (Roche Diagnostic) are frequently used. Templates for CFPS reactions should be prepared using standard commercial 'Midi' and 'Maxi' plasmid kits. Mini kits are usually not suitable due to the low quality of the obtained purified DNA. The resulting DNA should be dissolved in Milli-Q water and the stock concentration should be above 0.3 mg/mL.

Linear PCR templates obtained through a two-step overlap PCR and containing the T7 promoter and T7 terminator region can also be used. Linear PCR templates are less stable than plasmid templates; protective additives can be included, such as lambda gamS (Garamella et al., 2016).

CFPS Reactions (Overnight)

For membrane proteins, a CECF configuration should be employed for the CFPS reaction to obtain higher expression yields because of the relatively low yields of membrane proteins when compared with those of cytosolic proteins. Both analytical- and preparative-scale reactions can be performed using different dialysis setups. Although we use the custom-made reactors shown in **Figure 1**, commercial dialysis devices can also work, providing that the reaction mixture (RM) and the feeding mixture outside (FM) are in good contact, such as when using a D-tube dialyzer (Novagen) combined with an Eppendorf tube, as shown in **Figure 1**. The RM: FM ratio should be between 1:14 and 1:20.

1. Calculate the volume required for each compound to perform a set number of reactions.
2. Prepare a common master mixture (RFM) containing the compounds required for both the RM and the FM (see **Table 1**). Combine all the compounds into one reaction tube.
3. Aliquot the amount of RFM required for the RM and complete the RM and FM with the remaining compounds (see **Table 2**).
4. Optimize the Mg^{2+} and plasmid concentrations to obtain the maximum yield and use them for the ensuing assay. The range of Mg^{2+} screening is normally between 10 and 20 mM, while the range of plasmid concentration screening is between 0 and 50 μg/mL.
5. For D-CFPS (Cell-free Protein Synthesis in the presence of detergents) screening reactions, detergents should be included in both the RM and FM. For L-CFPS (Cell-free Protein Synthesis in the presence of liposomes) reactions, liposomes should only be included in the RM.
6. Transfer the RM and FM aliquots into reaction containers and incubate overnight at 30°C with continuous agitation, either by shaking or rolling, depending on the reaction container set up. This is necessary to ensure efficient substance exchange between the RM and FM through the

membrane. Shaking water baths or temperature-controlled cabinets with shaking plates (approximately 150–200 rpm) can be used.

Sample Preparation for Functional Analysis

Protein Resolubilization (Optional)

Aquaporins can be expressed without the addition of detergents or lipids. However, the expressed aquaporins are in pellet form and an additional step of resolubilization with detergents is required.

1. After overnight expression, collect the RM and transfer into a new tube (this should be done carefully to collect as much of the RM as possible because the precipitated proteins might stick to the membrane). Resuspend *via* pipetting up and down before aspiration.
2. Collect the pellet fraction *via* centrifugation at 20,000 × *g* for 30 min at 4°C.
3. Wash the pellet twice with Buffer-P-B without DDM and finally resuspend with a mild detergent in an equal volume of RM. Different detergents should be screened since they might result in different resolubilization efficiencies. Our previous results showed that Fos-Chlorine-12, Fos-Chlorine-16, and LPPG display relatively high resolubilization percentages for mAQP4 M23 (Kai et al., 2010).
4. Incubate the sample mixture at 37°C for 30 min with vigorous shaking.
5. Centrifuge at 20,000 × *g* for 30 min and collect the supernatant for the next purification step. Analyze the supernatant and pellet by SDS–PAGE and immunodetection to determine the resolubilization efficiency.

Protein Purification (3 H)

1. After overnight expression, collect the RMs from the D-CFPS reactions and transfer them into a new tube.
2. Centrifuge at 20,000 × *g* for 30 min at 4°C to collect the supernatant.
3. Combine the resulting supernatant or the supernatant resulting from the resolubilization step with nine volumes of Buffer P-B and mix with a proper amount of pre-equilibrated (with Buffer P-B) Ni-NTA resin slurry (in certain case Co^{+} was used instead of Ni^{+} for a better selectivity).
4. Incubate with gentle rolling at room temperature for 2 h and wash the resin with 10 volumes of Buffer P-W.
5. Elute with 3–5 volumes of Buffer P-E.
6. Analyze the results *via* SDS–PAGE or immunodetection.

Liposome Preparation (Overnight)

1. Solubilize the desired amount of lipid in chloroform or use presolubilized lipids in chloroform and transfer it into a round-bottom flask.
2. Form a thin lipid film *via* a vacuum rotary evaporator and leave the film under a complete vacuum overnight.
3. Reconstitute the dried lipid film in 1 mL of Buffer R to a final concentration of 20 mg/mL by vortexing for 15 min, forming multilamellar vesicles.
4. Extrude the resulting multilamellar vesicles 21 times through an Avanti Polar Lipids mini extruder holding a 200-nm Whatman polycarbonate membrane filter supported on both sides by two filter supports.

TABLE 1 | Common reagent mixture preparation (RFM) for CFPS with 1:16 ratio of RM: FM.

Compound	Stock	Final concentration	Volume[a] [μl]
RCWMDE	16.67 mM	1 mM	1,020
Amino acid mix	4 mM	0.5 mM	2,337.50
Acetyl phosphate	1 M	20 mM	340
Phospho(enol) pyruvic acid	1 M	20 mM	340
75 × NTP mix	90 mM ATP	1.2 mM	226.7
	60 mM G/C/UTP	0.8 mM	
1,4 dithiothreitol	500 mM	2 mM	68
Folinic acid	10 mg/ml	0.1 mg/ml	170
Complete® protease inhibitor	50×	1×	340
Hepes/EDTA buffer	24×	1 x	623.3
$Mg(OAc)_2$	1 M	11.1 16, mM[b]	274
KOAc	4 M	110, 270, mM[b]	382.5
PEG 8000	40%	2%	850
NaN_3	10%	0.05%	85
			Total: 7,057

[a]Volumes are calculated for 1 ml RM and 16 ml FM = 17 ml RFM. [b]Concentration is subject to optimization and might need to be adjusted for each individual target. Volumes are calculated for final total concentrations of Mg^{2+} of 16 mM, and K^{+} of 270 mM as additional amounts of 4.9 mM Mg^{2+} and 160 mM K^{+} are contributed by other compounds.

TABLE 2 | RM (1 ml) and FM (16 ml) preparation.

Compound	Stock	Final concentration	Volume
FM:			
RFM			6,642 μl
S30-C Puffer	1×	0.35×	5,600 μl
Amino acid mix	4 mM	0.5 mM	2,000 μl
MilliQ water			1,758 μl
			Total: 16 ml
RM:			
RFM			415 μl
Pyruvat kinase	10 mg/ml	0.04 mg/ml	4 μl
t-RNA (E. coli)	40 mg/ml	0.5 mg/ml	12.5 μl
T7RNAP	1.4 mg/ml	0.05 mg/ml	35,7 μl
Ribolock	40 U/μl	0.3 U/μl	7.5 μl
DNA template	0.2–0.5 mg/ml	0.015–0.03 mg/ml	60 μl
E. coli S30 extract	1×	0.35×	350 μl
MilliQ water			115.3 μl
			Total: 1 ml

5. The resulting unilamellar liposome solutions can be used for the next reconstitution step.
6. Caution: Do not freeze the liposomes; use freshly prepared liposomes. For short storage, store the preformed liposomes at 4°C but not for longer than 1 week as liposomes are not stable and tend to fuse during storage.

Aquaporin Reconstitution (Overnight)

1. Make up the following solution by the sequential addition of 10 × Buffer R (final dilution 1×), followed 10% Triton X-100 (a final concentration of 4 mM), 20 mg/mL preformed liposomes (a final concentration of 4 mg/mL), and finally purified aquaporin (a final concentration of approximately 100 μg/mL). A total volume of approximately 1.6 mL is recommended.
2. Incubate the reconstitution mixture at room temperature with gentle shaking for 30 min.
3. Remove the detergent using SM-2 beads according to the manual.
4. Centrifuge the reconstitution mixture at 500,000 × *g* for 45 min.
5. Wash the pellet again with Buffer R and finally resuspend with 1.6 mL of Buffer R.

Sample Preparation for L-CFPS (Overnight)

1. Collect the RMs from the L-CFPS reactions after overnight incubation and transfer into new Eppendorf tubes.
2. Centrifuge at 18,000 × *g* for 30 min and collect the pellet, which contains a mixture of multilamellar lipids with incorporated aquaporin, as well as precipitated aquaporin.
3. Add an equal volume of Buffer R with either 0.42% (*w*/*v*) Triton X-100 or 1% (*w*/*v*) β-OG per 4 mg/mL lipids to solubilize the incorporated aquaporin with lipids. The remaining pellets are precipitated aquaporin.
4. Centrifuge at 18,000 × *g* for 10 min, collect the supernatants, and remove the detergent using either SM-2 beads or dialysis.
5. Follow the instructions in the manual for the SM-2 beads to remove Triton X-100. For the removal of β-OG, perform step-wise dialysis against Buffer R with 0.5, 0.25, 0.125, and 0% β-OG. Each dialysis step should run for 6 h at 4°C with constant stirring, except for the final step which should run overnight.
6. The reformed proteoliposomes are extruded again through a 200-nm membrane filter before measurement by stopped-flow spectrophotometry.

Water Permeability Measurement Using a Stopped-Flow Spectrophotometer (1 H)

1. Determine the average diameter of the freshly prepared aquaporin-containing proteoliposomes by dynamic light scattering.
2. Measure the water permeability with a fixed light scattering angle (90 degrees) at 436 nm with temperature control.
3. When the system temperature is stable, the reconstituted proteoliposomes or reformed proteoliposomes from the L-CFPS reactions are mixed with an equal volume of hyperosmotic solution (Buffer R + 400 mM sucrose).
4. Data obtained from the spectrophotometer are fitted into an exponential rise equation. The initial shrinkage rate constant (k) is the average k vale of the best fitted exponential curve from 6–10 individual measurements.
5. Calculate the osmotic water permeability coefficients (P_f) of the corresponding samples using the following equation [1]:

$$P_f = \frac{k}{(S/V_0) \times V_w \times \Delta_{osm}} \quad (1)$$

where, S/V_0 is the vesicle's initial surface-to-volume ratio; V_w represents the partial molar volume of water (18 cm^3/mol); and Δ_{osm} is the osmotic driving force (200 mM if 400 mM sucrose is used for the hyperosmotic solution). The S/V_0 is calculated by determining the diameters of the proteoliposomes using dynamic light scattering.

6. (Optional) A reversible mercury inhibitory assay can be performed to show aquaporin-facilitated water transport. Incubate the reconstituted proteoliposomes with 300 μM $HgCl_2$ at the measuring temperature for 5 min. Then, follow steps 1–5 to obtain the P_f of the mercury-treated sample. The inhibition can be reversed by further treatment with 2 mM β-mercaptoethanol for 10 min.

ANTICIPATED RESULTS

Overexpression and Purification of Aquaporin From the CFPS System

The typical yield of aquaporins produced using the above protocol was approximately 1 mg/mL of RM. Routine optimization should be performed to achieve the maximum yield, including optimization of Mg^{2+} and plasmid concentrations. Aquaporin can be produced as a precipitate and resolubilized posttranslationally using mild detergents and remain functional (Schwarz et al., 2007). As shown in **Figures 2A,B**, 1% (*w*/*v*) Fos-Chlorine-12, 2% (*w*/*v*) Fos-Chlorine-16, and 1% (*w*/*v*) LPPG showed almost 100% resolubilization efficiency. One advantage of this expression mode was that the pellet fraction was mainly obtained from newly expressed membrane proteins, which led to higher purity without the requirement for extensive purification steps. However, an additional resolubilization step was required, and it remained unclear whether the resolubilized aquaporin was fully or only partially functional. The D-CFPS expression mode provides a direct hydrophobic environment co-translationally and avoids the resolubilization step. Several detergents, such as Brij®35 and Brij®58, showed high efficiency in solubilizing newly expressed aquaporins. For instance, in 0.2% Brij®35, 90% of the mAQP4 M23 was expressed in the soluble fraction (Kai et al., 2010) (**Figures 2C,D**), while

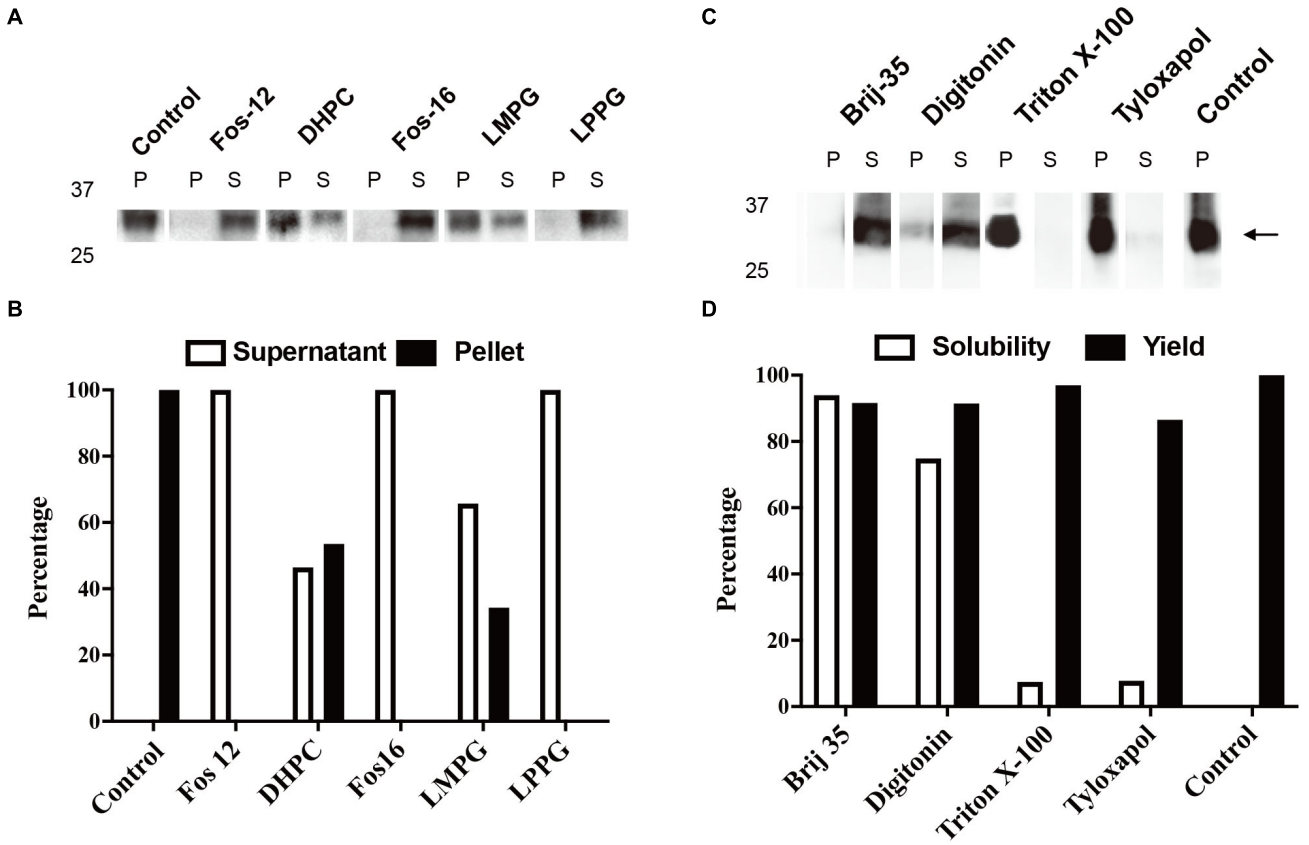


FIGURE 2 | Cell-free protein synthesis (CFPS) of mAQP4 M23. **(A,B)** Detergent resolubilization screening of mAQP4 M23; **(C,D)** detergent screening of mAQP4 M23 for the D-CFPS (CFPS in the presence of detergent) reaction. **(A,C)** Immunoblotting using anti-His antibodies. **(B,D)** Quantification of the corresponding immunoblotting by densitometry. The detergent concentrations (indicated by *w/v* percentage) were: 1% Fos-Chlorine-12; 2% DHPC; 2% Fos-Chlorine-16; 2% LMPG {1-myristoyl-2-hydroxy-sn-glycero-3-(phospho-rac-[1-glycerol])}; 1% LPPG {1-palmitoyl-2-hydroxy-sn-glycero-3-(phospho-rac-[1-glycerol])}; 0.2% Brij-35; 0.4% digitonin; 0.1% Triton X-100; and 0.05% tyloxapol. The figure was adapted and reused from Kai et al. (2010), PloS ONE, under the terms of the Creative Commons Attribution License.

1% Brij®S20 could support the soluble AqpZ expression (Yue et al., 2019). The optimum detergent might be target-dependent. Nevertheless, a general detergent selection criterion is that the detergent used should not decrease productivity, while still supporting a high solubility of the expressed protein in the CFPS environment.

Purification of resolubilized or D-CFPS-obtained aquaporin was performed through one affinity chromatography step and analyzed *via* SDS–PAGE. As indicated in the Methods section, during this step, a detergent exchange can be introduced because the detergents used in D-CFPS or resolubilization are often not suitable for the reconstitution step. First, dilution with the second detergent was performed to reduce the original detergent concentration. Then, a second washing step was carried out with a large volume of buffer containing the second detergent to further exchange and wash out the original detergent, leading to complete detergent exchange.

L-CFPS Reactions

As stated in Section "Sample Preparation for L-CFPS (Overnight)," a mixture of lipids and incorporated aquaporins can be directly isolated with medium-speed centrifugation. Homogenous proteoliposome can be reformed through detergent resolubilization and removal, which should greatly accelerate the speed of sample preparation for functional analysis. The liposome concentration must be above 8 mg/mL in the RM to support higher incorporation (Roos et al., 2013; Yue et al., 2019). However, due to the volume limit of the CFPS system, it may be impossible to achieve such a high lipid concentration in the final RM. The solution is to either increase the liposome stock concentration or the concentration of the amino acid mixture (25 mM for each of the 20 amino acids).

Water Permeability Measurements

As shown in **Figure 3**, functional aquaporin proteoliposomes should show a faster increase in the light scattering signal than that of the empty, control liposomes when mixed with a hyperosmotic solution. The corresponding water permeability can be calculated *via* equation [1] depicted in Section "Water Permeability Measurement Using a Stopped-Flow Spectrophotometer (1 H)." As an example,

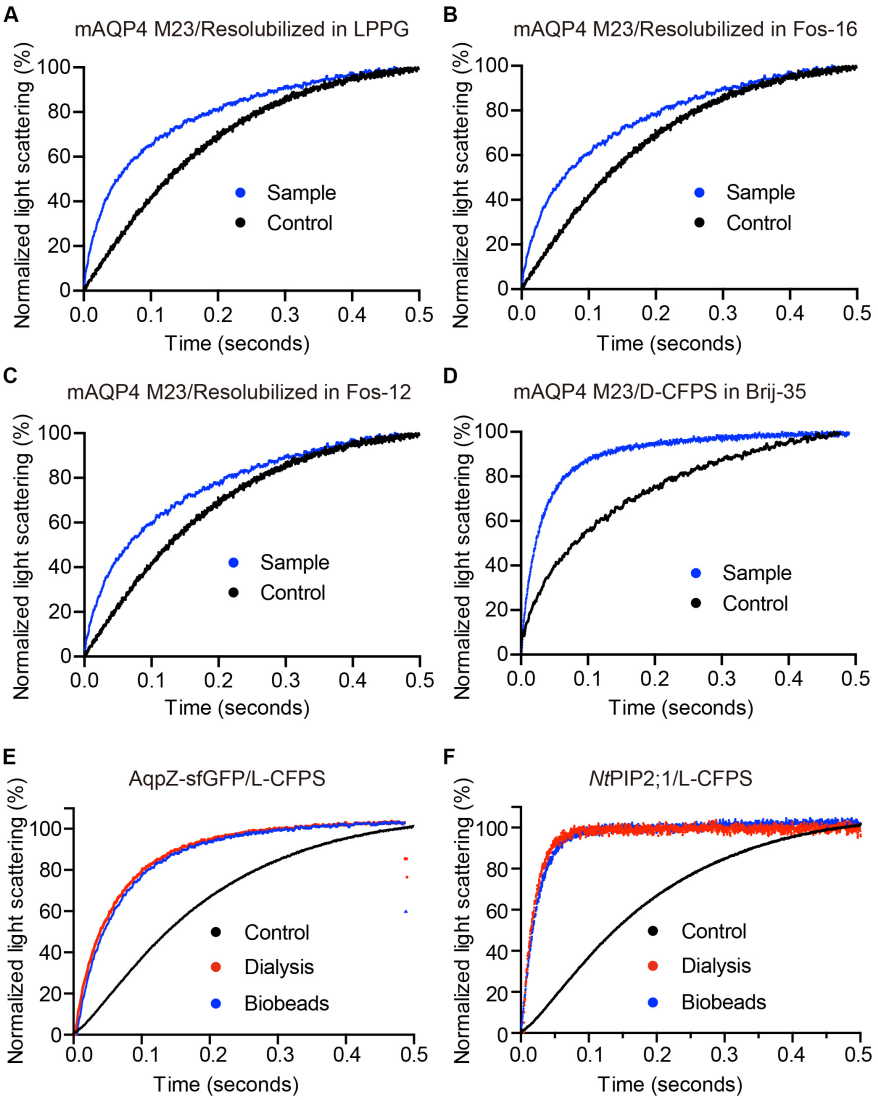


FIGURE 3 | The water permeability of aquaporin-containing proteoliposomes prepared from different Cell-free Protein Synthesis (CFPS) modes. The scattered light intensities (arbitrary units) from each measurement obtained *via* stopped-flow spectrophotometry were normalized between 0 and 100 (the smallest value was defined as 0, the largest value as 100). Each curve shown in this figure represents an averaged and normalized curve of 6–10 individual measurements. The water transport experiments were performed at 10°C. **(A–C)** mAQP4 M23 expressed as pellet fractions and resolubilized in different detergents; **(D)** mAQP4 M23 expressed using D-CFPS (CFPS in the presence of detergent); **(E,F)** *Escherichia coli* AqpZ-sfGFP and *Nt*PIP2;1 (from *Nicotiana tabacum*) expressed using L-CFPS (CFPS in the presence of liposomes) and treated with different detergent removal methods to reform the proteoliposomes. Panels **(A–D)** were adapted and reused from Kai et al. (2010), PloS ONE, under the terms of the Creative Commons Attribution License; **(E,F)** were adapted and reused from Yue et al. (2019), Cells, under the terms of the Creative Commons Attribution License.

the water permeabilities of the tested aquaporins (**Figure 3**) are summarized in **Table 3**. The binding of mercury ions to key cysteine residues inhibits the activities of most aquaporins (Savage and Stroud, 2007; Yukutake et al., 2008). Furthermore, a strong reducing reagent, such as β-mercaptoethanol, can be used to rescue this inhibition by competing with mercury for the binding to the key cysteine residues. Therefore, a reversible inhibitory assay can further confirm aquaporin-facilitated water transportation (Savage and Stroud, 2007; Yukutake et al., 2008; Kai et al., 2010). mAQP4 M23 was used as a model protein to test this reversible inhibitory assay (**Figure 4**). After treatment with mercury, the water permeability dropped from 160.01 ± 5.5 μm/s (k = 30.61 ± 1.05) to 94.9 ± 4.2 μm/s (k = 19.39 ± 0.81). However, the function of the blocked mAQP4 M23 could be rescued *via* β-mercaptoethanol treatment, resulting in a permeability of 137.5 ± 4.9 μm/s (k = 26.3 ± 0.94). The water permeability of empty, mercury-treated liposomes was set as control, displaying a water permeability of 50.3 ± 2.3 μm/s (k = 9.62 ± 0.45). No differences were detected between empty, non-treated liposomes and the mercury-treated samples (data not shown).

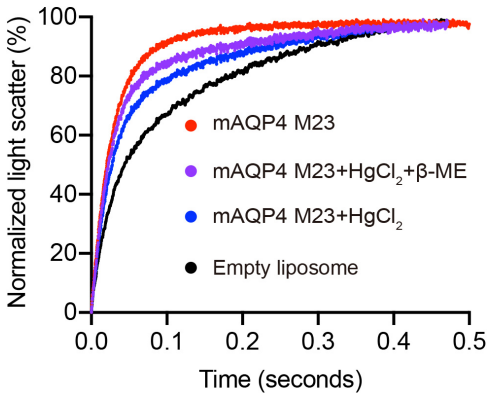


FIGURE 4 | Reverse inhibitory assay to confirm mAQP4 M23-mediated water transport. Proteoliposomes containing mAQP4 M23 obtained by D-CFPS (Cell-free Protein Synthesis in the presence of detergent) were treated with 300 μM $HgCl_2$ for 5 min at 23°C. For the recovery assay, 2 mM beta-mercaptoethanol was added, followed by incubation for 10 min. Empty liposomes with and without $HgCl_2$ treatment were used as controls.

The system temperature should be kept constant between individual measurements.

DISCUSSION

Although aquaporins can be successfully overexpressed in different host organisms such as *E. coli*, yeast, *X. laevis* oocytes, and insect cells, aquaporin sample preparation for functional and structural analysis remains a great challenge. An alternative recombinant protein production method, namely, the CFPS system, shows advantages over cellular expression systems with regard to membrane proteins (Rues et al., 2018). The CFPS system, which is devoid of living host cells, provides a flexible co-translation environment for the production of aquaporin, without the concern of maintaining living host cells. In addition, processes such as aquaporin extraction from the membrane of host cells or refolding of inclusion bodies can be completely avoided. Instead, the CFPS system allows the screening of translational conditions to achieve better folding. This is normally done through the introduction of additives, such as detergents or lipids, directly into the CFPS reaction mixture.

In this protocol, we have provided a systematic strategy for aquaporin sample preparation. This strategy is suitable for water permeability measurements and other applications that require functional aquaporin either in detergent micelles or liposomes, i.e., water filtration membranes (Wang et al., 2015). Three different expression modes were introduced: (i) CFPS without the addition of hydrophobic reagents; (ii) the D-CFPS mode; and (iii) the L-CFPS mode. The first expression mode is suitable for the first round of yield optimization, such as optimization of Mg^{2+} and plasmid concentrations. As indicated in the Results section, the membrane protein pellet was often purer than that obtained by D-CFPS. Although the pellet fractions could not always be functionally resolubilized, the D-CFPS mode provided co-translational solubilization of aquaporin, offering a direct hydrophobic environment for the folding of the target proteins. The D-CFPS mode was designed to obtain large quantities of aquaporins. The L-CFPS setup is the most efficient for the fast characterization of the functionality of specific aquaporins based on the proteoliposome assay. First, no additional purification step is needed. Second, the liposome lipid biolayer is a better mimic of the biological membrane than detergent micelles for aquaporin folding. However, after overnight expression, the resulting proteoliposomes were fused and formed a multilamellar lipid clot. Even at low speeds, centrifugation could still pellet all the lipid fractions (Yue et al., 2019). Therefore, it was not possible to directly use these samples for water permeability assays, as homogenously sized proteoliposomes are required for this process. Although it has been reported that low centrifugation can still provide proteoliposome suspensions that can be re-extruded through a filter membrane to reform homogenously sized samples (Hovijitra et al., 2009), this was not possible in our hands,

TABLE 3 | Summary of calculated water permeability from samples shown in **Figure 3**.

Sample[a]	Diameter (nm)	k[b]	Pf (μm/s)b	Significant differences[c]	References
Control (A–D)	113	4.87 ± 0.31	25.5 ± 1.6	a	Kai et al., 2010
Sample A	113	9.71 ± 0.52	50.7 ± 2.7	b	Kai et al., 2010
Sample B	113	9.45 ± 0.38	49.4 ± 2.0	b	Kai et al., 2010
Sample C	113	10. 41 ± 0.48	54.4 ± 2.5	b	Kai et al., 2010
Sample D	113	25.44 ± 1.07	133.1 ± 5.6	c	Kai et al., 2010
Control (E and F)	180	5.38 ± 0.16	44.88 ± 1.33	a	Yue et al., 2019
AqpZ-sGFP-Dialysis	210	15.49 ± 0.96	150.60 ± 9.33	b	Yue et al., 2019
AqpZ-sGFP-Biobeads	210	15.08 ± 0.62	146.61 ± 6.03	b	Yue et al., 2019
NtPIP2;1-Dialysis	195	54.23 ± 0.99	489.58 ± 8.94	c	Yue et al., 2019
NtPIP2;1-Biobeads	195	39.49 ± 0.59	356.51 ± 5.33	d	Yue et al., 2019

*[a]Control (A–D): empty liposomes shown in **Figures 3A–D**; Sample A–D: different mAQP4 M23 proteo-liposome samples shown in **Figures 3A–D**; Control (E and F): empty liposomes samples shown in **Figures 3E,F**. [b]Both the rate constant and calculated P_f value of a corresponding sample were given as mean ± SD. [c]Significant differences among different samples (indicated by different letters) were analyzed by one-way ANOVA followed by Tukey's honest significant difference test (n = 6–10, $P < 0.05$).*

which may have been due to several reasons. For instance, high Mg^{2+} and PEG concentrations may have promoted liposome fusion, or the relatively long reaction times and high yields may have resulted in additional protein precipitation before their incorporation into liposomes. Consequently, a large amount of non-incorporated aquaporin was present in the pellet fraction derived from the L-CFPS reaction. The key step involving mild detergent resolubilization provided a solution for the separation of the precipitated protein and the incorporated protein–lipid mixture. Homogenous aquaporin proteoliposome samples can be reformed following a detergent removal step. Both SM-2 beads and dialysis can then be used to remove the detergent. However, this may lead to small differences in the final water permeability of individual aquaporins, depending on which method is used. With the above-described methods, we successfully characterized mAQP4 M23 from the mouse (Kai et al., 2010), AqpZ from *E. coli*, *Nt*PIP2;1 from *Nicotiana tabacum* (Kai and Kaldenhoff, 2014; Yue et al., 2019), human AQP3 (Muller-Lucks et al., 2013), *Pf* AQP from *Plasmodium falciparum* (Muller-Lucks et al., 2012), and AqpB from *Dictyostelium discoideum* (von Bulow et al., 2012). In particular, the mercury inhibitory assay showed that reconstituted mAQP4 M23 was inhibited by mercury, which was different from AQP4 expressed in cell membranes. In line with the results of a previous study (Yukutake et al., 2008) it is reasonable to hypothesize that the reconstitution conditions followed in our study leads to bidirectional insertion of mAQP4 M23, exposing the critical cysteine (C178, loop D) to the outside of the liposome, and therefore accessible to the mercury ion (Kai et al., 2010). In addition, through the same sample preparation procedure, *Nt*AQP1 produced through D-CFPS could be applied to a CO_2 permeability assay using a modified black lipid membrane system (Kai and Kaldenhoff, 2014). Furthermore, this protocol can be easily adapted and modified to determine other aquaporin channel activities if the functional assay, such as that for glycerol, is based on proteoliposomes (Hovijitra et al., 2009). Because the CFPS system introduced here was in a reduced environment, the system must be adjusted through the introduction of a redox buffer, as well as DsbC, to allow the formation of correct disulfide bridges in specific aquaporins (Hachez et al., 2013; Matsuda et al., 2013; Rues et al., 2018; Dopp and Reuel, 2020).

CONCLUSION

We described a CFPS system-based sample preparation approach for the functional analysis of aquaporins. Three expression modes were applied that were suitable for different aquaporins. In particular, the L-CFPS mode approach is a fast, efficient, and convenient method for the functional characterization of aquaporins. Finally, this protocol can also be adapted for the preparation of samples for use in plenary or vesicular lipid-based functional assays involving other channel proteins or transporters.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

KY and LK performed the experiments. All authors contributed to the conceptualization and writing and editing of this manuscript.

FUNDING

KY received funding from the China Postdoctoral Science Foundation-funded project (Grant No. 2019M661776). LK is thankful for the support of the Natural Science Research of Jiangsu Higher Education Institutions of China (Grant No. 17KJB180003), the Natural Science Foundation of Jiangsu Normal University (Grant No. 17XLR037), and the Jiangsu Specially-Appointed Professors Program.

ACKNOWLEDGMENTS

The authors thank the support of Dr. Frank Bernhard from Goethe-Universität Frankfurt, Prof. Dr. Ralf Kaldenhoff from Technische Universität Darmstadt, and Prof. Dr. Zhinan Xu from Zhejiang University in developing this method.

REFERENCES

Agre, P., Sasaki, S., and Chrispeels, M. J. (1993). Aquaporins: a family of water channel proteins. *Am. J. Physiol. Renal Physiol.* 265(3 Pt 2):F461.

Bertl, A., and Kaldenhoff, R. (2007). Function of a separate NH3-pore in Aquaporin TIP2;2 from wheat. *FEBS Lett.* 581, 5413–5417. doi: 10.1016/j.febslet.2007.10.034

Dopp, J. L., and Reuel, N. F. (2020). Simple, functional, inexpensive cell extract for *in vitro* prototyping of proteins with disulfide bonds. *bioRxiv [Preprint]* doi: 10.1101/2019.12.19.883413

Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016). The All *E. coli* TX-TL Toolbox 2.0: a Platform for Cell-Free Synthetic Biology. *ACS Synth. Biol.* 5, 344–355. doi: 10.1021/acssynbio.5b00296

Hachez, C., Besserer, A., Chevalier, A. S., and Chaumont, F. (2013). Insights into plant plasma membrane aquaporin trafficking. *Trends Plant Sci.* 18, 344–352. doi: 10.1016/j.tplants.2012.12.003

Henrich, E., Hein, C., Dotsch, V., and Bernhard, F. (2015). Membrane protein production in *Escherichia coli* cell-free lysates. *FEBS Lett.* 589, 1713–1722. doi: 10.1016/j.febslet.2015.04.045

Ho, J. D., Yeh, R., Sandstrom, A., Chorny, I., Harries, W. E. C., Robbins, R. A., et al. (2009). Crystal structure of human aquaporin 4 at 1.8 angstrom and its mechanism of conductance. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7437–7442. doi: 10.1073/pnas.0902725106

Hovijitra, N. T., Wuu, J. J., Peaker, B., and Swartz, J. R. (2009). Cell-free synthesis of functional aquaporin Z in synthetic liposomes. *Biotechnol. Bioeng.* 104, 40–49. doi: 10.1002/bit.22385

Junge, F., Haberstock, S., Roos, C., Stefer, S., Proverbio, D., Dotsch, V., et al. (2011). Advances in cell-free protein synthesis for the functional and structural analysis of membrane proteins. *N. Biotechnol.* 28, 262–271. doi: 10.1016/j.nbt.2010.07.002

Kai, L., and Kaldenhoff, R. (2014). A refined model of water and CO2 membrane diffusion: effects and contribution of sterols and proteins. *Sci. Rep.* 4:6665. doi: 10.1038/srep06665

Kai, L., Kaldenhoff, R., Lian, J., Zhu, X., Dotsch, V., Bernhard, F., et al. (2010). Preparative scale production of functional mouse aquaporin 4 using different cell-free expression modes. *PLoS One* 5:e12972. doi: 10.1371/journal.pone.0012972

Kai, L., Orban, E., Henrich, E., Proverbio, D., Dotsch, V., and Bernhard, F. (2015). Co-translational stabilization of insoluble proteins in cell-free expression systems. *Methods Mol. Biol.* 1258, 125–143. doi: 10.1007/978-1-4939-2205-5_7

Kai, L., Roos, C., Haberstock, S., Proverbio, D., Ma, Y., Junge, F., et al. (2012). Systems for the cell-free synthesis of proteins. *Methods Mol. Biol.* 800, 201–225. doi: 10.1007/978-1-61779-349-3_14

Li, Y., Wang, E. D., and Wang, Y. L. (1999). A modified procedure for fast purification of T7 RNA polymerase. *Protein Expr. Purif.* 16, 355–358. doi: 10.1006/prep.1999.1083

Matsuda, T., Watanabe, S., and Kigawa, T. (2013). Cell-free synthesis system suitable for disulfide-containing proteins. *Biochem. Biophys. Res. Commun.* 431, 296–301. doi: 10.1016/j.bbrc.2012.12.107

Muller-Lucks, A., Bock, S., Wu, B. H., and Beitz, E. (2012). Fluorescent in situ folding control for rapid optimization of cell-free membrane protein synthesis. *PLoS One* 7:e42186. doi: 10.1371/journal.pone.0042186

Muller-Lucks, A., Gena, P., Frascaria, D., Altamura, N., Svelto, M., Beitz, E., et al. (2013). Preparative scale production and functional reconstitution of a human aquaglyceroporin (AQP3) using a cell free expression system. *N. Biotechnol.* 30, 545–551. doi: 10.1016/j.nbt.2013.03.007

Pettersson, N., Hagstrom, J., Bill, R. M., and Hohmann, S. (2006). Expression of heterologous aquaporins for functional analysis in Saccharomyces cerevisiae. *Curr. Genet.* 50, 247–255. doi: 10.1007/s00294-006-0092-z

Preston, G. M., Carroll, T. P., Guggino, W. B., and Agre, P. (1992). Appearance of Water Channels in Xenopus Oocytes Expressing Red-Cell Chip28 Protein. *Science* 256, 385–387. doi: 10.1126/science.256.5055.385

Roos, C., Kai, L., Proverbio, D., Ghoshdastider, U., Filipek, S., Dotsch, V., et al. (2013). Co-translational association of cell-free expressed membrane proteins with supplied lipid bilayers. *Mol. Membr. Biol.* 30, 75–89. doi: 10.3109/09687688.2012.693212

Rues, R. B., Dong, F., Dotsch, V., and Bernhard, F. (2018). Systematic optimization of cell-free synthesized human endothelin B receptor folding. *Methods* 147, 73–83. doi: 10.1016/j.ymeth.2018.01.012

Rues, R. B., Henrich, E., Boland, C., Caffrey, M., and Bernhard, F. (2016). Cell-Free production of membrane proteins in *Escherichia coli* lysates for functional and structural studies. *Methods Mol. Biol.* 1432, 1–21. doi: 10.1007/978-1-4939-3637-3_1

Savage, D. F., and Stroud, R. M. (2007). Structural basis of aquaporin inhibition by mercury. *J. Mol. Biol.* 368, 607–617. doi: 10.1016/j.jmb.2007.02.070

Schwarz, D., Dotsch, V., and Bernhard, F. (2008). Production of membrane proteins using cell-free expression systems. *Proteomics* 8, 3933–3946. doi: 10.1002/pmic.200800171

Schwarz, D., Junge, F., Durst, F., Frolich, N., Schneider, B., Reckel, S., et al. (2007). Preparative scale expression of membrane proteins in *Escherichia coli*-based continuous exchange cell-free systems. *Nat. Protoc.* 2, 2945–2957. doi: 10.1038/nprot.2007.426

Tornroth-Horsefield, S., Wang, Y., Hedfalk, K., Johanson, U., Karlsson, M., Tajkhorshid, E., et al. (2006). Structural mechanism of plant aquaporin gating. *Nature* 439, 688–694. doi: 10.1038/nature04316

Verkman, A. S., and Mitra, A. K. (2000). Structure and function of aquaporin water channels. *Am. J. Physiol. Renal Physiol.* 278, F13–F28.

von Bulow, J., Muller-Lucks, A., Kai, L., Bernhard, F., and Beitz, E. (2012). Functional characterization of a novel aquaporin from Dictyostelium discoideum amoebae implies a unique gating mechanism. *J. Biol. Chem.* 287, 7487–7494. doi: 10.1074/jbc.M111.329102

Wang, M., Wang, Z., Wang, X., Wang, S., Ding, W., and Gao, C. (2015). Layer-by-layer assembly of aquaporin Z-incorporated biomimetic membranes for water purification. *Environ. Sci. Technol.* 49, 3761–3768. doi: 10.1021/es5056337

Yang, B. X., and Verkman, A. S. (1997). Water and glycerol permeabilities of aquaporins 1-5 and MIP determined quantitatively by expression of epitope-tagged constructs in *Xenopus oocytes*. *J. Biol. Chem.* 272, 16140–16146. doi: 10.1074/jbc.272.26.16140

Ye, R. G., and Verkman, A. S. (1989). Simultaneous optical measurement of osmotic and diffusional water permeability in cells and liposomes. *Biochemistry* 28, 824–829. doi: 10.1021/bi00428a062

Yue, K., Trung, T. N., Zhu, Y. Y., Kaldenhoff, R., and Kai, L. (2019). Co-translational insertion of aquaporins into liposome for functional analysis via an *E. coli* based cell-free protein synthesis system. *Cells* 8:1325. doi: 10.3390/cells8111325

Yukutake, Y., Tsuji, S., Hirano, Y., Adachi, T., Takahashi, T., Fujihara, K., et al. (2008). Mercury chloride decreases the water permeability of aquaporin-4-reconstituted proteoliposomes. *Biol. Cell* 100, 355–363. doi: 10.1042/Bc20070132

Zeidel, M. L., Ambudkar, S. V., Smith, B. L., and Agre, P. (1992). Reconstitution of functional water channels in liposomes containing purified red-cell Chip28 protein. *Biochemistry* 31, 7436–7440. doi: 10.1021/bi00148a002

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

ORIGINAL RESEARCH
published: 19 August 2020
doi: 10.3389/fbioe.2020.00941

The Genetic Code Kit: An Open-Source Cell-Free Platform for Biochemical and Biotechnology Education

Layne C. Williams[1,2†], Nicole E. Gregorio[1,2†], Byungcheol So[1,2], Wesley Y. Kao[1,2], Alan L. Kiste[1], Pratish A. Patel[3], Katharine R. Watts[1,2] and Javin P. Oza[1,2*]*

[1] Department of Chemistry & Biochemistry, California Polytechnic State University, San Luis Obispo, CA, United States, [2] Center for Applications in Biotechnology, California Polytechnic State University, San Luis Obispo, CA, United States, [3] Department of Finance, Orfalea College of Business, California Polytechnic State University, San Luis Obispo, CA, United States

OPEN ACCESS

Edited by:
Yong-Chan Kwon,
Louisiana State University,
United States

Reviewed by:
Dong-Myung Kim,
Chungnam National University,
South Korea
Vijai Singh,
Indrashil University, India

***Correspondence:**
Katharine R. Watts
krwatts@calpoly.edu
Javin P. Oza
joza@calpoly.edu

[†] 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: 02 June 2020
Accepted: 21 July 2020
Published: 19 August 2020

Citation:
Williams LC, Gregorio NE, So B, Kao WY, Kiste AL, Patel PA, Watts KR and Oza JP (2020) The Genetic Code Kit: An Open-Source Cell-Free Platform for Biochemical and Biotechnology Education. Front. Bioeng. Biotechnol. 8:941. doi: 10.3389/fbioe.2020.00941

Teaching the processes of transcription and translation is challenging due to the intangibility of these concepts and a lack of instructional, laboratory-based, active learning modules. Harnessing the genetic code *in vitro* with cell-free protein synthesis (CFPS) provides an open platform that allows for the direct manipulation of reaction conditions and biological machinery to enable inquiry-based learning. Here, we report our efforts to transform the research-based CFPS biotechnology into a hands-on module called the "Genetic Code Kit" for implementation into teaching laboratories. The Genetic Code Kit includes all reagents necessary for CFPS, as well as a laboratory manual, student worksheet, and augmented reality activity. This module allows students to actively explore transcription and translation while gaining exposure to an emerging research technology. In our testing of this module, undergraduate students who used the Genetic Code Kit in a teaching laboratory showed significant score increases on transcription and translation questions in a post-lab questionnaire compared with students who did not participate in the activity. Students also demonstrated an increase in self-reported confidence in laboratory methods and comfort with CFPS, indicating that this module helps prepare students for careers in laboratory research. Importantly, the Genetic Code Kit can accommodate a variety of learning objectives beyond transcription and translation and enables hypothesis-driven science. This opens the possibility of developing Course-Based Undergraduate Research Experiences (CUREs) based on the Genetic Code Kit, as well as supporting next-generation science standards in 8–12th grade science courses.

Keywords: biochemical education, learn by doing, cell-free protein synthesis (CFPS), *in vitro* transcription and translation, synthetic biology (synbio), central dogma of molecular biology (CDMB), chemical education and teaching, augmented reality (AR)

Abbreviations: CFPS, cell-free protein synthesis; CUREs, course-based undergraduate research experiences; sfGFP, superfolder green fluorescent protein.

INTRODUCTION

Transcription and translation are fundamental cellular processes typically taught in high school and undergraduate science courses and utilized extensively in research settings. As such, students are expected to have an intimate grasp of these concepts to support both their academic and career goals. However, there is evidence that misconceptions about transcription and translation often persist for students even after they have completed these courses (Wright et al., 2014; Newman et al., 2016; Queloz et al., 2017). This issue likely stems from the intangibility of the microscopic processes of the "central dogma" when taught through lecture alone. In the absence of active learning modules, students are unable to visualize and represent these processes for further learning (Kozma et al., 2000; Duncan and Reiser, 2007). To address these limitations and allow students to interact with the individual steps of transcription and translation in the classroom, a variety of model-, analogy-, and virtual- based simulations have been developed (Pigage, 1991; Rotbain et al., 2008; Altiparmak and Nakiboglu Tezer, 2009; Debruyn, 2012; Takemura and Kurabayashi, 2014; Marshall, 2017; Dorrell and Lineback, 2019; Ibarra-Herrera et al., 2019; Chang et al., 2020) (**Figure 1**). Efforts to develop such activities represent educators' broad interest in providing students with active-learning modules to improve student learning outcomes. However, chemistry and biology curricula generally rely on laboratory practicals for active learning, as they help students connect scientific concepts and practices. Unfortunately, current wet-lab procedures for teaching transcription and translation are based on bacterial expression of fluorescent proteins, which precludes students from directly accessing and manipulating the genetic code machinery (Ward et al., 2000; Bassiri, 2011; Newman and Wright, 2013; Deutch, 2019) (**Figure 1**). While all these existing activities are generally low-cost and useful learning tools to help students understand the broad scope and details of transcription and translation, no single activity enables in-depth, hands-on, inquiry-based laboratory learning. The limitations of existing approaches underscore the need for an active learning laboratory-based module that allows students to interrogate transcription and translation in a learn-by-doing fashion.

Active learning has been demonstrated to increase student test scores and decrease the odds of failing classes in science, technology, engineering, and mathematics (STEM) (Nogaj, 2013; Freeman et al., 2014). In addition to these learning benefits, active learning is more engaging for students, ultimately promoting positive attitudes toward their education (Armbruster et al., 2009). Prior work also suggests that active learning may engage underrepresented students more than lecture-based courses, helping to narrow the achievement gap in STEM courses (Haak et al., 2011; Theobald et al., 2020). Curriculum design at our own university has led to the development of studio classrooms for general chemistry courses, which integrate the laboratory and lecture portions of the course into one space and time period. The studio classroom helps students to explicitly connect concepts taught in lecture through experimentation, resulting in improved exam scores, more expert-like learning attitudes, and positive assessments of the active learning environment from both students and instructors (Kiste et al., 2017). In order to apply these findings and address the lack of active learning opportunities for transcription and translation in our biochemistry curriculum, we sought to incorporate cell-free protein synthesis (CFPS) into our classroom laboratories. Toward this end, we developed the "Genetic Code Kit," a classroom-ready, modular CFPS kit that is amenable to broad dissemination. Importantly, we sought to determine whether implementing the Genetic Code Kit improves student performance on content-based assessments, as well as students' self-assessed comfort and confidence with experimental procedures.

Advancements in the CFPS platform over the last few decades have enabled a multitude of novel applications in biotechnology, including rapid prototyping for engineering biological systems and easy-to-use point of care diagnostics and biosensors (Pardee et al., 2016; Salehi et al., 2017; Benítez-Mateos et al., 2018; Bundy et al., 2018; Dopp and Reuel, 2018; Takahashi et al., 2018; Gräwe et al., 2019; Gregorio et al., 2019; Silverman et al., 2019; Jung et al., 2020). CFPS generally relies on a cell-extract containing the cellular machinery that supports transcription and translation *in vitro* and is supplemented with additional reagents that provide the necessary energy and precursors. The open nature of the CFPS system is one of the main advantages of the platform as it allows the user to produce proteins on-demand without relying on living cells. Thus, CFPS permits the user to directly manipulate the environment of protein synthesis to suit their needs without the limitation of cellular viability constraints, as is the case for *in vivo* protein expression. The unique advantages of CFPS are also what makes it well suited for active, inquiry-based learning in ways that can transform biochemical and biotechnology education, while simultaneously exposing students to experimental procedures associated with an emerging biotechnology. The pioneering work by BioBits and myTXTL have provided the proof-of-concept in adapting CFPS to classroom settings and engaging students at various grade levels (Huang et al., 2018; Stark et al., 2018, 2019; Collias et al., 2019). Additionally, CFPS remains robust in a variety of chemical environments (Yin and Swartz, 2004; Seki et al., 2008; Dopp et al., 2019; Gregorio et al., 2019; Karim et al., 2020) providing extensive flexibility in accommodating a broad range of learning objectives. These advantages make CFPS a next-generation educational technology to help meet the next-generation science standards. Moving beyond the proof-of-concept, we focus on using CFPS to teach the fundamental processes of transcription and translation and assess the extent and context of learning gains at the undergraduate level.

Transitioning the CFPS platform from a research-focused technology to one that is broadly accessible to high school and university classrooms required extensive simplification, reduced costs, and improved reagent stabilization. Our work to date has taken incremental steps toward these milestones by reducing the number of pipetting steps in CFPS setup (Levine et al., 2019a), creating a less-labor intensive cell extract preparation workflow (Levine et al., 2019b), and identifying a low-cost formulation of additives that enables storage and transport of cell-free extract at room temperature (Gregorio et al., 2020). These advances are part of a concerted effort by the research community to make CFPS

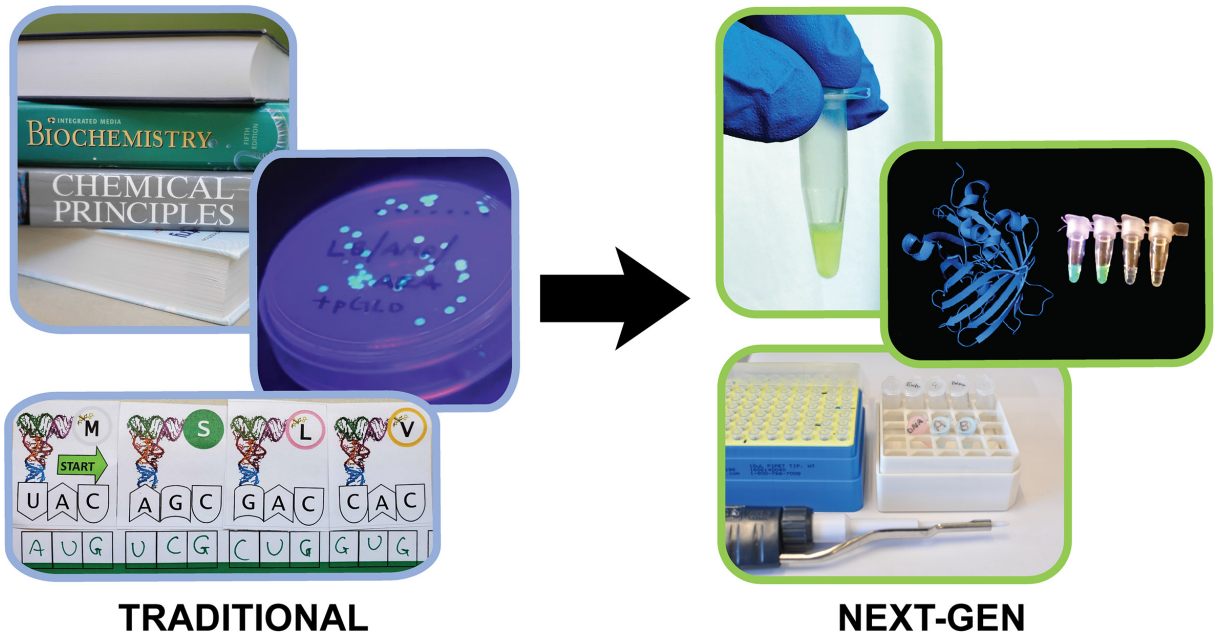


FIGURE 1 | Traditional central dogma teaching tools and the next-generation Genetic Code Kit. The Genetic Code Kit utilizes cell-free protein synthesis and augmented reality to teach the processes of transcription and translation.

accessible to classrooms around the world (Huang et al., 2018; Stark et al., 2018, 2019; Collias et al., 2019). As a result, instructors and institutions now have many options for obtaining CFPS resources for implementation in their classrooms. Each option has its respective advantages that allow instructors to support their learning objectives. Given these combined advancements in accessibility, CFPS is becoming even easier to broadly implement in the teaching laboratory, with the potential for supporting 100s to 1000s of students per quarter.

Here, we report the Genetic Code Kit, an implementation of CFPS used to teach transcription and translation. This kit is intended to be low-cost and open source to support accessibility and broad dissemination, especially to schools and programs with limited funding. To accommodate a variety of curricular limitations, the Genetic Code Kit can be completed within a single 3-h laboratory period, and does not require instructors to dedicate time in a subsequent day to collect data. The kit utilizes crude, *E. coli*-based extract and a DNA template encoding sfGFP, which together have been broadly demonstrated to support robust and reliable protein expression (Park et al., 2017; Levine et al., 2019a,b). Importantly, the sfGFP product resulting from a successful CFPS reaction is easy to visualize in real-time with minimal equipment or processing, and introduces students to a workhorse reporter broadly used in research and industry. The Genetic Code Kit contains 4 components: (1) a tube containing cell extract in which the reaction mixture is to be assembled, (2) the sfGFP DNA template, (3) "solution A" containing cofactors and substrates, and (4) "solution B" containing the energy system. The liquid transfer of just three reagents ranging from 4.2 to 11.4 μL allows students to gain micro-pipetting experience while reducing the likelihood of failed reactions. In our implementation, this setup proved reliable and forgiving, with all students able to obtain visible titers of sfGFP within 90 min. Requiring students to manually add all reagents necessary for transcription and translation is an important aspect of the Genetic Code Kit, as it provides the opportunity to identify and discuss the importance of each class of reagent (e.g., DNA template, energy reagents, and building blocks). This aspect of the kit also provides the flexibility to modify the kit based on the desired learning objectives, allowing for other inquiry-based learning opportunities, as well as CUREs.

We have also developed laboratory materials to accompany the Genetic Code Kit, which help students connect the microscopic processes taking place inside their CFPS reactions to the macroscopic outcome. This includes the laboratory manual and student worksheet (**Supplementary Data Sheets 1, 2**). Additionally, we created an augmented reality activity that allows students to interrogate the structure-function relationships of GFP to understand the basis for green fluorescence as a function of protein synthesis in their tubes (**Supplementary Data Sheet 3**). In addition to these specific pedagogical goals related to the central dogma, students also gain exposure to research techniques such as pipetting, reagent handling, the importance of negative and positive controls in experimental design, reaction setup, and data analysis. Importantly, we conducted a controlled study to investigate improvements in student understanding of transcription and translation and their self-assessed comfort with performing an emergent research technique as a function of their hands-on experience with the Genetic Code Kit. Our work demonstrates that implementing CFPS as a hands-on laboratory module leads to significant learning gains associated with transcription and translation learning objectives, as well

as positive self-assessment of comfort and confidence with research techniques.

MATERIALS AND METHODS

Cell Growth and Extract Preparation

Escherichia coli cell extract was generated using our previously reported CFAI workflow (Levine et al., 2019b). A culture was prepared by inoculating a loopful of BL21* DE3 cells into a 2 L baffled flask containing 1 L of Cell-free Autoinduction media (5.0 g of sodium chloride, 20.0 g of tryptone, 5.0 g of yeast, 14.0 g of potassium phosphate dibasic, 6.0 g of potassium phosphate monobasic, 6.0 mL of glycerol, 4.0 g of D-lactose, 0.5 g of D-glucose, and nanopure water to 1.0 L). The culture was incubated at 30°C and 200 rpm for approximately 15 h. Subsequently, the culture was centrifuged at 4°C and 5,000 *g* for 10 min. Harvested cells were resuspended in 30 mL of S30 buffer [10 mM Tris OAc, pH 8.2, 14 mM $Mg(OAc)_2$, 60 mM KOAc, 2 mM DTT] by vortexing, then spun down at 4°C and 5,000 *g* for 10 min. Supernatant was removed and cell pellets were flash frozen and stored at −80°C or used immediately for extract preparation. Cell pellets were resuspended in 1 mL of S30 buffer per 1 g of cells. 1.4 mL of resuspended cells were aliquoted into a 1.5 mL microfuge tube. The resuspension was sonicated using a Qsonica Q125 Sonicator with a 3.175 mm probe, with the cell resuspension surrounded by an ice water bath. Three pulses of 45 s on and 59 s off, at 50% amplitude were carried out. Immediately after sonication, 4.5 μL of 1.0 M DTT was spiked into the lysate and the tube was inverted several times to mix. Lysate was centrifuged at 4°C and 18,000 *g* for 10 min. The resulting supernatant is the cell extract. The mixture was flash frozen in liquid nitrogen and stored at −80°C until Genetic Code Kit preparation.

DNA Purification

DNA template pJL1-sfGFP was purified from DH5α cells using an Invitrogen PureLink HiPure Plasmid Maxiprep Kit. DNA was eluted using warm molecular biology grade water instead of the provided TE buffer for compatibility with the CFPS system. DNA plasmid was diluted with molecular biology grade water to a concentration of 42.1 ng/μL, such that no additional water was needed to prepare 30 μL CFPS reactions with a final DNA concentration of 16 ng/μL. DNA was stored at −20°C until Genetic Code Kit preparation.

Solution A and B Preparation

Solution A (containing cofactors and substrates) was prepared with the following specified concentrations of reagents: 8.14 mM ATP, 5.77 mM GTP, 5.77 mM UTP, 5.77 mM CTP, 153.8 mg/mL folinic acid, 771.9 mg/mL tRNA, 2.71 mM NAD, 1.81 mM CoA, 27.1 mM oxalic acid, 6.79 mM putrescine, 10.2 mM spermidine, 386.9 mM HEPES buffer. Solution B (containing the energy system) was prepared with the following specified concentrations of reagents: 71.6 mM magnesium glutamate, 71.6 mM ammonium glutamate, 930.8 mM potassium glutamate, 14.3 mM 20 amino acids, and 238.1 mM phosphoenolpyruvate. All reagents were dissolved in molecular biology grade water. Both solutions were stored at −80°C until Genetic Code Kit preparation, however, these solutions are also stable at −20°C for 3 months (**Supplementary Figure 1**).

Genetic Code Kit Preparation and Reaction Setup

Each kit contained the appropriate amount of pre-aliquoted reagents for the laboratory size and was stored at −20°C for up to 5 days until student use. Each pair of students was provided a strip of four PCR tubes, each containing 10 μL of extract. Each group of four students shared a set of PCR tubes containing molecular biology grade water, pJL1-sfGFP DNA plasmid, solution A, and solution B. Students added 11.4 μL of water, 4.4 μL of solution A, and 4.2 μL of solution B to two tubes as negative controls and 11.4 μL of DNA plasmid, 4.4 μL of solution A, and 4.2 μL of solution B to two tubes as positive controls. All reagents were kept on ice throughout reaction setup. The completed reactions were placed in a 37°C incubator and checked intermittently for green fluorescence. Necessary equipment includes a p20 pipette, pipette tips, an incubator, and a blue or black light. More details can be found in the laboratory manual (**Supplementary Data Sheet 1**).

Development of Lab Materials

The lab manual and worksheet (**Supplementary Data Sheets 1, 2**) for the Genetic Code Kit were developed with the following student learning objectives as a framework: (A) illustrate and describe the processes of transcription and translation; (B) identify the minimally necessary genetic components, enzymes, and reagents necessary for transcription and translation *in vitro*; (C) predict and visualize the outcomes of adding, or not adding, various components to CFPS reactions; (D) define CFPS and its advantages over *in vivo* protein synthesis; (E) paraphrase how energy metabolism sustains transcription and translation in a CFPS reaction. Background on CFPS, the processes of transcription and translation, including the necessary components for each of these processes, and the energy metabolism system operating in CFPS reactions was provided in the lab manual (**Supplementary Data Sheet 1**).

The student worksheet contained open-ended questions corresponding to each of the learning objectives; some questions also required students to draw a schematic to represent their understanding of a topic (**Supplementary Data Sheet 2**). For example, for learning objective B, students were asked to illustrate the templates for transcription and translation, including genetic elements like a promoter and ribosomal binding site, and their relative locations to one another on a DNA template. Students were asked to consider the outcome of the experiment if certain elements were missing, such as dNTPs or a particular amino acid, in order to address learning objective C. Questions related to learning objective E focused on steps that require energy input, and how the levels of high-energy molecules like ATP change throughout the CFPS reaction.

The student questionnaire contained 16 content-based questions and 12 attitudinal questions (**Supplementary Data Sheet 4**). All questions were multiple choice. The content-based

section contained three baseline questions that tested knowledge independent of the intervention's learning objectives and were not expected to be impacted by this laboratory exercise. They acted as a control for differences in baseline aptitudes between the pre- and post- questionnaires. Of the remaining 13 content-based questions, four questions tested transcription knowledge and nine tested translation knowledge. Transcription questions focused on key enzymes and required genetic elements on the DNA template for initiation and termination of transcription. Translation questions were focused on the basic mechanism of the ribosome, including how tRNA and mRNA interact, and the required genetic elements on the mRNA template for initiation and termination of translation. The 12 attitudinal questions asked students to rank their knowledge of transcription and translation vocabulary and comfort with research techniques.

The augmented reality activity utilized Augment[1], a smart phone application, to project the three-dimensional structure of sfGFP onto student benchtops for an exploration of protein structure, structure-function relationships, and the structural basis for fluorescence (**Supplementary Data Sheet 3**). However, our pre- and post- questionnaire did not assess student understanding of sfGFP structure or structure-function relationships, so the impacts of this activity on student learning cannot be reported here.

Implementation of the Genetic Code Kit and Data Collection

The Genetic Code Kit and relevant assessments were implemented in the laboratory component of our non-majors' "Survey of Biochemistry and Biotechnology" course (CHEM 313) taught by biochemistry faculty. The prerequisite for enrollment was the completion of an introductory organic chemistry course. Our curriculum allows students to select either Organic Chemistry I (CHEM 216), which is the first quarter of a year-long organic chemistry sequence or Survey of Organic Chemistry (CHEM 312), which is a one-quarter survey of organic chemistry (**Table 1**). The students involved in this study represent a breadth of educational backgrounds, with diverse majors from four colleges at Cal Poly SLO (**Table 1**). All student data was used with written consent of the participants in the study, based on Institutional Review Board (IRB) approval obtained prior to execution.

Implementation occurred over a 3-week period, with each lab section meeting once a week for 3 h. As a "pre-questionnaire" in week 1, all students completed the questionnaire described above (**Supplementary Data Sheet 4**). In week 2, students in the intervention group used the Genetic Code Kit in their regularly scheduled lab section (**Supplementary Data Sheets 1–3**). The control group did not meet and did not perform the experiment or augmented reality activity due to a holiday. However, they were provided with the lab manual and completed the same post-lab worksheet. In week 3, all students repeated the same questionnaire administered in week 1, representing the "post-questionnaire." A total of 69 students completed both pre- and post- laboratory questionnaires, with 15 in the control group and 54 in the intervention group.

Intervention group students performed the Genetic Code Kit lab module in a single 3 h lab period. They were provided the lab manual at least 3 days prior to performing the experiment. After a brief introduction to the experiment in the lab period, students were asked to follow the instructions for reaction setup described in the lab manual, commencing *in vitro* transcription and translation. Reaction tubes were then placed in a 37°C mini-incubator for 1–1.5 h (**Supplementary Figure 2**). During the incubation period, students completed the post-lab worksheet and augmented reality activity (**Supplementary Data Sheets 2, 3**) and listened to a short lecture from instructors on the basics of transcription and translation. This brief lecture reviewed information on transcription and translation that was also covered in the 4-h per week lecture portion of the course, and introduced the components of each of the solutions in the Genetic Code Kit that correspond to these processes. This information was also available to students in the control group in the form of the introduction in the lab manual, and in the course textbook. At the end of the incubation period, students visualized fluorescence with the naked eye, and enhanced visibility was achieved using a handheld blue light before the lab period was over.

Statistical Methods

Student responses were collected and all anonymized assessment scores and responses can be found in **Supplementary Table 1**. Content-based questions were divided into baseline (1, 5, 6) and transcription and translation (2–4, 7–16) categories based on each question topic. Each category was analyzed by comparing the pre- and post- questionnaire scores for the control and intervention groups, visualized via box and whisker plots generated in SigmaPlot. Paired *t*-tests were run for both groups using JMP, and *p*-values were recorded with a significance level of 0.05. These categories were also analyzed by calculating the normalized learning gain and effect size for both student groups to understand the magnitude of the effect of the Genetic Code Kit. Normalized gain enables the comparison of groups that start at different levels of performance, as it calculates the score increases with respect to the window of potential learning based on pre-questionnaire scores (Hake, 1998). Effect size provides an additional metric that accounts for the number of students tested and the variation in scores among the students (Cohen, 1988). Question-based normalized gain was calculated to determine student performance on each of the 16 questions individually. This metric uses the same equation as normalized gain, however, the average pre- and post- scores are replaced by the percentage of students who answered the question correctly on the pre- and post- questionnaires. Additionally, the content-based data was matched to student major and previous course completion data in the form of an Excel dashboard that allows the user to analyze trends that occur within these subgroups (**Supplementary Table 2**). The dashboard also allows for a statistical comparison of the control group relative to the intervention group. Due to the different sample sizes, the comparison was performed using the

[1]www.augment.com

TABLE 1 | Student population distributions by major and completed courses.

College	Major	Control	Intervention	Ochem I	Survey of Ochem
College of Science and Mathematics	Biological Sciences	6	16	17	7
	Kinesiology	0	1	0	1
	Marine Science	0	1	1	0
	Microbiology	1	3	4	0
College of Agriculture, Food, and Environmental Science	Animal Science	2	7	2	7
	Food Science	0	3	0	3
	Nutrition	2	14	2	13
	Wine and Viticulture	2	5	0	6
College of Engineering	Biomedical Engineering	1	2	3	0
	Materials Engineering*	0	1	1	1
College of Liberal Arts	Psychology	1	1	1	1
	Total students	15	54	31	39

**The materials engineering student took both Ochem I and Survey of Ochem.*

Fisher's Z Test. Point biserial analysis was performed using the Akindi software[2].

Attitudinal questions were analyzed by comparing the trends in the percentage of students that selected each answer choice on the pre- and post- questionnaires. For statistical analysis, student answers were converted to numerical values, where $A = 1$ and $E = 5$. Paired t-tests comparing pre- and post- scores for each question were run using JMP and p-values were recorded with a significance level of 0.05.

RESULTS

Content-Based Assessment of Student Learning

The content-based section of the questionnaire contained 16 questions (3 baseline, 13 transcription and translation). For baseline questions unrelated to the learning objectives, there was a minimal increase in the mean percentage of correct answers; the control group's mean score increased from 35.6 to 37.8% and the intervention group's mean score increased from 34.0 to 41.3% (**Figure 2A**). However, a two-sided paired t-test showed that neither of these increases were significant (p-value > 0.05). Thus, we concluded that neither group became significantly better at answering the post-questionnaire as a result of previous exposure in the pre-questionnaire. On transcription and translation questions for the control group, we observed minimal increases in the mean score, from 41.5 to 48.7%. Comparatively, the intervention group had a larger increase in the average score on transcription and translation questions, from 49.6 to 63.8% (**Figure 2B**). One-sided paired t-tests within the control and intervention groups comparing pre- and post- student scores indicated no significant increase (p-value > 0.5) for the control group and a significant increase (p-value < 0.001) for the intervention group. This indicates that completing the hands-on Genetic Code Kit experiment significantly improves students' ability to correctly answer questions regarding transcription and translation.

[2]www.akindi.com

In addition to observing improvements in average assessment scores, we also wanted to better understand the magnitude of the effect of the intervention on student learning gains. Toward this goal, we evaluated both normalized learning gains and effect sizes, since both are commonly used metrics in STEM education. The extent of normalized learning gains is categorized as low (gain < 0.3), medium ($0.7 >$ gain ≥ 0.3), and high (gain ≥ 0.7) (Hake, 1998). On baseline questions, the control and intervention groups demonstrated low gains of 0.03 and 0.11, respectively, as expected (**Figure 3A**). For the transcription and translation questions, the control group demonstrated a normalized gain of 0.12 while the intervention group demonstrated a gain of 0.28. Effect sizes were also calculated as an additional metric to understand the magnitude of learning gains, while accounting for the student sample size and variation. Effect sizes are categorized as small (effect = 0.2), medium (effect = 0.5), and large (effect = 0.8) (Cohen, 1988). For the baseline questions, we observed small effect sizes of 0.07 for the control group and 0.28 for the intervention group (**Figure 3B**). Effect sizes on the transcription and translation questions were 0.32 for the control (small-medium) and 0.60 for the intervention (medium-large). As with the normalized gain analysis, the intervention group's ability to correctly answer questions related to transcription and translation after using the Genetic Code Kit module was much greater than the control group, who did not carry out the activity.

Lastly, we analyzed the question-based normalized gains for each of the 16 questions individually (**Figure 3C**). This analysis was intended to indicate student performance on individual questions, allowing us to identify questions that were poorly designed or not well-addressed by the Genetic Code Kit. The outcome of question-based normalized gain assessment was the identification of questions 7 and 11 as particularly challenging for the intervention group. In fact, the control group outperformed the intervention group on those two questions, and the normalized gain for the intervention group was negative for question 11. Quantitatively, the point-biserial correlation coefficient values for questions 7 and 11 were above 0.2, suggesting that they are "fair" questions. Qualitatively, it is possible that these questions were written ineffectively, were mismatched with our learning objectives, or that CFPS

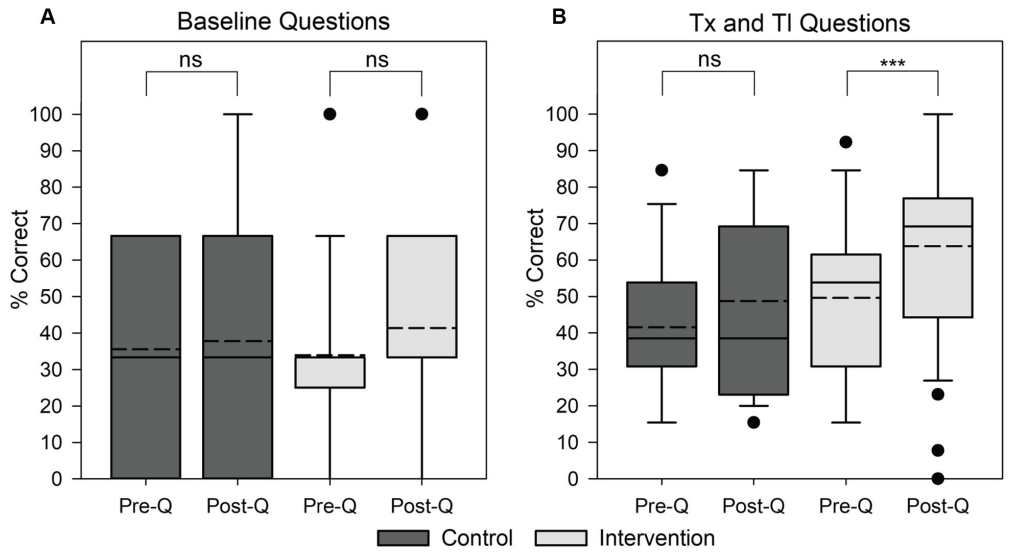


FIGURE 2 | Impact of the Genetic Code Kit on student performance on content-based questions involving baseline or transcription (Tx) and translation (Tl) questions. Student score distributions are depicted as follow: solid lines indicate median, dotted lines indicate mean, boxes demarcate the 25th and 75th percentiles, whiskers represent the 10th and 90th percentiles, and points represent outliers. Control group scores represent a population of 15 students and intervention group scores represent a population of 54 students. The content-based portion of the questionnaire contained 16 questions, 3 baseline and 13 transcription and translation. Specific questions can be found in **Supplementary Data Sheet 4**. **(A)** Student score distributions for baseline questions. Pre- and post- scores for the control group and intervention group were compared using a two-sided paired *t*-test (ns indicates *p*-value > 0.05) with a null hypothesis that pre- and post- scores will be equal. **(B)** Student score distributions for transcription and translation questions. Pre- and post- scores for the control group and intervention group were compared using a one-sided paired *t*-test (ns indicates *p*-value > 0.05, *** indicates *p*-value < 0.001) with a null hypothesis that pre- and post- scores will be equal.

was not able to resolve student misconceptions regarding the macromolecular interactions involved in translation. In fact, non-covalent interactions involved in translation were not explicitly covered in the pre-lab lecture, worksheet, or lab manual.

Given that we observed meaningful normalized learning gains and effect sizes upon intervention despite questions 7 and 11, we remained curious about the learning gains observed in the remaining questions. In a follow-up analysis (**Supplementary Tables 2**, **3**), we removed questions 7 and 11 from the group of transcription and translation questions and used this narrower scope to evaluate learning gains by student demographics. We observed that students who had previously taken Ochem I, the first quarter in a year-long series of organic chemistry, had significantly higher learning gains compared to the control group (p-value < 0.05), while those who had taken Survey of Ochem did not significantly benefit (p-value > 0.05) from the Genetic Code Kit intervention compared to the control group (**Supplementary Table 3**). The intervention group students that did not significantly benefit were mostly from the College of Agriculture, Food, and Environmental Science, who have historically underperformed in the Survey of Biochemistry and Biotechnology course. While this observation is only suggestive when we removed questions 7 and 11 from the analysis, it represents an intriguing starting point for using CFPS to consider preparation gaps and achievement gaps within our student populations. These results suggest that if question design can be improved and sample size can be increased, implementation of CFPS has the potential to explore the basis for preparation and achievement gaps in biochemical education. Regardless, these additional findings are contingent on solving the learning issues identified in questions 7 and 11, as these differences in prerequisite preparation only appear when they are removed from the analysis.

Overall, significant increases in the average scores on content-based questions (**Figure 2**), a normalized learning gain around 0.3, and an effect size of 0.6 for the intervention group (**Figure 3**) indicate that implementing the Genetic Code Kit improved students' ability to comprehend and answer questions relating to transcription and translation. As no significant increase (p-value > 0.05) in the performance on baseline questions was observed, we propose that the observed increase in assessment scores for transcription and translation questions was a result of the Genetic Code Kit rather than repeated exposure to the questionnaire.

Attitudinal-Based Assessment of Student Learning

The pre- and post- questionnaires completed by both the control and intervention groups contained a total of 12 attitudinal questions. These questions prompted students to self-assess their recognition and knowledge of transcription and translation vocabulary, as well as their comfort with laboratory techniques used in CFPS. Prior work has documented students' deficiency in metacognitive skills and found that active learning pedagogies can strengthen these skills (National Research Council, 2000). Our attitudinal-based questions allow us to examine how students' perceptions of their learning correlate with their results

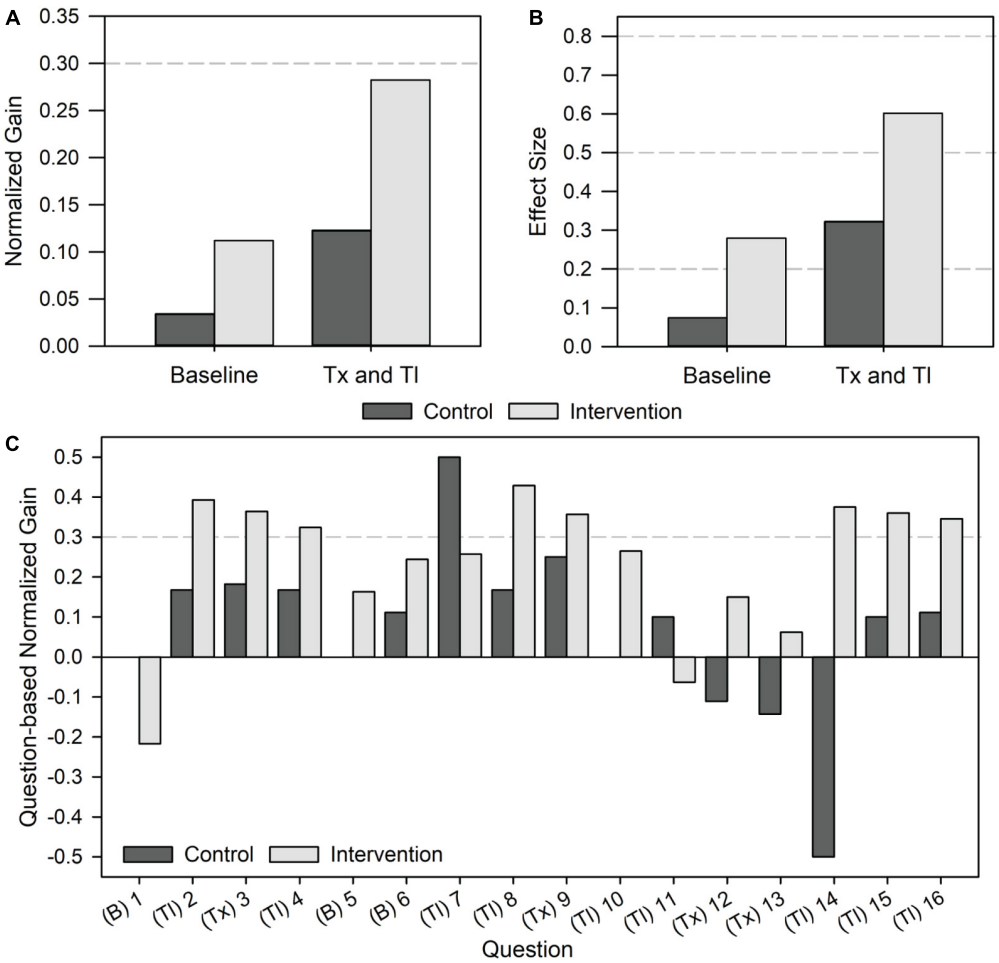


FIGURE 3 | Magnitude of student learning gains on content-based questions upon implementing the Genetic Code Kit. The control group represents a population of 15 students and the intervention group represents a population of 54 students. The content-based portion of the questionnaire contained 16 questions, 3 baseline and 13 transcription and translation. Specific questions can be found in **Supplementary Data Sheet 4**. **(A)** Normalized gain by question category. Normalized gains > 0.3 indicate a medium gain activity. **(B)** Effect size by question category. Effect sizes of 0.2 indicate small effects, 0.5 indicate medium effects, and 0.8 indicate large effects. **(C)** Question-based normalized gain for each question. Question categories are indicated as follows: (B) baseline, (Tx) transcription, and (TI) translation. Normalized gains > 0.3 indicate a medium gain activity.

on the content-based assessment (**Supplementary Figure 3**). We found that both the control and intervention groups showed positive correlations on pre- and post- questionnaires, with an increase in the slope from pre- to post- questionnaire. For the control group, the pre-questionnaire R^2-value was 0.02 and post-questionnaire was 0.36. For the intervention, the pre-questionnaire R^2-value was 0.10 and post-questionnaire was 0.30. The relatively low pre-questionnaire R^2 is noteworthy: it shows that students' knowledge and attitudes are, effectively, uncorrelated. The increase in post-questionnaire R^2 indicates that knowledge and attitudes move in the same direction. Overall, this analysis indicates that students' self-reported confidence correlated with their performance on content-based questions. As a result, we pursued more detailed analysis of the attitudinal-based questions.

We first considered the possibility that improvements in students' self-assessment of their confidence were an outcome of their recognition of vocabulary terms through repeated exposure to the questionnaire rather than as a result of improved conceptual understanding of the terms. To address this concern, we chose to perform detailed per-question analysis for the attitudinal-based assessment on questions that involved comfort with CFPS as an indicator of how beneficial the activity was in introducing a novel biotechnology. For the intervention group, we observed significant increases (p-value < 0.05) between pre- and post- scores for questions 23 and 25–27 using a one-sided paired t-test (**Figure 4**). When prompted with "I know what CFPS is" (question 23) on the pre-questionnaire, over 50% of the intervention group students indicated that they had no idea what the term meant and ∼11% indicated that they knew what the term meant (**Figure 4A**). After conducting the experiment, this changed to less than 5% and greater than 50%, respectively. The control group saw a similar, but less extensive shift in the trend with almost 40% of students reporting that they knew

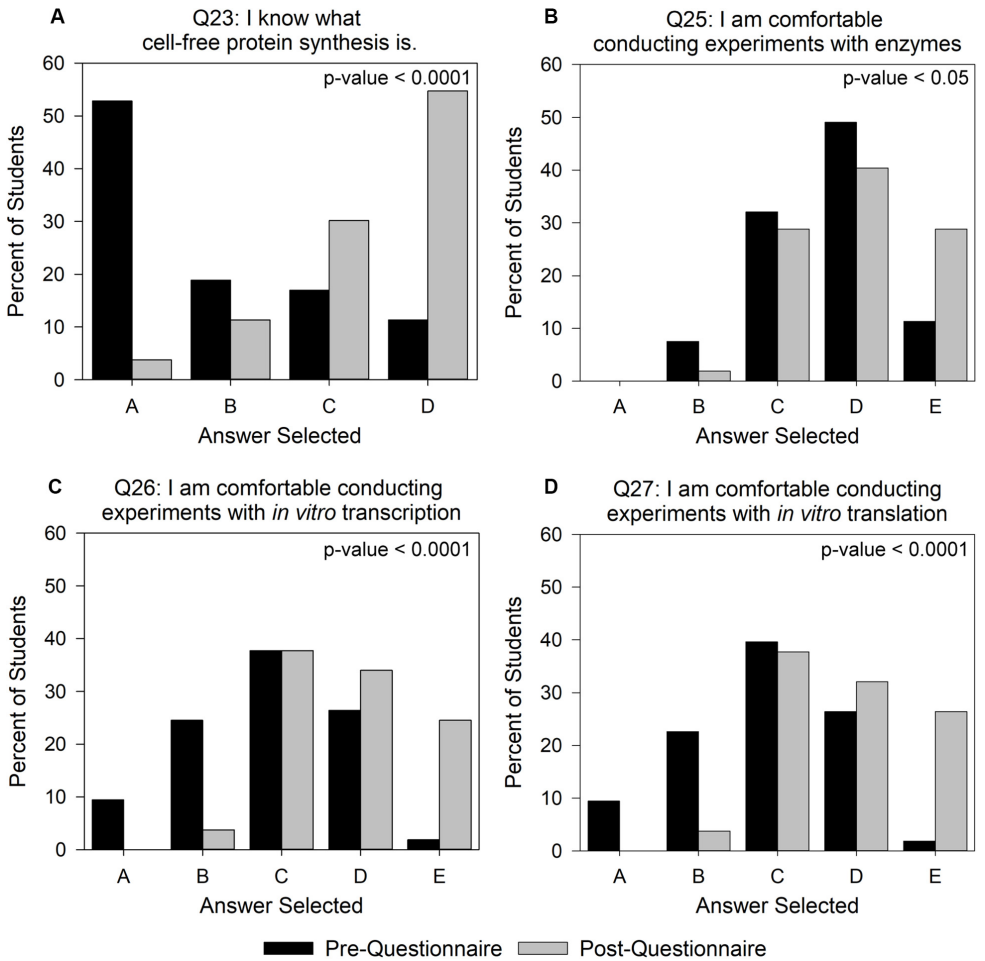


FIGURE 4 | Changes in intervention group student attitudes toward CFPS and conducting CFPS-based experiments. Answer choices for **(A)** ranged from A – "I have no idea what this term means" to D – "I know what this term means." Answer choices for **(B–D)** ranged from A – "Strongly disagree" to E – "Strongly agree." Student answers were converted to a numerical value where $A = 1$ and $E = 5$, in order to calculate p-values using a one-sided paired t-test with a null hypothesis that pre- and post- scores would be equal. The intervention group contained 52 students. This is less than the number of students in the content analysis, as some students did not complete the attitudinal section of the post-questionnaire. All possible answer categories can be found in **Supplementary Data Sheet 4**.

what the term meant in the post-questionnaire (**Supplementary Figure 4**). The comparable shift in the control and intervention groups is likely due to the background information that they received on CFPS through the lab manual alone. For question 25 (**Figure 4B**), "I am comfortable conducting experiments with enzymes," pre-questionnaire comfort was generally high for the intervention group, but only ∼11% of students indicated that they "strongly agreed." However, the Genetic Code Kit increased intervention student confidence in working with enzymes, such that almost 30% of students said they "strongly agreed" on the post-questionnaire. This was a noteworthy observation, since the Genetic Code Kit was implemented at the end of the quarter, and students had worked with enzymes in numerous previous laboratory modules. The intervention group's comfort with conducting experiments involving *in vitro* transcription and translation (questions 26 and 27; **Figures 4C,D**) also showed notable improvement, with the number of students answering "strongly agree" increasing to ∼25% from less than 2%. Comparatively, the control group had less than 8% of students say that they "strongly agreed" in response to questions 25–27 (**Supplementary Figure 4**). These data indicate that the intervention group's hands-on exposure to the CFPS reaction improved their comfort with these laboratory skills over the control group.

DISCUSSION

The CFPS platform has seen significant development and widespread use as a biotechnology tool in recent years. CFPS harnesses the genetic code in a test-tube, in a flexible and tunable biochemical milieu, making it poised to be a transformative educational technology. Specifically, CFPS allows students to probe the processes of transcription and translation in a way that improves their learning outcomes, while providing them the technical skills for careers in

biotechnology. Here, we report the implementation of our Genetic Code Kit, a simplified, yet modular CFPS reaction, in college-level biochemistry curriculum. Importantly, the Genetic Code Kit improved students' understanding of transcription and translation for undergraduate students in a survey of biochemistry course. Our results suggest that the tactile process of setting up a CFPS reaction by adding solutions containing the building blocks, energy system, and DNA template to *E. coli* extract, and observing the real-time production of a fluorescent protein increases students' comprehension of transcription and translation. Our observations are consistent with the extensive literature on the benefits of a physical experience in student learning (Bopegedera, 2011; Zacharia et al., 2012; Kontra et al., 2015; Kiste et al., 2016). Moreover, the Genetic Code Kit may help resolve common student misconceptions surrounding transcription and translation. For example, physically supplementing the CFPS reaction vessel with amino acids may eliminate potential confusion on the source of amino acids or the misconception that translation produces amino acids (Fisher, 1985). Additionally, requiring students to add both DNA and nucleotides to the CFPS reaction vessel could help resolve student misconceptions that DNA is converted into RNA via a chemical reaction instead of being used as a template for a new nucleotide strand (Wright et al., 2014).

The shifts in responses to attitudinal-based questions showcase the usefulness of the Genetic Code Kit to prepare students for future careers in laboratory science. Notably, these benefits to students extend beyond the learning gains in the content-based questions to support increased student confidence with the laboratory techniques used for CFPS. This work suggests that improvements to familiarity with biotechnologies and comfort in implementing biotechnology-based experiments provide fundamental advances toward workforce development. Prior work has documented that exposing students to research as part of science curriculum has improved student engagement in research outside of the classroom (Lindsay and McIntosh, 2000). Furthermore, undergraduate involvement in research experiences is known to increase student interest in obtaining a Ph.D. and pursuing a STEM field, especially when students are invested and interested in their research (Russell et al., 2007).

In order to enable all students to access these learning outcomes, the Genetic Code Kit is designed to be a low-cost, easy to assemble and implement, highly tailorable platform for various curricula and learning objectives, and requires minimal training and equipment. The Genetic Code Kit costs $4.08 per student, based on 4x CFPS reactions per student (**Table 2**). The cost of $1.02 per 30 μL reaction is inclusive of all materials, reagents, and labor at an estimated rate of $25/hr for the technician's efforts. The development of the previously reported CFAI workflow has allowed us to significantly reduce the time required for cell extract preparation, reducing the cost associated with labor (Levine et al., 2019b). For example, preparing kits for 375 students requires under 25 person-hours. Notably, our kit preparation can be completed entirely by undergraduate students, as was done in this work, which significantly reduces the cost of implementation. The Genetic Code Kit preparation is also highly scalable. In fact, preparing larger quantities becomes more cost-effective. After the cost of labor, the next largest expense is the energy reagents that drive the PANOxSP-based CFPS reaction, but prior work has shown that this cost could be further reduced by leveraging glucose metabolism (Calhoun and Swartz, 2005). Instructors and institutions now benefit from a variety of CFPS options for their classrooms and **Table 2** provides a list of options to choose from. We include cost comparisons in **Table 2**, since this may be one possible driver for selecting a path to implementing CFPS. However, we urge instructors to review the benefits of all listed options, as they may outweigh costs, particularly for convenience of implementation or suitability to specific learning objectives.

The Genetic Code Kit can be tailored to meet a variety of learning objectives beyond teaching transcription and translation. The open nature of the system makes it poised to support inquiry-based learning at a variety of grade levels and CUREs through minor modifications to the reaction setup or DNA template described here. These possibilities can help tailor the kit to the desired grade level and course learning objectives, and include (1) the sequence-function relationships of various genetic elements such as promoters, ribosome binding sites, and codon optimization, (2) riboswitches and aptamers, (3) genetic circuits, (4) CRISPR, (5) probing the mechanisms of various antibiotics, such as protein synthesis inhibitors, and many more. Some unique applications of CFPS for classroom instruction have already been developed for the BioBits kits (Huang et al., 2018; Stark et al., 2018, 2019). Lastly, the Genetic Code Kit can be implemented as a free-standing laboratory module to fit within a single 3-h lab course, but it can also be integrated into existing curricula. For example, this lab could be preceded by molecular

TABLE 2 | Cell-free protein synthesis (CFPS) reaction costs for in-house and commercially available kits.

Product	Vol/rxn (μL)	Cost/rxn	Cost/student	Cost/100 students	References
Genetic Code Kit	30	$ 1.02	$ 4.08	$ 408	Levine et al., 2019b
miniPCR BioBits	7	$ 2.97	$ 11.88	$ 1,235	Stark et al., 2018
Bioneer AccuRapid Midi	30	$ 2.94	$ 11.76	$ 1,544	–
Promega S30 for Circular DNA	30	$ 9.86	$ 39.44	$ 3,944	–
Arbor myTXTL	12	$ 10.65	$ 42.60	$ 4,260	Collias et al., 2019
NEBExpress	30	$ 10.20	$ 40.80	$ 5,100	–
Thermo Expressway Maxi	25	$ 13.20	$ 52.80	$ 5,280	–
Sigma iPE-Quick Kit	30	$ 12.42	$ 49.68	$ 5,400	–

biology labs including PCR or CRISPR and followed by analysis of the protein product via other traditional biochemical methods such as western blotting, ELISA, or SDS-PAGE.

Overall, this work represents the first controlled study of student learning gains resulting from a hands-on, learn-by-doing intervention based on CFPS. While this study's findings are limited by a small sample size and focus on undergraduate students from a single institution, we observed significant gains for learning objectives relating to transcription and translation. Thus, the results of this work provide the foundation to expand assessments of learning gains to various educational levels, pursue multi-institutional efforts that include large student sample sizes, and iterate on the design of the kit to further improve student learning gains for a broad range of learning objectives. We propose that the expansion of this work will further validate the important role of CFPS in biochemical education while supporting workforce development for the growing biotechnology industry.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by human subjects in Research Institutional Review Board at Cal Poly. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LW and NG wrote the manuscript, performed statistical analysis, and generated the figures. LW, NG, BS, and WK prepared all reagents for teaching lab implementation. BS and WK performed reagent storage and time-course testing. AK helped to design the teaching lab implementation and guided the analysis and visualization of assessments of student learning gains. PP helped with statistical analysis of the data. JO conceived the project. KW and JO designed and executed the teaching lab implementation. All authors helped to revise the manuscript and agreed to the accuracy of the work reported.

FUNDING

The authors acknowledge funding support from the Bill and Linda Frost Fund for scholarships to LW, NG, BS, and WK, for supporting general research expenses, and for funding open access publication fees. The authors also acknowledge the Center for Applications in Biotechnology's Chevron Biotechnology Applied Research Endowment Grant, Cal Poly Research, Scholarly, and Creative Activities Grant Program (RSCA 2017), and the National Science Foundation (NSF-1708919).

ACKNOWLEDGMENTS

The authors acknowledge Dr. Eric Jones, Dr. Anya Goodman, and Dr. Steven Wilkinson for supporting the implementation of the Genetic Code Kit in their teaching laboratories for this work; Andrea Laubscher for technical support; Dr. Derek Gragson for support in obtaining historical student data; and Dr. Aaron Engelhart (University of Minnesota), Dr. Seth Bush, Max Levine, Philip Smith, and Logan Burrington for helpful discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.00941/full#supplementary-material

Supplementary Data Sheet 5 contains the main Supplementary Materials documentation. **Supplementary Data Sheet 6** contains the information for **Supplementary Table 2**.

REFERENCES

Altiparmak, M., and Nakiboglu Tezer, M. (2009). Hands on group work paper model for teaching DNA structure, central dogma and recombinant DNA. *US China Educ. Rev.* 6, 19–23.

Armbruster, P., Patel, M., Johnson, E., and Weiss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE Life Sci. Educ.* 8, 203–213. doi: 10.1187/cbe.09-03-0025

Bassiri, E. A. (2011). pGLO mutagenesis: a laboratory procedure in molecular biology for biology students. *Biochem. Mol. Biol. Educ.* 39, 432–439. doi: 10.1002/bmb.20538

Benítez-Mateos, A. I., Llarena, I., Sánchez-Iglesias, A., and López-Gallego, F. (2018). Expanding one-pot cell-free protein synthesis and immobilization for on-demand manufacturing of biomaterials. *ACS Synth. Biol.* 7, 875–884. doi: 10.1021/acssynbio.7b00383

Bopegedera, A. M. R. P. (2011). Putting the laboratory at the center of teaching chemistry. *J. Chem. Educ.* 88, 443–448. doi: 10.1021/ed100045z

Bundy, B. C., Porter Hunt, J., Jewett, M. C., Swartz, J. R., Wood, D. W., Frey, D. D., et al. (2018). Cell-free biomanufacturing. *Curr. Opin. Chem. Eng.* 22, 177–183.

Calhoun, K. A., and Swartz, J. R. (2005). Energizing cell-free protein synthesis with glucose metabolism. *Biotechnol. Bioeng.* 90, 606–613. doi: 10.1002/bit.20449

Chang, P.-S., Lee, S.-H., and Wen, M. L. (2020). Metacognitive inquiry activities for instructing the central dogma concept: 'button code' and 'beaded bracelet making.'. *J. Biol. Educ.* 54, 47–62. doi: 10.1080/00219266.2018.1546756

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn, New York, NY: Routledge.

Collias, D., Marshall, R., Collins, S. P., Beisel, C. L., and Noireaux, V. (2019). An educational module to explore CRISPR technologies with a cell-free transcription-translation system. *Synth. Biol.* 4:ysz005. doi: 10.1093/synbio/ysz005

Debruyn, J. M. (2012). Teaching the central dogma of molecular biology using jewelry. *J. Microbiol. Biol. Educ.* 13, 62–64. doi: 10.1128/jmbe.v13i1.356

Deutch, C. E. (2019). Transformation of *Escherichia coli* with the pGLO plasmid: going beyond the kit. *Am. Biol. Teach.* 81, 52–55. doi: 10.1525/abt.2019.81.1.52

Dopp, B. J. L., Tamiev, D. D., and Reuel, N. F. (2019). Cell-free supplement mixtures: elucidating the history and biochemical utility of additives used to support in vitro protein synthesis in E. *coli* extract. *Biotechnol. Adv.* 37, 246–258. doi: 10.1016/J.BIOTECHADV.2018.12.006

Dopp, J. L., and Reuel, N. F. (2018). Process optimization for scalable *E. coli* extract preparation for cell-free protein synthesis. *Biochem. Eng. J.* 138, 21–28. doi: 10.1016/J.BEJ.2018.06.021

Dorrell, M. I., and Lineback, J. E. (2019). Using shapes & codes to teach the central dogma of molecular biology: a hands-on inquiry-based activity. *Am. Biol. Teach.* 81, 202–209. doi: 10.1525/abt.2019.81.3.202

Duncan, R. G., and Reiser, B. J. (2007). Reasoning across ontologically distinct levels: students' understandings of molecular genetics. *J. Res. Sci. Teach.* 44, 938–959. doi: 10.1002/tea.20186

Fisher, K. M. (1985). A misconception in biology: amino acids and translation. *J. Res. Sci. Teach.* 22, 53–62. doi: 10.1002/tea.3660220105

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8410–8415. doi: 10.1073/pnas.1319030111

Gräwe, A., Dreyer, A., Vornholt, T., Barteczko, U., Buchholz, L., Drews, G., et al. (2019). A paper-based, cell-free biosensor system for the detection of heavy metals and date rape drugs. *PLoS One* 14:210940. doi: 10.1371/journal.pone.0210940

Gregorio, N. E., Kao, W. Y., Williams, L. C., Hight, C. M., Patel, P., Watts, K. R., et al. (2020). Unlocking applications of cell-free biotechnology through enhanced shelf life and productivity of *E. coli* extracts. *ACS Synth. Biol.* 9, 766–778. doi: 10.1021/acssynbio.9b00433

Gregorio, N. E., Levine, M. Z., and Oza, J. P. (2019). A user's guide to cell-free protein synthesis. *Methods Protoc.* 2:24. doi: 10.3390/mps2010024

Haak, D. C., HilleRisLambers, J., Pitre, E., and Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science* 332, 1213–1216. doi: 10.1126/science.1204820

Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Cit. Am. J. Phys.* 66:64. doi: 10.1119/1.18809

Huang, A., Nguyen, P. Q., Stark, J. C., Takahashi, M. K., Donghia, N., Ferrante, T., et al. (2018). BioBitsTM explorer: a modular synthetic biology education kit. *Sci. Adv.* 4:eaat5105. doi: 10.1126/sciadv.aat5105

Ibarra-Herrera, C. C., Carrizosa, A., Yunes-Rojas, J. A., and Mata-Gómez, M. A. (2019). Design of an app based on gamification and storytelling as a tool for biology courses. *Int. J. Interact. Des. Manuf.* 13, 1271–1282. doi: 10.1007/s12008-019-00600-8

Jung, J. K., Alam, K. K., Verosloff, M. S., Capdevila, D. A., Desmau, M., Clauer, P. R., et al. (2020). Cell-free biosensors for rapid detection of water contaminants. *Nat. Biotechnol.* 6, 1–9. doi: 10.1038/s41587-020-0571-7

Karim, A. S., Rasor, B. J., and Jewett, M. C. (2020). Enhancing control of cell-free metabolism through pH modulation. *Synth. Biol.* 5:ysz027. doi: 10.1093/synbio/ysz027

Kiste, A. L., Hooper, R. G., Scott, G. E., and Bush, S. D. (2016). Atomic tiles: manipulative resources for exploring bonding and molecular structure. *J. Chem. Educ.* 93, 1900–1903. doi: 10.1021/acs.jchemed.6b00361

Kiste, A. L., Scott, G. E., Bukenberger, J., Markmann, M., and Moore, J. (2017). An examination of student outcomes in studio chemistry. *Chem. Educ. Res. Pract.* 18, 233–249. doi: 10.1039/C6RP00202A

Kontra, C., Lyons, D. J., Fischer, S. M., and Beilock, S. L. (2015). Physical experience enhances science learning. *Psychol. Sci.* 26, 737–749. doi: 10.1177/0956797615569355

Kozma, R., Chin, E., Russell, J., and Marx, N. (2000). The roles of representations and tools in the chemistry laboratory and their implications for chemistry learning. *J. Learn. Sci.* 9, 105–143. doi: 10.1207/s15327809jls0902_1

Levine, M. Z., Gregorio, N. E., Jewett, M. C., Watts, K. R., and Oza, J. P. (2019a). *Escherichia coli*-based cell-free protein synthesis: protocols for a robust, flexible, and accessible platform technology. *J. Vis. Exp.* 25:e58882.

Levine, M. Z., So, B., Mullin, A. C., Watts, K. R., and Oza, J. P. (2019b). Redesigned upstream processing enables a 24-hour workflow from *E. coli* cells to cell-free protein synthesis. *bioRxiv* [Preprint], doi: 10.1101/729699

Lindsay, H. A., and McIntosh, M. C. (2000). Early exposure of undergraduates to the chemistry research environment: a new model for research universities. *J. Chem. Educ.* 77, 1174–1177. doi: 10.1021/ed077p1174

Marshall, P. A. (2017). A Hands-on activity to demonstrate the central dogma of molecular biology via a simulated VDJ recombination activity. *J. Microbiol. Biol. Educ.* 18:1277. doi: 10.1128/jmbe.v18i2.1277

National Research Council (2000). *How People Learn*. Washington, DC: Academies Press.

Newman, D. L., Snyder, C. W., Fisk, J. N., and Wright, L. K. (2016). Development of the central dogma concept inventory (CDCI) assessment tool. *CBE Life Sci. Educ.* 15, 1–14. doi: 10.1187/cbe.15-06-0124

Newman, D. L., and Wright, L. K. (2013). Using PCR to target misconceptions about gene expression. *J. Microbiol. Biol. Educ.* 14, 93–100. doi: 10.1128/jmbe.v14i1.539

Nogaj, L. A. (2013). Using active learning in a molecular biology. *J. Coll. Sci. Teach.* 42, 50–55.

Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., et al. (2016). Rapid, low-cost detection of zika virus using programmable biomolecular components. *Cell* 165, 1255–1266. doi: 10.1016/J.CELL.2016.04.059

Park, Y. J., Lee, K.-H., and Kim, D.-M. (2017). Assessing translational efficiency by a reporter protein co-expressed in a cell-free synthesis system. *Anal. Biochem.* 518, 139–142. doi: 10.1016/J.AB.2016.11.019

Pigage, H. K. (1991). The central dogma in action. *Am. Biol. Teach.* 53, 436–437. doi: 10.2307/4449352

Queloz, A., Klymkowsky, M. W., Stern, E., Hafen, E., and Köhler, K. (2017). Diagnostic of students' misconceptions using the biological concepts instrument (BCI): a method for conducting an educational needs assessment. *PLoS One* 12:e0176906. doi: 10.1371/journal.pone.0176906

Rotbain, Y., Marbach-Ad, G., and Stavy, R. (2008). Using a computer animation to teach high school molecular biology. *J. Sci. Educ. Technol.* 17, 49–58. doi: 10.1007/s10956-007-9080-4

Russell, S. H., Hancock, M. P., and McCullough, J. (2007). Benefits of undergraduate research experiences. *Science* 316, 548–549. doi: 10.1126/science.1140384

Salehi, A. S. M., Shakalli Tang, M. J., Smith, M. T., Hunt, J. M., Law, R. A., Wood, D. W., et al. (2017). Cell-free protein synthesis approach to biosensing hTRβ-specific endocrine disruptors. *Anal. Chem.* 89, 3395–3401. doi: 10.1021/acs.analchem.6b04034

Seki, E., Matsuda, N., Yokoyama, S., and Kigawa, T. (2008). Cell-free protein synthesis system from *Escherichia coli* cells cultured at decreased temperatures improves productivity by decreasing DNA template degradation. *Anal. Biochem.* 377, 156–161. doi: 10.1016/J.AB.2008.03.001

Silverman, A. D., Akova, U., Alam, K. K., Jewett, M. C., and Lucks, J. B. (2019). Design and optimization of a cell-free atrazine biosensor. *bioRxiv* [Preprint], doi: 10.1101/779827

Stark, J. C., Huang, A., Hsu, K. J., Dubner, R. S., Forbrook, J., Marshalla, S., et al. (2019). BioBits health: classroom activities exploring engineering, biology, and human health with fluorescent readouts. *ACS Synth. Biol.* 8, 1001–1009. doi: 10.1021/acssynbio.8b00381

Stark, J. C., Huang, A., Nguyen, P. Q., Dubner, R. S., Hsu, K. J., Ferrante, T. C., et al. (2018). BioBitsTM bright: a fluorescent synthetic biology education kit. *Sci. Adv.* 4:eaat5107. doi: 10.1126/sciadv.aat5107

Takahashi, M. K., Tan, X., Dy, A. J., Braff, D., Akana, R. T., Furuta, Y., et al. (2018). A low-cost paper-based synthetic biology platform for analyzing gut microbiota and host biomarkers. *Nat. Commun.* 9, 1–12. doi: 10.1038/s41467-018-05864-4

Takemura, M., and Kurabayashi, M. (2014). Using analogy role-play activity in an undergraduate biology classroom to show central dogma revision. *Biochem. Mol. Biol. Educ.* 42, 351–356. doi: 10.1002/bmb.20803

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Nicole Arroyo, E., Behling, S., et al. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proc. Natl. Acad. Sci. U.S.A.* 117, 6476–6483. doi: 10.1073/pnas.1916903117

Ward, W. W., Swiatek, G. C., and Gonzalez, D. G. (2000). Green fluorescent protein in biotechnology education. *Methods*

Enzymol. 305, 672–680. doi: 10.1016/S0076-6879(00)05518-X

Wright, K., Fisk, J., and Newman, D. L. (2014). DNA→RNA: what do students think the arrow means? *CBE Life Sci. Educ.* 13, 338–348. doi: 10.1187/cbe.CBE-13-09-0188

Yin, G., and Swartz, J. R. (2004). Enhancing multiple disulfide bonded protein folding in a cell-free system. *Biotechnol. Bioeng.* 86, 188–195. doi: 10.1002/bit.10827

Zacharia, Z. C., Loizou, E., and Papaevripidou, M. (2012). Is physicality an important aspect of learning through science experimentation among kindergarten students? *Early Child. Res. Q.* 27, 447–457. doi: 10.1016/j.ecresq.2012.02.004

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

PERSPECTIVE
published: 19 August 2020
doi: 10.3389/fbioe.2020.00927

Roadmap to Building a Cell: An Evolutionary Approach

*Zhanar Abil and Christophe Danelon**

Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, Netherlands

Laboratory synthesis of an elementary biological cell from isolated components may aid in understanding of the fundamental principles of life and will provide a platform for a range of bioengineering and medical applications. In essence, building a cell consists in the integration of cellular modules into system's level functionalities satisfying *a* definition of life. To achieve this goal, we propose in this perspective to undertake a semi-rational, system's level evolutionary approach. The strategy would require iterative cycles of genetic integration of functional modules, diversification of hereditary information, compartmentalized gene expression, selection/screening, and possibly, assistance from open-ended evolution. We explore the underlying challenges to each of these steps and discuss possible solutions toward the bottom-up construction of an artificial living cell.

Keywords: synthetic cell, artificial cell, liposome, bottom-up synthetic biology, directed evolution, system's level evolution

OPEN ACCESS

Edited by:
Yuan Lu,
Tsinghua University, China

Reviewed by:
Pasquale Stano,
University of Salento, Italy
Julio Augusto Freyre-Gonzalez,
National Autonomous University of Mexico, Mexico
Juli Peretó,
University of Valencia, Spain

***Correspondence:**
Christophe Danelon
c.j.a.danelon@tudelft.nl

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

Received: *23 May 2020*
Accepted: *20 July 2020*
Published: *19 August 2020*

Citation:
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

INTRODUCTION

All known life forms are composed of cells as the elementary unit. Nevertheless, the staggering complexity of even the simplest known cells is prohibitive to our understanding of the most basic principles of life. Building a biological cell from scratch, i.e., from a minimal set of loose components, such as *in vitro* synthesized or purified biomolecules, would serve as a forward engineering approach, thus illuminating the design principles of life (Elowitz and Lim, 2010). Such construction of a minimal, or synthetic, cell bottom-up is considered today to be one of the grand challenges in synthetic biology. Although the challenge first arose as an approach to answering origin of life questions, the field grew considerably in recent decades, and is now also tapping into the perspectives for improving applications in biotechnology and biomedicine (Ausländer et al., 2017). Approaches that have been used so far toward building synthetic cells and their potential applications have been extensively reviewed elsewhere (Forster and Church, 2006; Luisi et al., 2006; Noireaux et al., 2011; Nourian et al., 2014; Xu et al., 2016; Jia et al., 2017; Sikkema et al., 2019; Stano, 2019). Most studies attempt to define the molecular hardware, i.e., enumerate relevant genes, proteins, biological mechanisms, and metabolic pathways that may compose a minimal cell. Yet, how practically the myriad of components will assemble into a functional cell remains a largely unexplored area. This perspective addresses this challenge by conceptualizing an evolutionary synthetic biology route. Specifically, we argue that *in vitro* evolution must be employed as an engineering tool to accelerate the optimization of individual modules as well as their integration into system's level functionalities. Numerous iterations of module integration, genetic diversification, and phenotype assessment would gradually increase the complexity and autonomy of the system, i.e., its degree of "aliveness."

MORE THAN THE SUM OF ITS PARTS

A conceptual difficulty in constructing a synthetic cell bottom-up lies in the recognition that life is not simply a construct composed of a mixture of chemicals that possesses certain physical properties, but rather is a highly complex physico-chemical phenomenon with yet to be fully understood emergent properties. In other words, to construct a living system from scratch, it is not enough to simply emulate some of life's properties, such as compartmentalization, growth, or information replication, but bring all the processes and functions together into a phenomenon that cannot be reduced to its parts (**Figures 1A,B**).

One of the features that are considered to be defining of life is the functional integration of elementary processes into system's level properties. Some of life's functionalities, such as self-organization, self-maintenance, and information continuity, have been extensively studied and attempts have been made to at least partially reproduce them in the laboratory (Caschera and Noireaux, 2014). However, *integration* of these functionalities in a single synthetic cell still remains far out of reach (Caschera and Noireaux, 2014). Another, related, property of living systems and a goal to achieve in the efforts of building a minimal cell is *autonomy*. By autonomy, we understand a property of a far-from equilibrium, thermodynamically open system that enables it to establish an organizational identity of its own as a result of a set of endergonic-exergonic couplings of internal processes (Ruiz-Mirazo et al., 2004). Hence, a system would be considered autonomous if it is able to maintain its far-from equilibrium state by means of intrinsically governed building of its components and operation of vital processes, provided there is an inflow of necessary substrates and outflow of byproducts. Finally, while basic autonomous systems might be capable of some degree of adaptability to changing environmental conditions, this might not be enough for the survival of a cellular population in a less stable environment. Thus, there is a belief among a number of experts that mechanisms for reliable information continuity and hereditary adaptability, or *evolution,* are required for the survival of a minimal cell in the long run, and constitute one of the defining features of life (Joyce, 1994; Koshland, 2002; Ruiz-Mirazo et al., 2012). Thus, by aiming to construct a minimal cell, we are seeking to construct a functionally integrated phenomenon capable of novel, *system's level functionalities*, such as autonomy and evolvability.

Another, practical, difficulty in building a minimal cell is the sheer volume of the variety of molecules, modules, functions, and their inter-relationship that needs to be explored. The first challenge is that, once chosen, the individual modules might not work as intended in the new environment (altered expression, degradation, loss of activities in the new environment, and broken regulation). This is particularly expected if the constituents originate from different organisms (**Figure 1B**). Additional engineering of individual parts would thus be necessary. Furthermore, a synthetic cell will require a functional integration of multiple modules, which implies further engineering of functional orthogonality, compatibility, and cross-regulation of parts. This approach traditionally involves detailed analysis of individual modules and laborious fine-tuning processes to make the genetic circuits, pathways, and other complex networks to function as desired. However, with increased complexity, the task becomes prohibitively difficult.

A SYSTEM'S LEVEL EVOLUTIONARY APPROACH

We propose to address these difficulties of building a synthetic cell by using a system's level evolutionary approach (**Figures 1C,D**), which offers an achievable alternative to the traditional rational engineering strategy. We know this thanks to the substantial contribution, in the last few decades, of directed evolution methods to engineering of a vast number and variety of proteins (Arnold, 2019) and *in vitro* evolution methods for engineering ribozymes and aptamers (Wilson and Szostak, 1999). Specifically, successful evolution of proteins *in vitro* and especially in liposomes have been demonstrated (Matsuura and Yomo, 2006; Nishikawa et al., 2012). One of the great advantages of the evolutionary approach to engineering of complex systems is that it does not require *a priori* knowledge of structure-function relationships, and, in our case, of inter-relation of multiple components in pathways, genetic circuits, or functional modules in even larger synthetic systems. Thus, we believe that due to our very limited understanding of the nature of a minimal living system, the system's level evolutionary approach is the only practical way we can progress toward the goal of building a synthetic cell.

The system's level evolutionary approach likewise underlines our current understanding of the origin of the first biological cell via chemical evolution (Ruiz-Mirazo et al., 2017; Takeuchi et al., 2017). Reminiscent of how chemical evolution contributed to the stepwise increase in a protocell's complexity and dynamic kinetic stability (Pross, 2013), we envision artificial evolution driving the gradual increase in complexity of a minimal synthetic cell to a system that has a higher degree of "aliveness." Our strategy of building a cell thus embraces the idea of fuzzy logic that allows partial set membership rather than clear-cut definition of life with strict criteria (Bruylants et al., 2010). This is not unlike how prebiotic systems are considered in a continuum of grades of "aliveness." Therefore, building a synthetic cell via numerous intermediates from a more primitive to a more complex autonomous system could shed light on the fundamental principles of the emergence of life.

Finally, as we mentioned above, evolution is often considered as one of life's defining features; and thus, engineering a synthetic cell precursor for evolvability may be an attractive goal in itself. Hence, the development of a continuous *in vitro* evolution approach, which will be discussed below, will not only accelerate the engineering of a synthetic cell, but will become a synthetic cell feature that ensures increased robustness and adaptability to varying environmental conditions, and will require the seamless integration of information replication, diversification, expression, selection, and distribution among daughter synthetic cells.

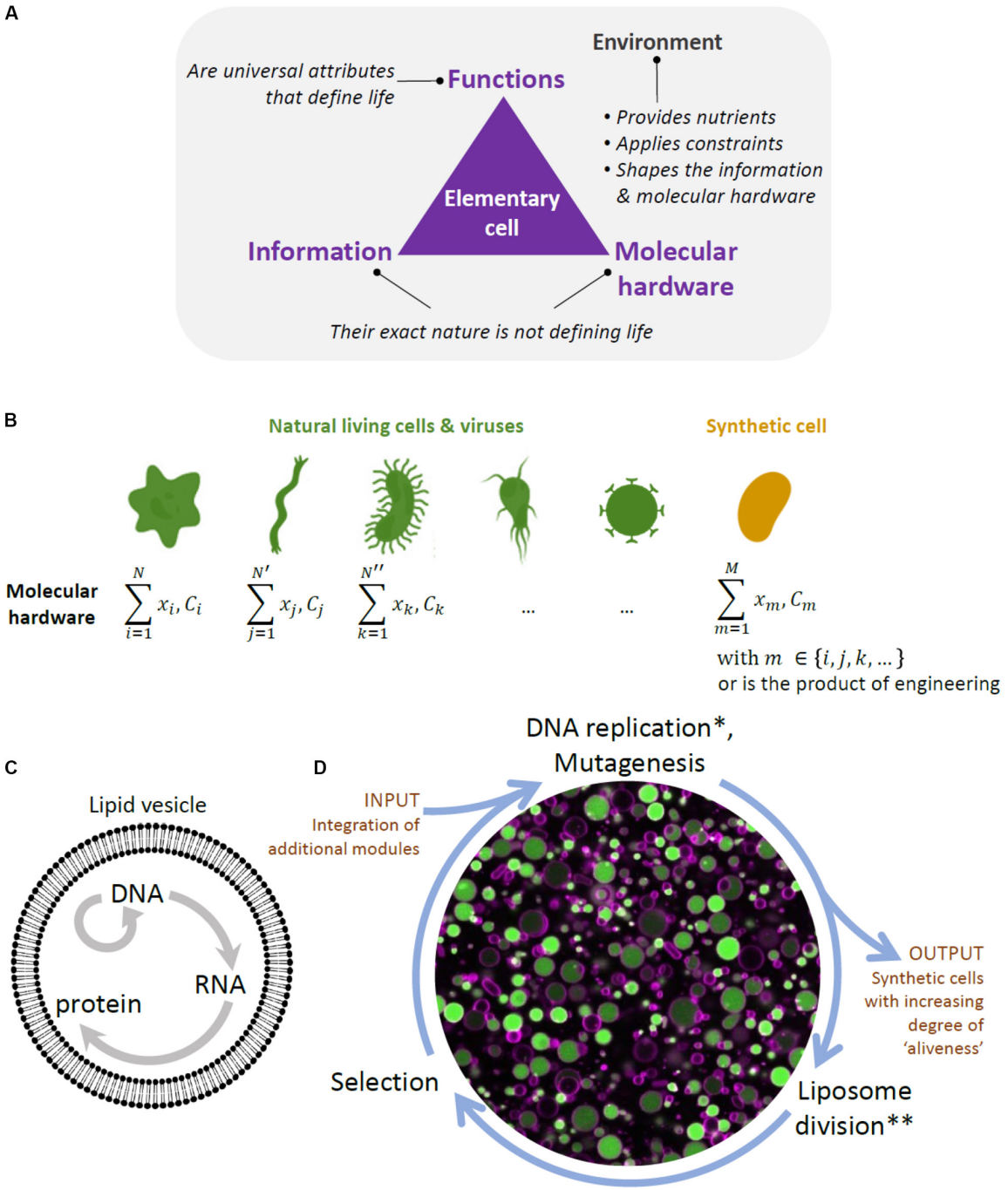


FIGURE 1 | Framework to designing a synthetic cell. **(A)** Building a cell is also a way to question the fundamental principles of life. The diagram depicts our view of the key ingredients constituting a biological cell, and by extension an elementary cell. Cellular unit and environment are entangled and they both constitute the "living system". We consider information (i.e., the sequence of nucleotides in the DNA and RNA of present-day organisms and viruses) and molecular hardware (i.e., the memory-carrying molecule and the set of proteins, lipids, cofactors, etc, composing the cell) as defining elements of life, but their exact nature is *not*, as they are both subject to diversity across organisms. In contrast, minimal cellular life can be seen as a set of universal functions that are common to all living things. **(B)** The molecular hardware is defined as the nature and concentration (*C*) of each cell's constituents (*x* refers to their localization). It differs from one organism to the other and, even, in the modern view of biology, between individual cells from a clonal population. The molecular hardware of a synthetic cell may be a combination of natural parts derived from existing organisms and viruses, and engineering products through *in vitro* evolution or *de novo* design. The complexity (or simplicity) of the molecular hardware depends on the composition of the environment. Synthesizing a cell following this approach will yield never-born cells, whose initial conditions supporting life do not exist in nature. **(C)** Liposome and the key reactions forming the central dogma in biology. **(D)** Scheme of the *in vitro* continuous evolution cycle applied to synthetic lipid vesicles. Center, microscopy image of liposomes (colored in magenta) expressing the protein GFP (green) from its gene. *In discontinuous evolution mode, this step would consist of DNA isolation and bulk amplification. **In discontinuous evolution mode, the DNA amplified in bulk is re-encapsulated in new liposomes.

In the following, we would like to discuss the challenges for successful implementation of the evolutionary approach and propose possible solutions to tackle them.

CHALLENGES AND POSSIBLE SOLUTIONS

Choice of Scaffold

A number of scaffolds for the bottom-up construction of a minimal cell have been explored. Membranous vesicles, emulsion droplets, and coacervates all have their own strengths and disadvantages (Buddingh' and Van Hest, 2017), and will undoubtedly contribute to the ultimate goal conceptually and as parallel experimental platforms. However, we believe that liposomes, aqueous vesicles encapsulated by a phospholipid bilayer, may be the most promising choice in the long run (**Figure 1C**). First, liposomes are vesicles that resemble modern cells the most, and thus have the most potential in shedding light on the mechanisms of life as we know it. Second, they are biocompatible, meaning that they can be more easily integrated with other building blocks, which will most likely come from existing biomolecule types, such as membrane proteins and polysaccharides. Biocompatibility also entails that they could be used in combination with other currently characterized biomolecules for biotechnological or biomedical applications (i.e., bioreactors for sugar production or nucleic acid delivery).

Integration of Modules for System's Level Evolution

The system's level evolution of a synthetic cell may be directed and facilitated by applying a semi-rational approach. Therein, new polypeptide domains, new genes, and entire functional modules are added intermittently throughout the evolution campaign, followed by multiple rounds of additional genetic diversification and phenotypic enrichment. Multiple modules will need to be integrated in a functional way. However, the goal is not to reconstitute an existing living organism like *Escherichia coli* or a mycoplasma, but building a novel, minimal system, composed of a variety of modules, and inspired from multiple contemporary organisms and viruses (**Figure 1B**).

The order of addition of functional modules and concomitant re-design of the environment will need to be carefully considered. Since we will be undertaking an evolutionary approach, the primary and central module would be the one ensuring genetic heredity. One could start with a simple hereditary system capable of self-replication, such as a DNA polymerase and its auxiliary proteins (Van Nies et al., 2018). The initial set of genes could be expressed using an *in vitro* transcription-translation system, such as PURE system (Shimizu et al., 2001). Some of the modules that would require subsequent integration are those responsible for rudimentary anabolism, catabolism, various internal and external feedback and regulatory mechanisms, as well as controlled division mechanisms. When an additional module is functionally integrated (for production of a biomolecule, for example), it would remove the need for external supply of the corresponding component or physical intervention of the researcher. This will be used as the newly added selection pressure. Hence, in general, with the growing autonomy of the synthetic cell, the complexity of the environment and the need for physical intervention should decrease.

It is important to note that by integration we do not mean simply adding the individual modules to the growing DNA chain, and applying the traditional directed evolution methods, such as genetic diversification and phenotype interrogation. A traditional approach would require total destruction of vesicles, as well as extraction and subsequent re-encapsulation of DNA. This discontinuous evolution mode may be applied in the early stages of synthetic cell evolution. However, we are not interested in simply obtaining an optimal synthetic genome. After all, being in possession of an entire *E. coli* genome does not equal to having a living organism. Therefore, we suggest to build a system capable of *in vitro* continuous evolution (**Figure 1D**), where hereditary material gets re-distributed between daughter cells, thus providing a physical continuity of a gene-expressing vesicle population in time. Initially, it would be enabled by a researcher's intervention (e.g., mechanical vesicle division), and eventually, will be the result of the system's intrinsic functions (such as self-encoded division mechanism).

Genetic Diversification of Large Synthetic Genomes

Genetic diversification drives directed evolution, and success of the campaign heavily depends on the nature, size, and quality of the genetic library (Bratulic and Badran, 2017). For single genes or shorter DNA fragments, the commonly used approaches of genetic diversification can be used (Gillam et al., 2014). Increasing complexity of a synthetic cell, however, would necessitate diversification of an increasingly larger synthetic genome, where traditional approaches may become impractical. Herein, we will discuss strategies that may become instrumental for genetic diversification of synthetic genomes, such as assembly and recombination, random, and semi-random mutagenesis methods.

Construction of synthetic genomes will first entail the utilization of commonly used *in vitro* and/or *in vivo* genetic assembly methods (Eriksen et al., 2018; **Figure 2A**). The initial optimization of gene cluster architecture can be performed by combinatorial variation of part choice, gene order, gene orientation, and operon occupancy (**Figure 2B**), as it was shown that these factors are highly useful for genetic optimization (Smanski et al., 2014). Gene assembly methods would also enable genetic recombination, thus ensuring combinatorial rearrangement of mutations between rounds of evolution (**Figure 2A**).

For efficient *in vitro* diversification of large nucleic acids, one can envision the use of a replication system with high processivity, high strand-displacement activity for isothermal

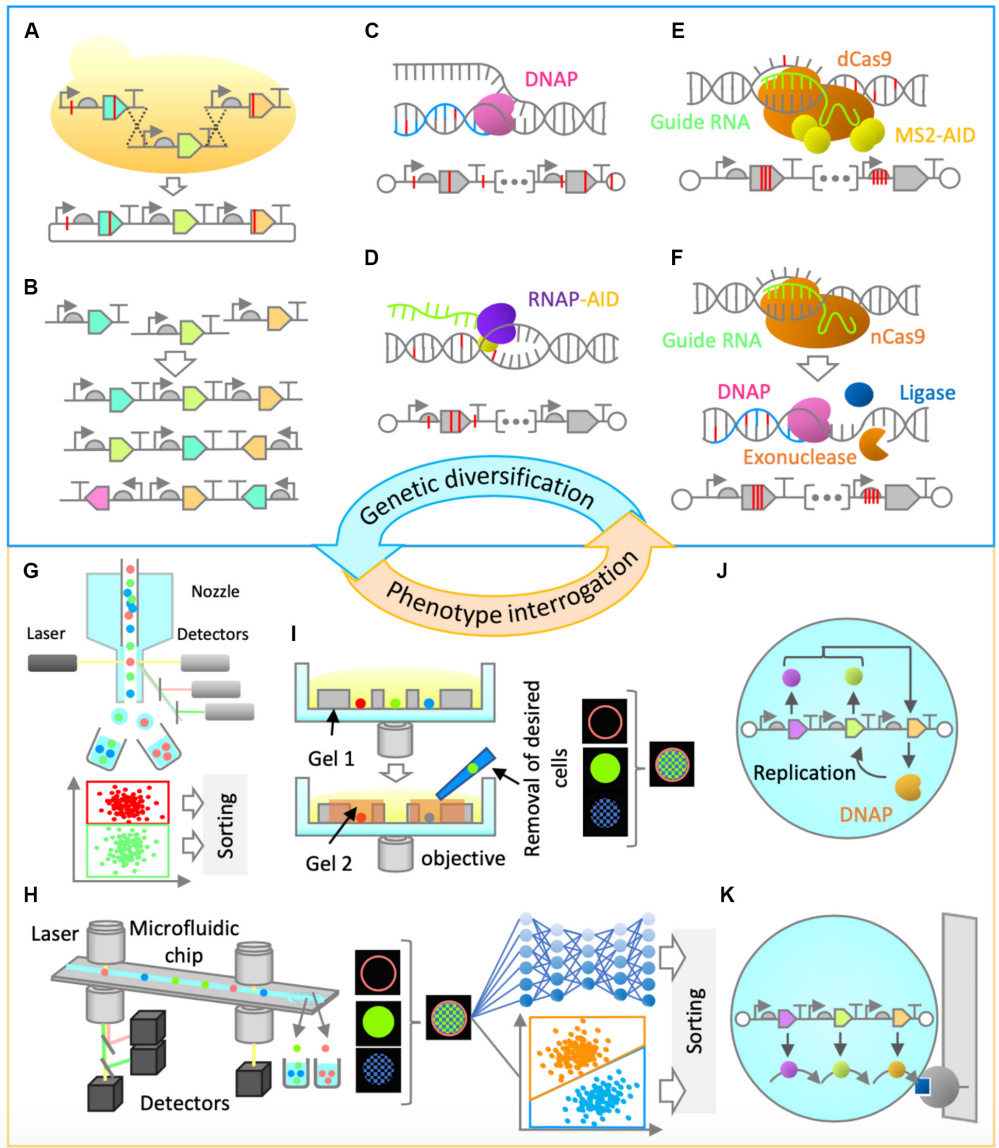


FIGURE 2 | Proposed genetic diversification and phenotype interrogation methods for system's level *in vitro* evolution of gene-expressing vesicles. **(A,B)** Combinatorial design and assembly methods **(A)** Gene assembly and recombination. Homologous recombination in yeast can be performed for gene assembly and recombination of large genetic clusters. **(B)** Gene order and orientation optimization. Constraint-based combinatorial design as well as *in vitro* assembly could be employed for gene expression optimization. **(C,D)** Random mutagenesis methods. **(C)** Low-fidelity DNA replication. Engineered Φ29 or a similar DNA polymerase can be used for error-prone replication of the entire synthetic genome, thus enabling in-vesicle random mutagenesis. **(D)** Base editing. Base editor, such as activation-induced deaminase (AID) can be fused to an RNA polymerase, enabling random mutagenesis in both strands of DNA in gene expressing regions. **(E,F)** Semi-random mutagenesis methods. **(E)** Programmable base editing. Base editors can be tethered to programmable DNA-binding proteins, which will target random mutagenesis to a narrow region of interest. For this purpose, cytidine or adenine deaminase domains (AID, APOBEC1, TadA) can be fused to MS2 coat protein and tethered to endonuclease-deficient versions of Cas9 (dCas9) or Cpf1 (dCpf1) proteins via MS2-hairpins linked to guide RNA. This flexible architecture would allow a random mutagenesis window of around 100 base pairs (Hess et al., 2016). **(F)** Error-prone *in vitro* nick repair. In this proposed method, a nick introduced by a reprogrammable nickase (such as nCas9 or nCpf1) at a region of interest is chewed back with an exonuclease (such as T5 exonuclease), filled in by error-prone DNA polymerase (such as human Poll), and ligated by DNA ligase (such as Taq DNA ligase) to resolve the nick. Such a method would likely constrain the random mutagenesis to up to 200 bases on the targeted strand. **(G–I)** Screening Methods. **(G)** Fluorescence-activated cell sorting (FACS). This high-throughput screening method relies on fluorescence measurement of individual particles and sorting by electrostatic deflection. **(H)** Intelligent image-activated cell sorting (IACS). This high-throughput screening method sorts particles based on their unique morphological features and is based on real-time image-based flow cytometry assisted by artificial intelligence. **(I)** Dual-photopolymerized microwell array sorting. This image-based single-cell sorting method relies on two-step photopolymerization process: the first to create an array of microwells to capture individual cells and the second to encapsulate undesired cells. The desired cells are removed by washing. **(J,K)** Selection methods. **(J)** Compartmentalized partnered replication. When adapted for *in vitro*, in-liposome use, this approach would rely on linking a phenotype of interest to the expression of a DNA polymerase gene. Differential amplification of a synthetic gene cluster by the expressed DNA polymerase in liposomes would be the basis of selection. **(K)** Affinity chromatography. Liposomes displaying a molecule of interest can be enriched by specific interaction with a binding partner on a solid support.

amplification, and low fidelity for accelerated mutagenesis. A suitable candidate with such characteristics is the Φ29 DNA polymerase, provided that its high fidelity can be reduced (**Figure 2C**). The fidelity of Φ29 or a similar DNA polymerase could be lowered by using increased Mn^{2+} ion concentrations (Fujii R. et al., 2014) or by utilizing exonuclease-deficient variants. Alternatively, increased mutagenesis levels could be achieved by adapting the use of DNA deaminase domains (such as cytidine or adenine deaminases) *in vitro* either alone (Al-Qaisi et al., 2018) or in fusion with T7 RNA polymerase (Chen et al., 2019; **Figure 2D**). Development of these and similar methods will be critical in the diversification of large synthetic genomes. However, the sequence space of a synthetic genome would be vast, and the use of random mutagenesis approaches alone may not be enough for efficient evolution campaigns.

More focused, or semi-random, mutagenesis approaches might be needed, assisted by relevant knowledge from biochemical and bioinformatical studies, or simply by modulation of relative gene expression levels in the multi-gene system. For example, to target the mutagenesis to a narrow region of interest, deaminase domains can be tethered to reprogrammable DNA-binding proteins, such as endonuclease-deficient Cas9 or Cpf1 (Eid et al., 2018; **Figure 2E**). The attractiveness of this approach lies in the fact that it does not require a double strand break or a complex cellular repair machinery, and it could be adapted for use *in vitro*. Alternatively, one could engineer a method whereby a nick introduced by a reprogrammable nickase at a region of interest is subsequently or simultaneously treated with an exonuclease to chew back the nicked strand, an error-prone DNA polymerase to mutagenically fill in the gap, and a ligase to covalently repair the nick (**Figure 2F**). This would be reminiscent of the commonly used cloning method (Gibson et al., 2009), but would not require *in vitro* assembly of multiple fragments followed by *in vivo* screening, and would be applicable for a more focused *in vitro* mutagenesis of large DNA fragments. These and other genetic diversification methods still require extensive development and implementation *in vitro*; however, the effort will be vital for successful system's level evolution and synthetic cell construction.

Selection/Screening Strategies

The next challenge is to develop efficient approaches to enrich a population of liposomes for synthetic cells' desired phenotypes (**Figure 1D**). For certain phenotypes, traditional screening and selection methods (Packer and Liu, 2015) can be implemented. For example, if a phenotype of interest can be linked to the expression level of a single gene, such as a fluorescent or luminescent protein, high-throughput screening by fluorescence activated cell sorting (FACS) can be implemented (Fujii S. et al., 2014). Fluorescence activated cell sorting, the most common method for sorting particles, allows for high-throughput and multiparameter sorting (**Figure 2G**). It could also be implemented in cases where a substrate is fluorogenic or affinity-based enrichment is required. Alternatively, it is possible to exploit selection strategies by linking a phenotype to the expression of a DNA polymerase (Abil et al., 2017), given that it is adapted for *in-liposome* amplification (**Figure 2J**) or to a destructive protein (negative selection by destabilizing the liposomes). For affinity-based enrichment, FACS could potentially be substituted with affinity chromatography (**Figure 2K**). In many cases, though, traditional approaches that are suitable for evolution of single genes or simple pathways may not be appropriate enough for system's level evolution, since there will likely be no single parameter defining the system's state for high-throughput screening or selection. Therefore, a paradigm shift in the approaches to tackle this problem is needed.

One way of enriching vesicles for system's level functionalities could be via imaging-based screening methods. Cells' (and by extension, vesicles') "physiological" functions often manifest in morphological identities and molecular localization footprints, and therefore an imaging-based screening approach would offer an unprecedented advantage over conventional high-throughput screening methods, such as FACS. One emerging technology that would enable enrichment based on multidimensional fluorescence images is intelligent image-activated cell sorting (IACS; **Figure 2H**; Nitta et al., 2018; Isozaki et al., 2020). Image-activated cell sorting integrates microfluidics, high-throughput cell microscopy, real-time intelligent image processing, decision making, and cell sorting. The strategy has the potential for opening the door for the exploration of phenotypes that are more directly connected to vesicles' system's level functions. A possible alternative to IACS is image-based sorting via dual-photopolymerized microwell arrays (**Figure 2I**; Sun et al., 2013). This method relies on microwell array-assisted spatial segregation, imaging and identification, and screening based on two-step photopolymerization. This approach offers the advantage of interrogation based on high-resolution imaging, as well as time-lapse and long-term dynamics, an advantage that could be uniquely suited for system's level evolution. Adapting this methodology for screening of micron-sized vesicles and improving it by significantly increasing the throughput capacity would be desirable.

Open Ended Evolution

As an alternative approach to the semi-rational evolutionary approach discussed above, one could choose to undertake a more exploratory, open ended, evolution. Herein, an intrinsic mechanism would be provided for the exploration of new state spaces by the system, thus allowing expansive and transformational evolutionary novelties (Banzhaf et al., 2016; Taylor, 2019). For example, new state spaces for exploration can be provided simply by modifying and re-engineering of the environment. Other mechanisms could be exploitation of exaptation, where a structure that was performing a certain function can switch to performing a different function; and compositional systems, where a new or enhanced function emerges from the combination of smaller part types. Finally, to aid in the discovery of difficult-to-identify features, one can explore the utilization of unsupervised machine learning (Lu et al., 2019). Open ended evolution could provide new unexpected solutions to evolutionary challenges. However, its requirements and mechanisms are still poorly understood and its implementation will require intensive investigation.

CONCLUSION

Herein, we proposed a system's level evolutionary approach for building a synthetic cell. We believe that it delineates the paradigm shift needed to reach this ambitious goal. A number of new methodologies still need to be developed for the implementation of this vision. However, each of these developments will not only bring us closer toward a synthetic cell, but also open the doors to the realization of novel applications in synthetic biology. Infusing cell-free biology with system's level evolution will provide testbeds for exploring the roots of evolutionary processes in living organisms, as well as a new perspective in our understanding of life's fundamental properties. It would also significantly contribute to the ongoing fluid interrelationship between theoretical definition of life and empirical assessment thereof (Amilburu et al., 2020).

The evolutionary strategy will still require expertise, considerable effort, and coordination of more than a single group of researchers. These efforts will be further complicated by the fact that numerous evolutionary routes will likely have to be explored to arrive at a more viable one. Therefore, to facilitate reproducibility and distribution of materials between different groups working on this unifying goal, it would be beneficial to establish or improve platforms for standard part repository (Beal et al., 2020).

A sceptic might wonder if building a synthetic cell via evolution (which encompasses a level of randomness), as opposed to rational engineering, can be useful for our efforts of understanding life and its origins. The first drawback of our approach is that the final system's structure and performance will be highly dependent on its evolutionary history, which would be contingent on experimenter's choices of modules and environmental design, as well as random mutagenesis and genetic drift. The second drawback is that the evolutionary path is extremely unlikely to repeat the prebiotic evolution of terrain life, mostly because of the difficulty to reconstitute in the laboratory the environmental conditions that prevailed on the early Earth. Nonetheless, if we manage to construct artificial life using *in vitro* continuous evolution, this would support the general idea that terrain life likely originated from similar evolutionary processes. If, on the other hand, evolution is found to be the *only* way of constructing a living organism, this fact alone would illuminate a novel view on the genesis of living matter from non-living as well as reveal a fundamental feature of life. The description of life as an *evolved* and *evolving* phenomenon would thus transform from a theoretical concept to an empirical observation.

AUTHOR CONTRIBUTIONS

Both authors contributed to conceptualization, writing, and funding acquisition.

FUNDING

This project has received funding from the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie grant agreement no 707404 and "BaSyC – Building a Synthetic Cell" Gravitation grant (024.003.019).

REFERENCES

Abil, Z., Ellefson, J. W., Gollihar, J. D., Watkins, E., and Ellington, A. D. (2017). Compartmentalized partnered replication for the directed evolution of genetic parts and circuits. *Nat. Protoc.* 12, 2493–2512. doi: 10.1038/nprot.2017.119

Al-Qaisi, T. S., Su, Y. C., and Roffler, S. R. (2018). Transient AID expression for in situ mutagenesis with improved cellular fitness. *Sci. Rep.* 8:9413. doi: 10.1038/s41598-018-27717-2

Amilburu, A., Moreno, Á, and Ruiz-Mirazo, K. (2020). Definitions of life as epistemic tools that reflect and foster the advance of biological knowledge. *Synthese* 1–21. doi: 10.1007/s11229-020-02736-7

Arnold, F. H. (2019). Innovation by evolution: bringing new chemistry to Life (Nobel Lecture). *Angew. Chemie Int. Ed.* 58, 14420–14426. doi: 10.1002/anie.201907729

Ausländer, S., Ausländer, D., and Fussenegger, M. (2017). Synthetic Biology—The Synthesis of Biology. *Angew. Chemie Int. Ed.* 56, 6396–6419. doi: 10.1002/anie.201609229

Banzhaf, W., Baumgaertner, B., Beslon, G., Doursat, R., Foster, J. A., Mcmullin, B., et al. (2016). Defining and simulating open-ended novelty: requirements, guidelines, and challenges. *Theory Biosci.* 135, 131–161. doi: 10.1007/s12064-016-0229-7

Beal, J., Goñi-Moreno, A., Myers, C., Hecht, A., del Vicente, M., Parco, M., et al. (2020). The long journey towards standards for engineering biosystems. *EMBO Rep.* 21:e50521. doi: 10.15252/embr.202050521

Bratulic, S., and Badran, A. H. (2017). Modern methods for laboratory diversification of biomolecules. *Curr. Opin. Chem. Biol.* 41, 50–60. doi: 10.1016/j.cbpa.2017.10.010

Bruylants, G., Bartik, K., and Reisse, J. (2010). Is it useful to have a clear-cut definition of life? On the use of fuzzy logic in prebiotic chemistry. *Orig. Life Evol. Biosph.* 40, 137–143. doi: 10.1007/s11084-010-9192-3

Buddingh', B. C., and Van Hest, J. C. M. (2017). Artificial cells: synthetic compartments with life-like functionality and adaptivity. *Acc. Chem. Res.* 50, 769–777. doi: 10.1021/acs.accounts.6b00512

Caschera, F., and Noireaux, V. (2014). Integration of biological parts toward the synthesis of a minimal cell. *Curr. Opin. Chem. Biol.* 22, 85–91. doi: 10.1016/j.cbpa.2014.09.028

Chen, H., Liu, S., Padula, S., Lesman, D., Griswold, K., Lin, A., et al. (2019). Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor. *Nat. Biotechnol.* 38, 165–168. doi: 10.1038/s41587-019-0331-8

Eid, A., Alshareef, S., and Mahfouz, M. M. (2018). CRISPR base editors: genome editing without double-stranded breaks. *Biochem. J.* 475, 1955–1964. doi: 10.1042/BCJ20170793

Elowitz, M., and Lim, W. A. (2010). Build life to understand it. *Nature* 468, 889–890. doi: 10.1038/468889a

Eriksen, D. T., Chao, R., and Zhao, H. (2018). "Applying advanced DNA assembly methods to generate pathway libraries," in *Synthetic Biology*, eds C. S. Sang, Y. Lee, J. Nielsen, and G. Stephanopoulos (Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA), 331–347. doi: 10.1002/9783527688104.ch16

Forster, A. C., and Church, G. M. (2006). Towards synthesis of a minimal cell. *Mol. Syst. Biol.* 2:45. doi: 10.1038/msb4100090

Fujii, R., Kitaoka, M., and Hayashi, K. (2014). "Error-prone rolling circle amplification greatly simplifies random mutagenesis," in *Directed Evolution Library Creation. Methods in Molecular Biology (Methods and Protocols)*, Vol. 1179, eds E. Gillam, J. Copp, and D. Ackerley (New York, NY: Springer), doi: 10.1007/978-1-4939-1053-3_2

Fujii, S., Matsuura, T., Sunami, T., Nishikawa, T., Kazuta, Y., and Yomo, T. (2014). Liposome display for in vitro selection and evolution of membrane proteins. *Nat. Protoc.* 9, 1578–1591. doi: 10.1038/nprot.2014.107

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Gillam, E. M. J., Copp, J. N., and Ackerley, D. (eds) (2014). *Directed Evolution Library Creation*, 2nd Edn. Totowa, NJ: Humana Press Inc.

Hess, G. T., Frésard, L., Han, K., Lee, C. H., Li, A., Cimprich, K. A., et al. (2016). Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat. Methods* 13, 1036–1042. doi: 10.1038/nmeth.4038

Isozaki, A., Mikami, H., Tezuka, H., Matsumura, H., Huang, K., Akamine, M., et al. (2020). Intelligent image-activated cell sorting 2.0. *Lab Chip* 20, 2263–2273. doi: 10.1039/d0lc00080a

Jia, H., Heymann, M., Bernhard, F., Schwille, P., and Kai, L. (2017). Cell-free protein synthesis in micro compartments: building a minimal cell from biobricks. *N. Biotechnol.* 39, 199–205. doi: 10.1016/j.nbt.2017.06.014

Joyce, G. F. (1994). "Foreword," in *Origins of Life: The Central Concepts*, eds D. W. Deamer and G. Fleischaker (Boston: Jones and Bartlett).

Koshland, D. E. (2002). The seven pillars of life. *Science* 295, 2215–2216.

Lu, A. X., Kraus, O. Z., Cooper, S., and Moses, A. M. (2019). Learning unsupervised feature representations for single cell microscopy images with paired cell inpainting. *PLoS Comput. Biol.* 15:e1007348. doi: 10.1371/journal.pcbi.1007348

Luisi, P. L., Ferri, F., and Stano, P. (2006). Approaches to semi-synthetic minimal cells: a review. *Naturwissenschaften* 93, 1–13. doi: 10.1007/s00114-005-0056-z

Matsuura, T., and Yomo, T. (2006). In vitro evolution of proteins. *J. Biosci. Bioeng.* 101, 449–456. doi: 10.1263/jbb.101.449

Nishikawa, T., Sunami, T., Matsuura, T., and Yomo, T. (2012). Directed evolution of proteins through in vitro protein synthesis in liposomes. *J. Nucleic Acids* 2012:923214. doi: 10.1155/2012/923214

Nitta, N., Sugimura, T., Isozaki, A., Mikami, H., Hiraki, K., Sakuma, S., et al. (2018). Intelligent image-activated cell sorting. *Cell* 175, 266.e13–276.e13. doi: 10.1016/j.cell.2018.08.028

Noireaux, V., Maeda, Y. T., and Libchaber, A. (2011). Development of an artificial cell, from self-organization to computation and self-reproduction. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3473–3480. doi: 10.1073/pnas.1017075108

Nourian, Z., Scott, A., and Danelon, C. (2014). Toward the assembly of a minimal divisome. *Syst. Synth. Biol.* 8, 237–247. doi: 10.1007/s11693-014-9150-x

Packer, M. S., and Liu, D. R. (2015). Methods for the directed evolution of proteins. *Nat. Rev. Genet.* 16, 379–394. doi: 10.1038/nrg3927

Pross, A. (2013). Dynamic kinetic stability (DKS) as a conceptual bridge linking chemistry to biology. *Curr. Org. Chem.* 17, 1702–1703. doi: 10.1007/s00239-013-9556-1

Ruiz-Mirazo, K., Briones, C., and De La Escosura, A. (2017). Chemical roots of biological evolution: the origins of life as a process of development of autonomous functional systems. *Open Biol.* 7, 170050. doi: 10.1098/rsob.170050

Ruiz-Mirazo, K., Moreno, A., Ruiz-Mirazo, K., and Moreno, A. (2012). Autonomy in evolution: from minimal to complex life. *Synthese* 185, 21–52. doi: 10.1007/s11229-011-9874-z

Ruiz-Mirazo, K., Peretó, J., and Moreno, A. (2004). A universal definition of life: autonomy and open-ended evolution. 34, 323–346. doi: 10.1023/b:orig.0000016440.53346.dc

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., et al. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755. doi: 10.1038/90802

Sikkema, H. R., Gaastra, B. F., Pols, T., and Poolman, B. (2019). Cell fuelling and metabolic energy conservation in synthetic cells. *ChemBioChem* 20, 2581–2592. doi: 10.1002/cbic.201900398

Smanski, M. J., Bhatia, S., Zhao, D., Park, Y. J., Woodruff, L. B. A., Giannoukos, G., et al. (2014). Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* 32, 1241–1249. doi: 10.1038/nbt.3063

Stano, P. (2019). Gene expression inside liposomes: from early studies to current protocols. *Chem. A Eur. J.* 25, 7798–7814. doi: 10.1002/chem.201806445

Sun, T., Kovac, J., and Voldman, J. (2013). Image-based single-cell sorting via dual-photopolymerized microwell arrays. *Anal. Chem.* 86, 977–981. doi: 10.1021/ac403777g

Takeuchi, N., Hogeweg, P., and Kaneko, K. (2017). Conceptualizing the origin of life in terms of evolution. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 375. doi: 10.1098/rsta.2016.0346

Taylor, T. (2019). Evolutionary innovations and where to find them routes to open-ended evolution in natural and artificial systems. *Artif. Life* 25, 207–224. doi: 10.1162/artl_a_00290

Van Nies, P., Westerlaken, I., Blanken, D., Salas, M., Mencía, M., and Danelon, C. (2018). Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun.* 9:1583. doi: 10.1038/s41467-018-03926-1

Wilson, D. S., and Szostak, J. W. (1999). In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* 68, 611–647. doi: 10.1146/annurev.biochem.\break68.1.611

Xu, C., Hu, S., and Chen, X. (2016). Artificial cells: from basic science to applications. *Mater. Today* 19, 516–532. doi: 10.1016/j.mattod.2016.02.020

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

OPINION
published: 26 August 2020
doi: 10.3389/fbioe.2020.00953

On the "Life-Likeness" of Synthetic Cells

Luisa Damiano[1] *and Pasquale Stano*[2]*

[1] *Research Group on the Epistemology of the Sciences of the Artificial (RG-ESA), Department of Ancient and Modern Civilizations, University of Messina, Messina, Italy,* [2] *Laboratory of Bio-Organic Chemistry, Department of Biological and Environmental Sciences and Technologies (DiSTeBA), University of Salento, Lecce, Italy*

Keywords: synthetic biology, synthetic cells, artificial cells, autopoiesis, self-organization, Turing test, molecular communications, life-likeness

One of the most exciting and rapidly expanding research area in contemporary science is the bottom-up construction of artificial cell-like systems, also known as synthetic cells. Such approaches are part of the synthetic biology research paradigm, which equates to understanding means constructing. Accordingly, these artificial systems are considered able to generate new knowledge based on explorative procedures that are complementary to traditional scientific investigations. Constructing synthetic cells aims at understanding the emergence of life from scratch at the cellular level, modeling chemically primitive cells for unveiling origins-of-life mysteries, and developing radically new biotechnological tools for medical, industrial, and research applications. The following article is dedicated to one of the most compelling open questions emerging from the rapid improvement of synthetic cell technology: "life-likeness" of synthetic cells. Based on prior work, we promote an 'organizational approach' to the assessment of life-likeness, and, coherently, we propose the transition from behavioral assays, like the Turing test, to systemic strategies, based on concepts such as organization, complexity, networks, and emergence.

OPEN ACCESS

Edited by:
Simon J. Moore,
University of Kent, United Kingdom

Reviewed by:
Yutetsu Kuruma,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Japan
Yifan Liu,
ShanghaiTech University, China

***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: *01 May 2020*
Accepted: *23 July 2020*
Published: *26 August 2020*

Citation:
Damiano L and Stano P (2020) On the "Life-Likeness" of Synthetic Cells. Front. Bioeng. Biotechnol. 8:953. doi: 10.3389/fbioe.2020.00953

1. UNDERSTANDING BY CONSTRUCTING: CELL MODELS IN THE AGE OF SYNTHETIC BIOLOGY

The current age of synthetic biology (SB), for the first time, makes possible novel ways of understanding the emergence of cellular life by means of "constructive" or "synthetic" approaches, also known as "bottom-up" or "understanding-by-building" methods. The operative paradigm is based on the assembly of molecular components into a spatially organized structure that resembles and behaves like a living cell, even if at a minimal complexity level. *Synthetic cell technology* is an innovative blend of microcompartment technology (in particular, but not exclusively, liposomes), cell-free systems, microfluidics, and numerical modeling. The ever increasing number of experimental reports referring to "synthetic cells" (SCs), "artificial cells," and "protocells" shows that the goal of constructing cell-like systems in the laboratory is now within the experimental reach. Such advances trigger new stimulating questions about practical and conceptual issues referring to SC construction, such as organizational issues (complexity, life-likeness), materials (primitive-like, modern biochemicals, fully synthetic), goals (basic or applied science). Motivated by the current discussions about the "science of the artificial" (Cordeschi, 2002; Damiano et al., 2011), here we aim to shortly comment on the first issue, in particular about life-likeness, with the goal of stimulating an open discussion in the community.

2. HOW TO QUANTIFY SC LIFE-LIKENESS?

2.1. The Imitation Game

In 2006 the members of the project CHELLnet (2005–2009) published a thought-provoking position paper focused on the relevant yet elusive question of *how much alive a SC is*. The Authors

start the discussion by remarking that no universally acceptable definition of life is available, while the progresses in SB are going to allow the construction of cell-like systems of increasing complexity. How to determine, then, their life-likeness? Following Turing's proposal for artificial intelligence (AI), the answer was a sort of extension of the Turing test [or "imitation game" (Turing, 1950)] to the realm of SCs, in the framework of the analogies indicated in **Figure 1A**.

In the case of AI, the long-debated question "can machines think?" was re-defined by Turing based on an operational scenario consisting in a machine that *imitates* the act of thinking, in such a way that a human interrogator, using a natural language, cannot distinguish between a machine's answer and a person's answer. In the SB context, the question becomes "are SCs alive?" and thus it can be re-defined by an operational scenario where SCs do perform well (or do not) in a sort of cellular imitation game. The provided example refers to the SCs capacity of establishing a chemical communication with natural biological cells—in the spirit of the original Turing test.

The 2006 paper on the SC Turing test (and the CRAACC workshop held in the same year in Venice, hosted by the ECLT) were illuminating and anticipated new directions in the SC research that we are starting to approach now. Although a real Turing test for SC may still seem unrealistic, these studies have the merit of having raised very intriguing questions.

2.2. Experimental Results

The original idea of applying the Turing test to SCs generated a first pioneer report a few years later. Ben Davis and collaborators published an experimental paper on chemical communications between SCs and biological cells in 2009 (Gardner et al., 2009). The so-called *formose* reaction was carried out inside liposomes. Some of the products, after escaping from the liposome and reacting with borate in the medium, reached *Vibrio harveyi* bacteria and activated a response.

SC technology, however, allows the construction—sometimes microfluidics-aided (Weiss et al., 2018)—of more sophisticated and programmable SCs based on the control of gene expression. These SCs can be seen as general-purpose systems that can be interfaced with biological cells, exchange signals, and be engaged in communicative interactions. An early elaboration of this scenario was obtained by merging pre-existing ideas from SB, bio-engineered nanofactories, and molecular communication networks (Luisi et al., 2006; Leduc et al., 2007; Nakano et al., 2011, 2013; Stano et al., 2012). Experimental results on SCs communicating with other SCs or with biological cells appeared more recently (Lentini et al., 2014, 2017; Adamala et al., 2017; Ding et al., 2018; Niederholtmeyer et al., 2018; Rampioni et al., 2018; Tang et al., 2018; Dupin and Simmel, 2019; Joesaar et al., 2019) showing that establishing chemical communication is now within the reach of several laboratories. Details on the molecular circuitry employed in these studies have been reviewed by Rampioni et al. (2019).

One of these studies is quite relevant and deserves a close attention, because it reports for the first time a *bidirectional* SCs-biological cells communication (Lentini et al., 2017), thus realizing, in a minimal form, the concepts envisioned in the above-mentioned 2006 paper (Cronin et al., 2006). In particular, SCs based on gene expression were able to perceive and decode a primary chemical signal sent by *V. fischeri*, then reply by synthesizing a response signal sent back to, and perceived by, *V. fischeri*. Results were explicitly presented from the viewpoint of the cellular Turing test and attracted the attention of science press (Urquhart, 2017).

Based on the two-way communication between SCs and bacteria, the life-likeness of the SCs employed in that study was estimated, consistently with the seminal idea of the Turing test mentioned above. It is informative, then, commenting on the proposed approach, in light of the present discussion. The strategy is based on measuring the *bacterial* gene expression profile when a successful bidirectional communication channel is established with SCs (see **Figure 1B**), estimating the SC life-likeness as 39% (technical details in **Figure 1B**). The result is at the same time exciting and debatable. The bidirectional communication pattern described above is generated by SCs that just express two genes: a receptor for sensing the signal molecules sent by the bacterium, and a synthase for producing the response signal molecules. How can a system of such (low) complexity displaying 39% life-likeness?

Mansy et al. carefully discussed their conclusions, with the correct intuition that, in addition to the two expressed genes, one should also consider the about 100 macromolecular elements of the transcription-translation (TX-TL) machinery that operate in the background, all essential for expressing the two genes. Accordingly, the value of 39% actually was considered as a proxy for the ratio between the number of genes of the TX-TL machinery and the about 200 genes recognized as the so-called *minimal genome* (Mushegian and Koonin, 1996; Gil et al., 2004; Islas et al., 2004; Forster and Church, 2006), essential for a minimal cellular life. In other words, it is suggested that the 39% value actually refers to a *hypothetical* SCs that could produce all TX-TL macromolecules and, consequently, the two proteins which are functional for the experimentally observed bidirectional molecular communication.

It is important to remark that the hypothetical SCs described above are not yet experimentally accessible, and they would be certainly much more complex (and more life-like) than the actual SCs used in that study. However, and this is the key point of the discussion here—even if that sort of SCs would exist, or other similar system of lower complexity, any analysis based on the gene expression of the *receiving V. fischeri* cells would still have 39% life-likeness.

This conclusion results from monitoring the differences induced in the biological partner for inferring how the synthetic partner *behaves*, irrespective of the structure and the functioning of the synthetic partner (by "functioning" here we mean the mechanisms that generate its behavior, i.e., its organization). Indeed, the gene expression profile of the receiving biological cells would not change if three different SCs with different complexity would engage the same communication with biological cells, as shown in detail in **Figure 1C**.

A

	Turing test for AI	Turing test for SCs
Imitate emergent property	Thought	Cellular functions (*e.g.*, metabolism, evolution, containment)
Embodiment of property	Computational digital machine	Chemical systems (synthetic cells)
Probing mechanism	Questions/answers mediated by natural language	Questions/answers mediated by (natural) physico-chemical language (*e.g.*, interconversion of chemical potentials, mechanical transduction, signalling)

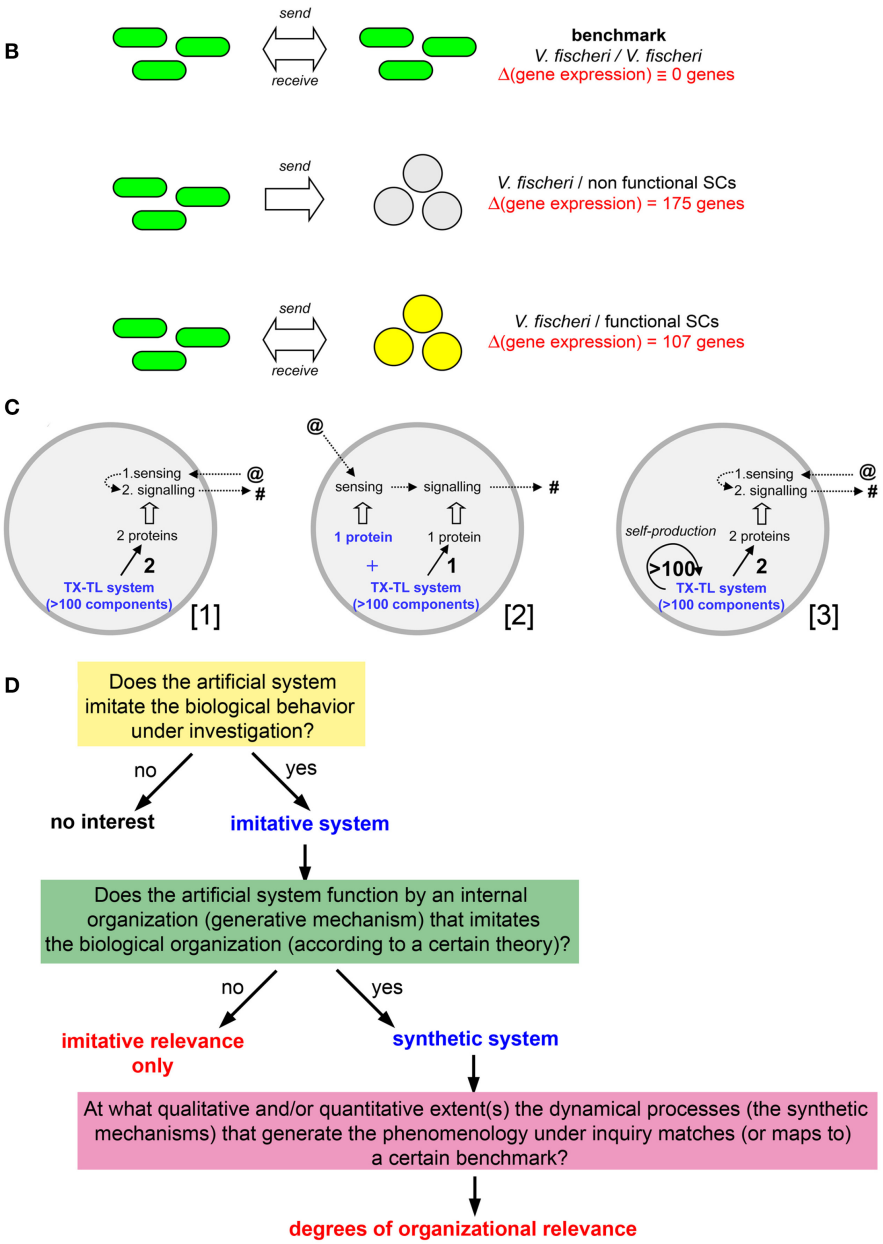


FIGURE 1 | (A) Comparison of Turing tests for intelligence and life, taken from Cronin et al. (2006). **(B)** Graphical summary of the method used by Lentini et al. (2017) to quantify life-likeness of SCs capable of bidirectional chemical communication with *V. fischeri*. (Top) The biological communication between when *V. fischeri* cells

(Continued)

FIGURE 1 | is taken as benchmark against which the synthetic/natural communication is contrasted. (Middle) When *V. fischeri* cells send a signal to SCs that are not able to reply, a change in the expression levels of 175 *V. fischeri* genes is observed. (Bottom) When *V. fischeri* cells send a signal to SCs that are able to reply with a response signal, a change in the expression levels of 107 *V. fischeri* genes is observed. Therefore, SCs behavior is capable of offsetting 68 (= 175 − 107) differences in *V. fischeri* gene expression levels, recovering the 39% (= 68/175) of the biological behavior. **(C)** Three examples of SCs capable of establishing the kind of molecular communication with *V. fischeri* of the same type as described by Lentini et al. (2017). (Left) The actual experimental system made in the laboratory: TX-TL machinery produces two proteins capable of sensing the primary signal (sent by the bacteria), and, as response, produces a secondary signal (sent to the bacteria); this SC is based on the expression of two genes. (Center) A hypothetical SC—simpler than the one shown on the left, where the receptor is not produced by the TX-TL machinery, but it is added from the beginning as ready-to-work protein; this SC is based on the expression of one gene. (Right) A hypothetical SC—very complex and currently out-of-reach—where TX-TL components (>100) are all produced by the TX-TL machinery itself, together with the two proteins needed for communication (this SCs is based on the expression of >100 genes). The three SCs clearly have different complexity ([3] >> [1] > [2]) and different life-likeness, but they would result in the same 39% value when evaluated by a behavioral assay (as in **B**). **(D)** A possible workflow diagram—inspired from Damiano et al. (2011)—for the evaluation of SC life-likeness based on the measure of "organizational relevance." The organizational relevance is evaluated after a first screening for the "imitative relevance." The question is how to devise a method to map the dynamical processes occurring within the artificial system (i.e., the synthetic mechanisms that generate the behavior under investigation) with a benchmark organization, stemming from a reference theory. This study suggests that systemic theories of life, such as autopoiesis, (M,R)-systems and relational biology, or chemoton are best candidate for it, but it left to future investigation the screening and the selection of the best criteria for their application.

2.3. Limitations of the Imitation Game

The Turing test for AI focuses only on the "recognizability," by a biological system, of some phenomenology (behaviors) produced by an artificial system, without referring to the biological plausibility of the underlying mechanisms operating within the artificial system. Shannon and McCarthy (1956), and later Block (1981), among others, presented arguments against the purely behavioral approach expressed by the Turing test. In particular it has been claimed that a further *non-behavioral* criterion, referring to the manner in which the artificial system works, is needed for a proper qualification of the artificial agent. In the mentioned SCs/*V. fischeri* communication example, the life-likeness of the artificial system is inferred indirectly, by measuring the response achieved in the biological system. The manner in which the artificial system works is not considered, leading to the inconsistencies mentioned above (**Figure 1C**).

The definition of SC complexity and life-likeness therefore require further elaborations, and the previous considerations suggest to focus on non-behavioral criteria. The original intention of the Turing test was to assess AI without referring to any theory about intelligence functioning. However, assessing life at the minimal unicellular level can be soundly based on several descriptions, disclosed chemical mechanisms, and systemic/synthetic theories. The latter, in particular, can provide a useful starting block for discussing SC life-likeness beyond behavioral imitation.

2.4. The Organizational Perspective

As soon as SC behavior is deliberately skipped, SC life-likeness can be advantageously assessed by their *organization*, despite the fact that SCs are material entities existing in the chemical domain and discussing their *structure* is obviously tempting[1]. SC organization can be conveniently discussed within the Rosen relational theory of (Metabolism,Repair)-systems (Rosen, 1958, 1991; Letelier et al., 2003), the Maturana-Varela autopoiesis (Varela et al., 1974; Maturana and Varela, 1980; Luisi, 2003), and the Gánti's chemoton (Gánti, 2003). These systemic theories focus on the organization, depicting the relations between the elements and the transformations of components within a system, and describe life as a property emerging not from specific elements, but from a peculiar type of organization (Bich and Damiano, 2012).

How to move, then, from mere behavior, on which the Turing test is focused, to the patterns of generative mechanisms? With this regard, the *sciences of the artificial* (Simon, 1996; Cordeschi, 2002) can help significantly. These include important fields like cybernetics, robotics, artificial intelligence, artificial life, and most recently, synthetic biology, all having in common the construction of artifacts based on the so-called synthetic (or constructive) approach. For example, within the context of the epistemology of the sciences of the artificial, the relevance of artificial models of biological and cognitive processes has been recently discussed in an innovative manner (Damiano et al., 2011), based on an organizational approach to the scientific characterization of life that re-proposes the "life = cognition" thesis (Bich and Damiano, 2012)[2]. In Damiano et al.

[1]Here we refer to the molecular structure of the elements constituting the SCs. In principle, different chemicals can lead to networks of transformations that map into the known biological ones, when causal relationships are considered [an example is the "Rasmussen protocell" (Rasmussen et al., 2004)]. Structure is instead relevant in topological sense, i.e., as a determinant of the unity of SCs in space and time, and must emerge from the physical self-organization of the system's components. For example, non-self-bounded systems should be excluded.

[2]We refer to the autopoietic description of the biological domain introduced by Maturana and Varela (1980), whose innovative definition of life will be discussed later in this paragraph. Autopoiesis rejects the classical analytical approach organizing scientific definitions of life as lists of properties considered essential for living systems. Recognizing that any such list may always be susceptible to criticism and revision, and be rejected by part of the scientific community, autopoiesis proposes a synthetic approach to the definition of life. This consists in directing the definitional effort not toward the identification of the essential properties of living beings, but toward the determination of an organizational mechanism capable of generating, from a plurality of components, the minimal living cell and, on this basis, all living phenomenology as we know it (Bich and Damiano, 2007). This mechanism, which Maturana and Varela characterized as the "autopoietic organization," is considered by them to be capable to create all biological processes and features, including the properties typically listed in analytical definitions of life – autonomy, production of the molecular components (including the replication of information-carrying molecules), growth-and-division (reproduction), sensing the environment, etc. In this sense, from the autopoietic point of view, all the properties we can scientifically ascribe to living systems are grounded in this specific organizational mechanism (the system's autopoiesis), which generates living processes as cognitive processes, and is responsible for the emergence of intelligence in higher-level biological domains (Bich and Damiano, 2012).

(2011) the authors argues that hardware, software and wetware artifacts can have different types of relevance, defined by the combination of two criteria, respectively assessing (a) an artificial system's capability of recreating the phenomenology typical of the natural process under study, and (b) the system's capability of reproducing, in its organization, the organizational mechanisms that in nature generate that phenomenology. In this perspective, an artifact designed for passing the AI Turing test by merely imitating a certain cognitive behavior is recognized to have a purely imitative relevance, and, in this sense, is considered less relevant, as a synthetic model of the target cognitive behavior, than an artifact based on plausible organizational mechanisms generating that behavior.

When translated in the current discussion, the above-mentioned insightful argument leads to an interesting direction. In particular, a distinction can be made between mere "imitative" relevance (low life-likeness) and deeper "organizational" relevance (high life-likeness) according to the scheme shown in **Figure 1D**. The criterion of organizational relevance (stratified after the imitative one) can measure the SC life-likeness if a suitable "benchmark," directly stemming from a theory of reference, is available.

Autopoiesis has inspired SCs research from its beginning (Luisi and Varela, 1989; Walde et al., 1994; Luisi, 2003) and thus it appears a suitable theory of reference for the conceptual and operational definition of organizational criteria in SC research. Let us recall that the autopoietic organization is defined as:

> (…) a network of processes of production (transformation and destruction) of components that produces the components which (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute (…) a concrete unity in the space in which they (the components) exist by specifying the topological domain of its realization as such a network (Maturana and Varela, 1980, p. 79).

A possible quantification of SC life-likeness can be based on the comparison between the actual SC organization and the minimal autopoietic organization. The latter, however, is not specified, there is nothing as a "standard autopoietic network" to refer to. This depends, ultimately, by the actual realization of the autopoietic system, and thus by the type of elements contributing to the autopoietic network, as well as by the foodstuff supplied by the environment[3]. Autopoietic systems can have in principle different realizations, provided that their autopoiesis and their self-bounding as unity is maintained. For example, if SCs are realized—as it often happens—with familiar biochemical molecules (DNA, RNA, proteins, lipids, …), their minimal complexity is determined (and constrained) by the known biochemistry, because the autopoietic network needs to generate all SC components.

[3]In a chemically "rich" environment, for example, autopoietic systems can be simpler than the one established in a chemically poor one (endosymbionts and parasites, living inside other cells, illustrates well this principle).

Let us imagine, then, being able to devise a proper and suitable benchmark autopoietic network, which can be realized, at least in principle, in the chemical domain, and endowed with the above-mentioned requirements. The next step would be the selection of criteria for mapping the actual SC organization with the autopoietic benchmark (such a question does not need to be answered here). Life-likeness would result from a comparison between the two organizations (SC vs. autopoietic). Consider for example the above-mentioned synthetic/natural bidirectional communication (Lentini et al., 2017). The benchmark autopoietic network would correspond to a hypothetic SC that produces all components of its organization (e.g., about 100 macromolecules), whereas it can rely on the full set of low MW compounds freely available in the environment. Such an organization consists of a self-bounding constraint and of hundreds of confined transformations, with the possibility of importing and exporting small molecules. Vice versa, the actual SC has been able to produce only a small number of the network components. Comparing these two networks leads estimate SC life-likeness as a very small number (<1%?), which seems more reasonable than the 39% estimated on the basis of the behavioral Turing test.

3. OUTLOOK AND OPEN QUESTIONS

It could be argued that referring to a theory (e.g., autopoiesis, or other systemic theories) to measure the SC life-likeness actually brings back the problem of definition of life into the question of determining the life-likeness: exactly against the general motivation behind the Turing test. While such a criticism should be seriously considered, there are two counter-arguments that can be taken into account. The first is that biological life, especially at the level of single cells, is more understood that human intelligence. Therefore, it is not unrealistic to approach life-likeness as discussed above. Secondly, referring to the systemic perspective is very attractive because theories, such as autopoiesis, (*M,R*)-systems, and chemoton do not attempt defining life by a list of attributes, but they provide high-level relational descriptions about how living systems function. Accordingly, this move will lead to fecund landscapes where the entire technical and conceptual toolbox of complex adaptive systems theory (complexity theory) can be applied to next steps of SC research, further enriching this rapidly evolving field.

AUTHOR CONTRIBUTIONS

PS conceived the research and wrote the manuscript. LD contributed to the epistemological aspects. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

Part of this study was presented at the 2017 Italian Workshop of Artificial Life and Evolutive Computation (WIVACE 2017, Venice, 19–21 September 2017); a more extensive discussion will be presented elsewhere (Damiano and Stano, in preparation).

REFERENCES

Adamala, K. P., Martin-Alarcon, D. A., Guthrie-Honea, K. R., and Boyden, E. S. (2017). Engineering genetic circuit interactions within and between synthetic minimal cells. *Nat. Chem*. 9, 431–439. doi: 10.1038/nchem.2644

Bich, L., and Damiano, L. (2007). Question 9: theoretical and artificial construction of the living: redefining the approach from an autopoietic point of view. *Orig. Life Evol. Biosph*. 37, 459–464. doi: 10.1007/s11084-007-9082-5

Bich, L., and Damiano, L. (2012). Life, autonomy and cognition: an organizational approach to the definition of the universal properties of life. *Orig. Life Evol. Biosph*. 42, 389–397. doi: 10.1007/s11084-012-9300-7

Block, N. (1981). Psychologism and behaviorism. *Philos. Rev*. XC, 5–43. doi: 10.2307/2184371

CHELLnet (2005–2009). Available online at: https://gow.epsrc.ukri.org/NGBOViewGrant.aspx?GrantRef=EP/D023343/1 (accessed August 12, 2020).

Cordeschi, R. (2002). *The Discovery of the Artificial. Behavior, Mind and Machines Before and Beyond Cybernetics. Studies in Cognitive Systems*. Dordrecht: Springer Netherlands.

Cronin, L., Krasnogor, N., Davis, B. G., Alexander, C., Robertson, N., Steinke, J. H. G., et al. (2006). The imitation game–a computational chemical approach to recognizing life. *Nat. Biotechnol*. 24, 1203–1206. doi: 10.1038/nbt1006-1203

Damiano, L., Hiolle, A., and Cañamero, L. (2011). "Grounding synthetic knowledge," in *Advances in Artificial Life, ECAL 2011*, eds T. Lenaerts, M. Giacobini, H. Bersini, P. Bourgine, M. Dorigo, and R. Doursat (Boston, MA: MIT Press), 200–207.

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. Mater. Interfaces* 10, 30137–30146. doi: 10.1021/acsami.8b10029

Dupin, A., and Simmel, F. C. (2019). Signalling and differentiation in emulsion-based multi-compartmentalized *in vitro* gene circuits. *Nat. Chem*. 11, 32–39. doi: 10.1038/s41557-018-0174-9

Forster, A. C., and Church, G. M. (2006). Towards synthesis of a minimal cell. *Mol. Syst. Biol*. 2:45. doi: 10.1038/msb4100090

Gánti, T. (2003). *The Principles of Life*. Oxford: Oxford University Press.

Gardner, P. M., Winzer, K., and Davis, B. G. (2009). Sugar synthesis in a protocellular model leads to a cell signalling response in bacteria. *Nat. Chem*. 1, 377–383. doi: 10.1038/nchem.296

Gil, R., Silva, F. J., Peretó, J., and Moya, A. (2004). Determination of the core of a minimal bacterial gene set. *Microbiol. Mol. Biol. Rev*. 68, 518–537. doi: 10.1128/MMBR.68.3.518-537.2004

Islas, S., Becerra, A., Luisi, P. L., and Lazcano, A. (2004). Comparative genomics and the gene complement of a minimal cell. *Orig. Life Evol. Biosph*. 34, 243–256. doi: 10.1023/B:ORIG.0000009844.90540.52

Joesaar, A., Yang, S., Bögels, B., van der Linden, A., Pieters, P., Kumar, B. V. V. S. P., et al. (2019). DNA-based communication in populations of synthetic protocells. *Nat. Nanotechnol*. 14, 369–378. doi: 10.1038/s41565-019-0399-9

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 Central Sci*. 3, 117–123. doi: 10.1021/acscentsci.6b00330

Lentini, R., Santero, S. P., Chizzolini, F., Cecchi, D., Fontana, J., Marchioretto, M., et al. (2014). Integrating artificial with natural cells to translate chemical messages that direct *E. coli* behaviour. *Nat. Commun*. 5:4012. doi: 10.1038/ncomms5012

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., and Varela, F. (1989). Self-replicating micelles–a chemical version of a minimal autopoietic system. *Orig. Life Evol. Biosph*. 19, 633–643. doi: 10.1007/BF01808123

Luisi, P. L. (2003). Autopoiesis: a review and a reappraisal. *Naturwissenschaften* 90, 49–59. doi: 10.1007/s00114-002-0389-9

Luisi, P. L., Ferri, F., and Stano, P. (2006). Approaches to semi-synthetic minimal cells: a review. *Naturwissenschaften* 93, 1–13. doi: 10.1007/s00114-005-0056-z

Maturana, H. R., and Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living, 1st Edn*. Dordrecht: D. Reidel Publishing Company.

Mushegian, A. R., and Koonin, E. V. (1996). A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. U.S.A*. 93, 10268–10273. doi: 10.1073/pnas.93.19.10268

Nakano, T., Eckford, A. W., and Haraguchi, T. (2013). *Molecular Communications*. Cambridge: Cambridge University Press.

Nakano, T., Moore, M., Enomoto, A., and Suda, T. (2011). "Molecular communication technology as a biological ICT," in *Biological Functions for Information and Communication Technologies, Number 320 in Studies in Computational Intelligence*, ed H. Sawai (Berlin; Heidelberg: Springer Berlin Heidelberg), 49–86. doi: 10.1007/978-3-642-15102-6_2

Niederholtmeyer, H., Chaggan, C., and Devaraj, N. K. (2018). Communication and quorum sensing in non-living mimics of eukaryotic cells. *Nat. Commun*. 9:5027. doi: 10.1038/s41467-018-07473-7

Rampioni, G., D'Angelo, F., Leoni, L., and Stano, P. (2019). Gene-expressing liposomes as synthetic cells for molecular communication studies. *Front. Bioeng. Biotechnol*. 7:1. doi: 10.3389/fbioe.2019.00001

Rampioni, G., D'Angelo, F., Messina, M., Zennaro, A., Kuruma, Y., Tofani, D., et al. (2018). Synthetic cells produce a quorum sensing chemical signal perceived by *Pseudomonas aeruginosa*. *Chem. Commun*. 54, 2090–2093. doi: 10.1039/C7CC09678J

Rasmussen, S., Chen, L. H., Deamer, D., Krakauer, D. C., Packard, N. H., Stadler, P. F., et al. (2004). Transitions from nonliving to living matter. *Science* 303, 963–965. doi: 10.1126/science.1093669

Rosen, R. (1958). A relational theory of biological systems. *Bull. Math. Biophys*. 20, 245–260. doi: 10.1007/BF02478302

Rosen, R. (1991). *Life Itself. A Comprehensive Inquiry Into the Nature, Origin, and Fabrication of Life. Complexity in Ecological Systems*. New York, NY: Columbia University Press.

Shannon, C. E., and McCarthy, J. (1956). *Automata Studies*. Princeton, NJ: Princeton University Press.

Simon, H. A. (1996). *The Sciences of the Artificial*. Cambridge MA: MIT Press.

Stano, P., Rampioni, G., Carrara, P., Damiano, L., Leoni, L., and Luisi, P. L. (2012). Semi-synthetic minimal cells as a tool for biochemical ICT. *Biosystems* 109, 24–34. doi: 10.1016/j.biosystems.2012.01.002

Tang, T.-Y. D., Cecchi, D., Fracasso, G., Accardi, D., Coutable-Pennarun, A., Mansy, S. S., et al. (2018). Gene-mediated chemical communication in synthetic protocell communities. *ACS Synth. Biol*. 7, 339–346. doi: 10.1021/acssynbio.7b00306

Turing, A. M. (1950). I. Computing machinery and intelligence. *Mind* LIX, 433–460. doi: 10.1093/mind/LIX.236.433

Urquhart, J. (2017). *Artificial Cells Pass the Turing Test*. Available online at: https://www.chemistryworld.com/news/synthetic-cells-pass-bacterial-turing-test/2500365.article (accessed August 12, 2020).

Varela, F., Maturana, H., 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

Walde, P., Wick, R., Fresta, M., Mangone, A., and Luisi, P. (1994). Autopoietic self-reproduction of fatty-acid vesicles. *J. Am. Chem. Soc*. 116, 11649–11654. doi: 10.1021/ja00105a004

Weiss, M., Frohnmayer, J. P., Benk, L. T., Haller, B., Janiesch, J.-W., Heitkamp, T., et al. (2018). Sequential bottom-up assembly of mechanically stabilized synthetic cells by microfluidics. *Nat. Mater*. 17, 89–96. doi: 10.1038/nmat5005

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: 28 October 2020
doi: 10.3389/fbioe.2020.584178

Modeling Cell-Free Protein Synthesis Systems—Approaches and Applications

*Jan Müller, Martin Siemann-Herzberg and Ralf Takors**

Institute of Biochemical Engineering, University of Stuttgart, Stuttgart, Germany

In vitro systems are ideal setups to investigate the basic principles of biochemical reactions and subsequently the bricks of life. Cell-free protein synthesis (CFPS) systems mimic the transcription and translation processes of whole cells in a controlled environment and allow the detailed study of single components and reaction networks. *In silico* studies of CFPS systems help us to understand interactions and to identify limitations and bottlenecks in those systems. Black-box models laid the foundation for understanding the production and degradation dynamics of macromolecule components such as mRNA, ribosomes, and proteins. Subsequently, more sophisticated models revealed shortages in steps such as translation initiation and tRNA supply and helped to partially overcome these limitations. Currently, the scope of CFPS modeling has broadened to various applications, ranging from the screening of kinetic parameters to the stochastic analysis of liposome-encapsulated CFPS systems and the assessment of energy supply properties in combination with flux balance analysis (FBA).

Keywords: *in vitro* protein synthesis, cell-free synthetic biology, mathematical model, *in silico*, ribosomes, transcription and translation, modeling

OPEN ACCESS

Edited by:
Simon J. Moore,
University of Kent, United Kingdom

Reviewed by:
Matthew Lux,
U.S. Army Edgewood Chemical Biological Center (ECBC), United States
Ashty Karim,
Northwestern University, United States
James T. MacDonald,
Imperial College London, United Kingdom

***Correspondence:**
Ralf Takors
ralf.takors@ibvt.uni-stuttgart.de;
takors@ibvt.uni-stuttgart.de

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

Received: *16 July 2020*
Accepted: *29 September 2020*
Published: *28 October 2020*

Citation:
Müller J, Siemann-Herzberg M and Takors R (2020) Modeling Cell-Free Protein Synthesis Systems—Approaches and Applications. Front. Bioeng. Biotechnol. 8:584178. doi: 10.3389/fbioe.2020.584178

INTRODUCTION

Cell-free protein synthesis (CFPS) technology has a long history in life sciences, which started with fundamental research on deducing the genetic code (Nirenberg and Matthaei, 1961). Over several decades, the system was developed stepwise into a polypeptide production machinery (Spirin et al., 1988). Since the early adaptations of the system for commercial use (Shimizu et al., 2005), an increasing number of applications have emerged in the market (e.g., PURExpress, PURE*frex*, PURE*frex*2.0, myTXTLkit). These are based either on synthetic transcription–translation systems with a well-defined composition or on crude cell extracts that contain a more complex component assembly. While the product titer and production volume of such systems have increased from a few microliters to hundreds of liters, several limitations of the system remain. To date, and even within the best commercial systems, the protein titer with CFPS systems is orders of magnitude lower than that of *in vivo* whole-cell production due to resource expense and reduced longevity (Carlson et al., 2012; Gregorio et al., 2019; Silverman et al., 2019). Here, purpose-driven modeling can be a crucial tool to push the boundaries forward and identify bottlenecks.

CFPS is an "open" experimental system that allows defined reaction setups, which is ideal for simulation approaches. It has emerged not only as a research tool for the processes of transcription and translation but as a biomanufacturing platform for rapidly prototyping production systems

in silico and *in vitro* (Laohakunakorn et al., 2020; Vilkhovoy et al., 2020). For a variety of transcriptional and translational components, kinetic parameters are known, allowing the study of their behavior. Yet, one of the most fundamental principles of modeling always restricts the approach; the accuracy of model predictions cannot exceed the granularity of the model itself. In other words, distinctions between experimental observations and simulations are likely to occur if model predictions, extrapolated data sets, or fundamental model structures do not reflect the real problem. Consequently, such discrepancies may motivate a more thorough study of the experimental problem. Hence, proper model design aims to reflect reality with sufficient granularity (e.g., should the maturation of a reporter be considered?), thereby building on a solid mixture of experimentally validated data supported by assumptions. In this regard, we provide a brief overview of the existing models for CFPS and related systems and how they are applied to specific cases. It must be stated that many models have been developed with different objectives regarding the system environment, model approach (deterministic, stochastic), and granularity. Therefore, the models are not categorized as "good" or "bad" but clustered and assessed with respect to their particular purpose. In contrast to the mini-review of Koch et al. (2018), which focuses on deterministic models for CFPS, we expand the scope to adjacent fields and highlight qualitative and quantitative model characteristics.

When developing a model, it is necessary to know the components that should be considered. A CFPS system typically consists of, at least, the core components of transcription and translation: a mRNA polymerase, ribosomes, translational factors, amino acyl-tRNA synthetases, amino acids, tRNAs, an energy regeneration system, and nucleotides (Shimizu et al., 2005). Additionally, the DNA substrate, the produced mRNA, and the product (in most cases presented here, GFP derivatives) must be considered. If a crude cell extract is used, the system becomes much more complex, as the concentration of many of the components is unknown. The modeling studies on CFPS presented in detail in section "**Development and Application of CFPS Models**" share the common goal of identifying key model parameters by parameter regression on experimental data. However, the complexity of the models differs. By trend, the models may be divided into four groups of different granularities (**Figure 1**):

"Minimum model": Minimal models, presented in section "Identifying Bottlenecks in CFPS Systems," take into account up to ten parameters or equations, mainly focusing on macromolecular components such as mRNA and DNA. They are the backbone for more detailed descriptions. Additionally, most of the genetic circuit models presented in section "**Extending the Scope of CFPS Modeling**" can be described as minimum models.

"Structured model": Medium-scale models that introduce structured descriptions of certain aspects of the transcription-translation network. In structured models, kinetic models such as Michaelis–Menten or Hill are implemented in combination with larger ODE systems of up to 100 equations.

"Unstructured model": Large-scale models are fine-grained. They are meant to describe the CFPS in a holistic way and comprise networks of several hundred reactions. These models typically use simple individual reactions without further structural elements in order to save computational costs (e.g., Matsuura et al., 2017).

"Hybrid model": A special case are hybrid models, connecting structured models to other networks such as metabolic networks, or unstructured parts of lumped elements. Intrinsically, the approach increases the model complexity and computational costs. However, it offers an in-depth analysis of CFPS.

DEVELOPMENT AND APPLICATION OF CFPS MODELS

Continuous models are typically used to simulate CFPS systems. They make use of ordinary differential equations (ODEs) and algebraic equations to dynamically describe model states. In a structured model, equations incorporate affinity constants and other parameters. To reduce complexity, models can be formulated following an unstructured black-box approach, considering only apparent kinetics (Bailey and Ollis, 1986). The differences between the model types are dynamic. Often, models are partly structured to focus on selected segments of the reaction network with particular interest. We call these approaches "hybrid models." The quality of CFPS mechanistic models relies heavily on the proper model structure and the correct identification of model parameters. Given the complex nature of CFPS, the precondition of independent datasets for parameter identification is challenging and may require repeated careful consideration for each regression analysis (Golightly and Wilkinson, 2005; Moore et al., 2018). Stochastic effects play only a minor role in most of the classical CFPS modeling approaches. In liposome or droplet-based CFPS, due to small reaction volumes, low numbers of molecules may cause rendering reactions between different molecules in stochastic events. Under such conditions, a description with a discrete and stochastic model is preferable (Gillespie, 1977; Frazier et al., 2009).

Identifying Bottlenecks in CFPS Systems

The most straightforward way of describing *in vitro* expression of GFP is to consider the macromolecular components, DNA, mRNA, and proteins in a black-box model (Karzbrun et al., 2011; Stögbauer et al., 2012; Chizzolini et al., 2017; Marshall and Noireaux, 2019). This allows fitting kinetic equations to the experimental results of GFP production, mRNA production, and mRNA degradation. RNA polymerase and ribosomes are considered as catalytic components. For simplification, we call such approaches "minimum models." Karzbrun et al. (2011) proposed a coarse-grained dynamic model consisting of four enzymatic reactions. Kinetic studies were performed for crude cytoplasmic extract from *Escherichia coli* to identify biosynthesis and degradation parameters. A similar granularity was chosen by Stögbauer et al. (2012) to simulate and analyze the results gathered with the PURExpress system. Here, the model neglects protein degradation but covers a broader experimental range, identifying the plateau phase when the translational system expires. A comparable ODE-based model was applied

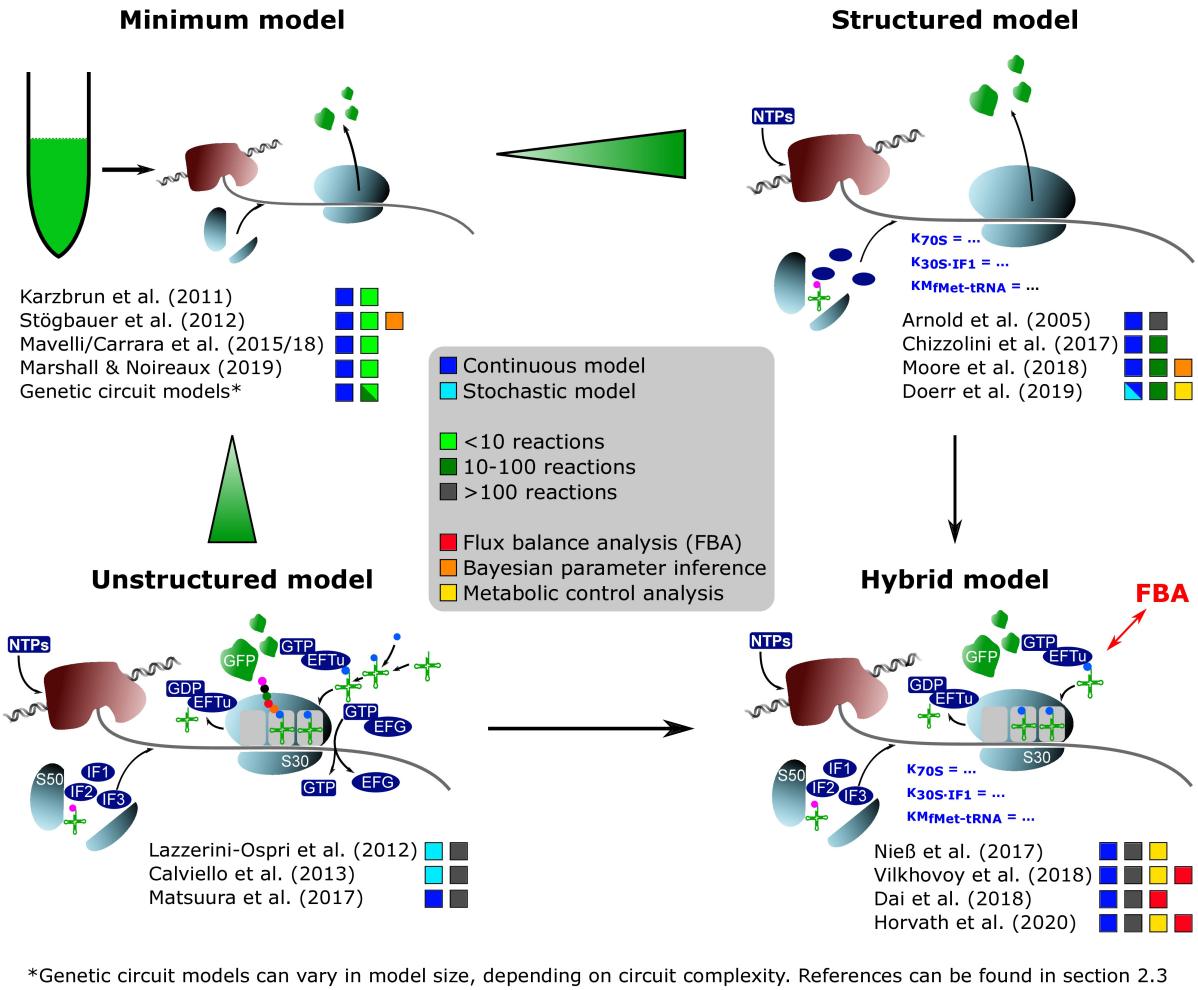


FIGURE 1 | On-trend categorization of CFPS models with respect to the defined model levels "minimum," "structured," "unstructured," and "hybrid." Color-coded squares indicate model classes, size, and particular features. The transition from "minimum" to "structured" considers the implementation of detailed kinetics. In contrast, the shifting from "minimum" to "unstructured" extends the reaction network and kinetic complexity. "Hybrid model" represents a tradeoff between "structured" and "unstructured" approaches. NTPs, nucleoside triphosphates; S50/S30, ribosomal subunits; EFTu, elongation factor Tu; EFG, elongation factor G; IF1/IF2/IF3, initiation factors; GFP, green fluorescent protein.

to a variety of regulatory elements (promotor strength) to identify limitations in the resources of the transcription–translation system (Marshall and Noireaux, 2019). Here, the commercially available "myTXTLkit" was used. Despite the application of different CFPS systems, all models revealed a saturation effect in GFP production under increased DNA template concentrations. Chizzolini et al. (2017) extended the minimal model to describe the expression of different fluorescence proteins under various regulatory elements. Here, limitations of current CFPS models were addressed, namely the specificity for only narrow experimental data sets, limited prediction capacity, and neglecting biophysical factors (e.g., RNA secondary structure).

With a model system of similar complexity, it was shown that limitations of CFPS may occur, which could not be mirrored by minimum models (Doerr et al., 2019). Using a comprehensive experimental data set of commercial *E. coli* CFPS, depletion of tRNAs and translation initiation were identified as limiting factors. By extending the minimum model to a structured description with additional terms for inactive mRNA states, it was possible to improve the prediction quality for the experimental data.

More fine-grained models were necessary to identify the challenging substrate limitations *in silico*. Such a model was introduced in our laboratory by Arnold et al. (2005) and recently renovated (Nieß et al., 2017). In this hybrid model, a simplified transcriptional model was connected to a detailed description of the translation process. The unique approach uses a ribosome flow model to simulate the movement of the ribosome along a one-dimensional discrete template (MacDonald and Gibbs, 1969; Heinrich and Rapoport, 1980). This approach enables a careful study of the influence of different components on the translation rate. The elongation factor Tu and tRNA concentration were identified as the most sensitive parameters hampering the

translation rate. As a key difference between *in vitro* and *in vivo* conditions, a control shift from the ternary complex to translation initiation was identified.

An equally complex system was developed to describe the synthesis of a short Met-Gly-Gly peptide in an *E. coli*-based *in vitro* system by incorporating 968 reactions and 241 components (Matsuura et al., 2017). The approach evaluated the stability of pseudo-steady states, revealing the temporal stability of metabolite clusters, their collapse, and re-merge, until a final steady state is reached. Interestingly, increasing tRNA supply also led to a slight increase in translation rates (observed as increased poly-peptide production), but the effect was much less dominant, as shown by Nieß et al. (2017).

Analysis and Prediction of Liposome-Encapsulated Protein Synthesis

A special case of *in vitro* protein synthesis is the encapsulation of CFPS components in liposomes. In this model, only a few stochastically distributed components may be balanced, creating different reaction conditions in the vesicle and outside the vesicle. The initial studies showed that GFP production kinetics strongly depend on liposome size and lipid composition (Sunami et al., 2010). In a first attempt to simulate protein synthesis inside liposomes, a medium-sized CFPS model considering 30 species and 106 reactions was connected to a stochastic model for encapsulation (Lazzerini-Ospri et al., 2012). Later, the model was extended to 280 species and 270 reactions, comprising a coupled transcription–translation model (Calviello et al., 2013). The approach allowed screening of GFP production with different start conditions either by looking at different liposome diameters or by considering different quantities of CFPS components inside the liposome. In agreement with continuum CFPS simulations, optimal DNA levels were identified for maximizing GFP formation. Oversaturation of the system with DNA decreased GFP yields. This finding reflected the enormous energy needs for the transcription process. Follow-up studies showed that some of the results could be achieved with a much less complicated model. Here, around 10 reactions were incorporated by lumping reactions for tRNA charging, transcription, translation, and energy regeneration. The simplified model described the behavior of the PURE system under 27 different compositions, rendering resource availability from standard conditions to limitation, remarkably well (Mavelli et al., 2015; Carrara et al., 2018).

Extending the Scope of CFPS Modeling

CFPS systems have emerged as ideal test beds for genetic circuits, allowing easier and faster prototyping than traditional in-cell engineering. Consequently, mathematical models to describe those systems have been developed (Niederholtmeyer et al., 2015; Takahashi et al., 2015). They cover a wide range of regulatory circuits: two-gene cascades (Siegal-Gaskins et al., 2014), sigma factor guided regulation (Adhikari et al., 2020), complex genetic ring oscillators (Niederholtmeyer et al., 2015), and experimentally verified RNA circuit controllers (Agrawal et al., 2018, 2019; Hu et al., 2018). The developed "minimum models" typically consist of three to ten ODEs and mass balances, considering mRNA, regulatory RNAs, or protein products as model species, and mass action, Hill, or Michaelis–Menten kinetics for regulatory descriptions. Even with coarse-grained models, the highly dynamic systems could be mirrored and predicted successfully. The model-guided circuit design significantly reduced development times.

The use of *in silico* models is not limited to well-defined model systems such as *E. coli* crude extract or commercial products. Moore et al. (2018) broadened their application to study the CFPS capacities of *Bacillus megaterium*, linking robotic liquid handling with a coarse-grained ODE model (26 parameters, 14 species, and 18 reactions) for the TX-TL system. Key kinetic parameters of the xylose-repressor system were approximated from DNA titration experiments. Simulations were performed using parameters identified by Bayesian parameter inference. Extending the model to describe the concurrent expression of two targets, plasmids carrying GFP and mCherry derivatives revealed competition for translational resources. In general, the reported translation elongation rates (between 0.10 and 0.02 aa s^{-1}) were slower than those reported for CFPS systems. The inefficient use of available energy accounted for the low performance. Another model approach for investigating resource competition in CFPS was formulated by Gyorgy and Murray (2016) with a minimal model for genetic circuits. Here, the authors could successfully quantify the burden of two targets expressed simultaneously on the resources of a CFPS system.

A constraint-based model to approximate energy and substrate supply from *E. coli* CFPS extract was presented by Varner and colleagues (Dai et al., 2018; Vilkhovoy et al., 2018; Horvath et al., 2020). They connected a simplified description of protein production (Allen and Palsson, 2003) and allosteric enzyme regulation (Wayman et al., 2015) with the metabolic network. Flux balance analysis (FBA) was applied to estimate the flux patterns of central carbon metabolism, amino acid biosynthesis, and energy metabolism using the objective function of maximizing the production rate of chloramphenicol acetyltransferase (CAT). Analysis of different amino acid supply scenarios *in silico* revealed inefficient energy yields of the experimental *in vivo* setup, most likely due to unfavorable side reactions. Similar scenarios may have also occurred in the experimental setup for *B. megaterium* described above.

DISCUSSION

A historical trend can be observed regarding the objectives of CFPS models. Early models (Arnold et al., 2005; Karzbrun et al., 2011; Stögbauer et al., 2012) are focused on the basic CFPS system using GFP as an experimental readout, beyond classical targets such as ß-galactosidase, chloramphenicol transferase, luciferase, or other likewise "easy to quantify" targets. In the past decade, an increasing number of diverging scientific branches have developed broadening the scope of model building and simulations. Yet, the GFP-based system is described in most detail and is the focus of current investigations. Currently, derivatives of the initial GFP are commonly used, such as its

TABLE 1 | Overview of the different granularities of CFPS models.

References	Features	No. of parameters	No. of species	No. of equations	k_{TX} [nt s^{-1}]	k_{TL} [aa s^{-1}]
Karzbrun et al. (2011)	Minimal model of the CFPS system	10	7	4	0.50	4.00
Stögbauer et al. (2012)	Refined minimal model of the CFPS system	8	5	4	2.20	0.03
Mavelli et al. (2015) and Carrara et al. (2018)	Simplified CFPS model for screening of different CFPS compositions	16	9	6	1.67	0.09
Marshall and Noireaux (2019)	Simulation of different transcription (promoter) and translation initiation (ribosome binding site) configurations	14	5	5	10.00	2.50
Chizzolini et al. (2017)	Study on different fluorescence protein targets, regulatory elements, and critical evaluation of model prediction	12	10	10	–	–
Moore et al. (2018)	Model description and kinetic parameter estimation for CFPS of non-model bacteria	26	14	18	8.13–11.47	0.09–0.11
Doerr et al. (2019)	Identification of translation initiation as bottleneck of CFPS, analysis of different commercial kits	13	10	10	–	–
Lazzerini-Ospri et al. (2012) and Calviello et al. (2013)	*In lipo* protein synthesis, stochastic distribution of components in liposomes	24–280	106–270	–	19.00	4.00
Matsuura et al. (2017)	Quasi-stationary state analysis of complex model networks	483	241	968	–	–
Nieß et al. (2017) (based on Arnold et al., 2005)	First detailed description of coupled transcription and translation model (Arnold). Comparison of *in vitro* and *in vivo* conditions, metabolic control analysis	>70	174 + no. of codons	>500	–	1.12
Dai et al. (2018); Vilkhovoy et al. (2018), and Horvath et al. (2020)	Coupling of CFPS to flux balance network of the central carbon metabolism, implementation of allosteric regulation	–	146	264	–	–

Each reference is listed with its unique features alongside the number of parameters, species, and equations incorporated in the model. If available, the transcriptional and translational rate constants k_{TX} and k_{TL} as readouts of the model are given. Models that share the same background are lumped together (note that genetic circuit models are not considered here).

"enhanced" and "super folder" variants (Pédelacq et al., 2006; Shin and Noireaux, 2010). The mRNA product is typically quantified with RNA aptamer reporters such as the malachite green RNA aptamer. The broadening of scientific approaches and the extension of cell-free genetic circuits will increase the need for easy and reliable reporter systems based on short nucleotide or peptide sequences (Wick et al., 2019).

When analyzing the different model granularities (**Figure 1**), major differences are observed. For coarse-grained models, compromises are made by assuming certain states of the model system by neglecting components or by lumping different metabolites (e.g., all amino acids) to one species. Fine-grained models consider these species in detail. System complexity has been increased from small systems with around 10 equations to large models with hundreds of reaction components (**Table 1**). Increasing complexity can offer the possibility to resolve bottlenecks by getting insights into reactions or reaction networks. Experimental access to all process elements is hardly possible, and only subsets of information are normally available, even for best investigated bacterial strains such as *E. coli*. As a result, complex models usually rely on multiple data resources covering different experiments (Matsuura et al., 2017; Nieß et al., 2017), whereas small models may be well identified by single experiments. Interestingly, it was shown that results gathered with complex systems can also be mimicked with reduced systems (e.g., Mavelli et al., 2015). Consequently, deciding a proper CFPS model structure should be driven by the questions to be answered, and should critically reflect the database for model identification.

The quality of CFPS models is checked by challenging model predictions with experimental observations. Typically, rates for transcription (k_{TX}), translation, and elongation (k_{TL}) are experimental readouts. However, the range of these parameters is broad (**Table 1**). k_{TX} has been reported from 0.5 (Karzbrun et al., 2011) to 19 nt s^{-1} (Calviello et al., 2013). k_{TL} ranged from 0.03 (Stögbauer et al., 2012) to 4.00 aa s^{-1} (Karzbrun et al., 2011; Calviello et al., 2013). The apparent differences may reflect the intrinsic problem of using relatively few experimental readouts to identify models of different complexities (Chizzolini et al., 2017). It has been shown that even simultaneously planned and performed CFPS experiments can lead to significant outcomes between different laboratory sites (Cole et al., 2019). As the modeling studies presented here are based on a wide range of commercial and homemade CFPS systems and extracts, this might explain the deviance of calculated parameters.

CONCLUSION

Currently, CFPS models can identify bottlenecks in the transcriptional and translational processes as well as infer kinetic parameters from model data. The consensus of most model predictions is the identification of the translational rather than the transcriptional process as one of the key targets for further developments in CFPS systems. Potential starting points are translation initiation, tRNA supply, and recycling. In most approaches, the modeled mechanisms of the translational process seem to be oversimplified. Inspiring approaches for *in vivo* translation have been published by Vieira et al. (2016) and Dykeman (2020) that could be adapted to *in vitro* descriptions. For many modeling purposes, hybrid models can be the ideal

tradeoff between complexity and acceptable computational costs. As CFPS systems and genetic circuits get more complex and consider multiple targets (RNAs/proteins), models that consider the joint burden on resources will come into focus (Gyorgy and Murray, 2016; Borkowski et al., 2018). A feedback loop between the model investigation and experimental setup has to be established. The works on genetic circuit models have proved that fast and easy prototyping is possible with CFPS. To unravel the key mechanisms for designing models, data from metabolomics and proteomics have to be integrated. Recent research addresses this need and offers a variety of datasets that could be harnessed by the CFPS modeling community (Garcia et al., 2018; Garenne et al., 2019; Miguez et al., 2019). This significantly increases the possibility to describe CFPS with an improved mechanistic resolution, up to a complete dynamic description of the CFPS system components. The development will open the door for a thorough application of tools of statistical systems analysis and metabolic control analysis to translate simulation results into system engineering advice.

AUTHOR CONTRIBUTIONS

JM reviewed the literature, designed the concept, wrote the manuscript, and prepared the figures. MS-H and RT co-edited and supervised the manuscript. All authors approved the manuscript for publication.

FUNDING

JM was supported by the UfIB Ph.D. program of "Bundesministerium für Bildung und Forschung – BMBF" (Grant No. 031B0725).

REFERENCES

Adhikari, A., Vilkhovoy, M., Vadhin, S., Lim, H. E., and Varner, J. D. (2020). Effective biophysical modeling of cell free transcription and translation processes. *BioRxiv* [Preprint]. doi: 10.1101/2020.02.25.964841

Agrawal, D. K., Marshall, R., Noireaux, V., and Sontag, E. D. (2019). In vitro implementation of robust gene regulation in a synthetic biomolecular integral controller. *Nat. Commun.* 10:5760. doi: 10.1038/s41467-019-13626-z

Agrawal, D. K., Tang, X., Westbrook, A., Marshall, R., Maxwell, C. S., Lucks, J., et al. (2018). Mathematical modeling of RNA-based architectures for closed loop control of gene expression. *ACS Synth. Biol.* 7, 1219–1228. doi: 10.1021/acssynbio.8b00040

Allen, T. E., and Palsson, B. (2003). Sequence-based analysis of metabolic demands for protein synthesis in prokaryotes. *J. Theor. Biol.* 220, 1–18. doi: 10.1006/jtbi.2003.3087

Arnold, S., Siemann-Herzberg, M., Schmid, J., and Reuss, M. (2005). Model-based inference of gene expression dynamics from sequence information. *Adv. Biochem. Eng. Biotechnol.* 100, 89–179. doi: 10.1007/b136414

Bailey, J. E., and Ollis, D. F. (1986). *Biochemical Engineering Fundamentals (2. ed)*. New York, NY: McGraw-Hill.

Borkowski, O., Bricio, C., Murgiano, M., Rothschild-Mancinelli, B., Stan, G. B., and Ellis, T. (2018). Cell-free prediction of protein expression costs for growing cells. *Nat. Commun.* 9:1457. doi: 10.1038/s41467-018-03970-x

Calviello, L., Stano, P., Mavelli, F., Luisi, P. L., and Marangoni, R. (2013). Quasi-cellular systems: stochastic simulation analysis at nanoscale range. *BMC Bioinformatics* 14:S7. doi: 10.1186/1471-2105-14-S7-S7

Carlson, E. D., Gan, R., Hodgman, C. E., and Jewett, M. C. (2012). Cell-free protein synthesis: applications come of age. *Biotechnol. Adv.* 30, 1185–1194. doi: 10.1016/j.biotechadv.2011.09.016

Carrara, P., Altamura, E., D'angelo, F., Mavelli, F., and Stano, P. (2018). Measurement and numerical modeling of cell-free protein synthesis: combinatorial block-variants of the pure system. *Data* 3, 1–12. doi: 10.3390/data3040041

Chizzolini, F., Forlin, M., Yeh Martín, N., Berloffa, G., Cecchi, D., and Mansy, S. S. (2017). Cell-free translation is more variable than transcription. *ACS Synth. Biol.* 6, 638–647. doi: 10.1021/acssynbio.6b00250

Cole, S. D., Beabout, K., Turner, K. B., Smith, Z. K., Funk, V. L., Harbaugh, S. V., et al. (2019). Quantification of interlaboratory cell-free protein synthesis variability. *ACS Synth. Biol.* 8, 2080–2091. doi: 10.1021/acssynbio.9b00178

Dai, D., Horvath, N., and Varner, J. (2018). Dynamic sequence specific constraint-based modeling of cell-free protein synthesis. *Processes* 6, 1–28. doi: 10.3390/pr6080132

Doerr, A., de Reus, E., van Nies, P., van der Haar, M., Wei, K., Kattan, J., et al. (2019). Modelling cell-free RNA and protein synthesis with minimal systems. *Phys. Biol.* 16:025001. doi: 10.1088/1478-3975/aaf33d

Dykeman, E. C. (2020). A stochastic model for simulating ribosome kinetics in vivo. *PLoS Comput. Biol.* 16:e1007618. doi: 10.1371/journal.pcbi.1007618

Frazier, J. M., Chushak, Y., and Foy, B. (2009). Stochastic simulation and analysis of biomolecular reaction networks. *BMC Syst. Biol.* 3:64. doi: 10.1186/1752-0509-3-64

Garcia, D. C., Mohr, B. P., Dovgan, J. T., Hurst, G. B., Standaert, R. F., and Doktycz, M. J. (2018). Elucidating the potential of crude cell extracts for producing pyruvate from glucose. *Synth. Biol.* 3, 1–9. doi: 10.1093/synbio/ysy006

Garenne, D., Beisel, C. L., and Noireaux, V. (2019). Characterization of the all-E. coli transcription-translation system myTXTL by mass spectrometry. *Rapid Commun. Mass Spectrom.* 33, 1036–1048. doi: 10.1002/rcm.8438

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81, 2340–2361. doi: 10.1021/j100540a008

Golightly, A., and Wilkinson, D. J. (2005). Bayesian inference for stochastic kinetic models using a diffusion approximation. *Biometrics* 61, 781–788. doi: 10.1111/j.1541-0420.2005.00345.x

Gregorio, N. E., Levine, M. Z., and Oza, J. P. (2019). A user's guide to cell-free protein synthesis. *Methods Protoc.* 2:24. doi: 10.3390/mps2010024

Gyorgy, A., and Murray, R. M. (2016). "Quantifying resource competition and its effects in the TX-TL system," in *Proceedings of the 2016 IEEE 55th Conference on Decision and Control, CDC 2016, 1(Cdc)*, Las Vegas, NV, 3363–3368.

Heinrich, R., and Rapoport, T. A. (1980). Mathematical modelling of translation of mRNA in eucaryotes; steady states, time-dependent processes and application to reticulocytest. *J. Theor. Biol.* 86, 279–313. doi: 10.1016/0022-5193(80)90008-9

Horvath, N., Vilkhovoy, M., Wayman, J. A., Calhoun, K., Swartz, J., and Varner, J. D. (2020). Toward a genome scale sequence specific dynamic model of cell-free protein synthesis in *Escherichia coli*. *Metab. Eng. Commun.* 10:e00113. doi: 10.1016/j.mec.2019.e00113

Hu, C. Y., Takahashi, M. K., Zhang, Y., and Lucks, J. B. (2018). Engineering a functional small RNA negative autoregulation network with model-guided design. *ACS Synth. Biol.* 7, 1507–1518. doi: 10.1021/acssynbio.7b00440

Karzbrun, E., Shin, J., Bar-Ziv, R. H., and Noireaux, V. (2011). Coarse-grained dynamics of protein synthesis in a cell-free system. *Phys. Rev. Lett.* 106, 1–4. doi: 10.1103/PhysRevLett.106.048104

Koch, M., Faulon, J. L., and Borkowski, O. (2018). Models for cell-free synthetic biology: make prototyping easier, better, and faster. *Front. Bioeng. Biotechnol.* 6:182. doi: 10.3389/fbioe.2018.00182

Laohakunakorn, N., Grasemann, L., Lavickova, B., Michielin, G., Shahein, A., Swank, Z., et al. (2020). Bottom-up construction of complex biomolecular systems with cell-free synthetic biology. *Front. Bioeng. Biotechnol.* 8:213. doi: 10.3389/fbioe.2020.00213

Lazzerini-Ospri, L., Stano, P., Luisi, P. L., and Marangoni, R. (2012). Characterization of the emergent properties of a synthetic quasi-cellular system. *BMC Bioinformatics* 13(Suppl. 4):S9. doi: 10.1186/1471-2105-13-S4-S9

MacDonald, C. T., and Gibbs, J. H. (1969). Concerning the kinetics of polypeptide synthesis on polyribosomes. *Biopolymers* 7, 707–725.

Marshall, R., and Noireaux, V. (2019). Quantitative modeling of transcription and translation of an all-E. coli cell-free system. *Sci. Rep.* 9, 1–12. doi: 10.1038/s41598-019-48468-8

Matsuura, T., Tanimura, N., Hosoda, K., Yomo, T., and Shimizu, Y. (2017). Reaction dynamics analysis of a reconstituted *Escherichia coli* protein translation system by computational modeling. *Proc. Natl. Acad. Sci. U.S.A.* 114, E1336–E1344. doi: 10.1073/pnas.1615351114

Mavelli, F., Marangoni, R., and Stano, P. (2015). A simple protein synthesis model for the pure system operation. *Bull. Math. Biol.* 77, 1185–1212. doi: 10.1007/s11538-015-0082-8

Miguez, A. M., McNerney, M. P., and Styczynski, M. P. (2019). Metabolic profiling of *Escherichia coli*-based cell-free expression systems for process optimization. *Ind. Eng. Chem. Res.* 58, 22472–22482. doi: 10.1021/acs.iecr.9b03565

Moore, S. J., MacDonald, J. T., Wienecke, S., Ishwarbhai, A., Tsipa, A., Aw, R., 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, 4340–4349. doi: 10.1073/pnas.1715806115

Niederholtmeyer, H., Sun, Z. Z., Hori, Y., Yeung, E., Verpoorte, A., Murray, R. M., et al. (2015). Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife* 4:e09771. doi: 10.7554/eLife.09771

Nieß, A., Failmezger, J., Kuschel, M., Siemann-Herzberg, M., and Takors, R. (2017). Experimentally validated model enables debottlenecking of in vitro protein synthesis and identifies a control shift under in vivo conditions. *ACS Synth. Biol.* 6, 1913–1921. doi: 10.1021/acssynbio.7b00117

Nirenberg, M. W., and Matthaei, J. H. (1961). The dependence of cell-free protein synthesis in *E. coli* upon naturally occuring or synthetic polyribonucleotides. *Biol. Chem.* 47, 1588–1602.

Pédelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* 24, 79–88. doi: 10.1038/nbt1172

Shimizu, Y., Kanamori, T., and Ueda, T. (2005). Protein synthesis by pure translation systems. *Methods* 36, 299–304. doi: 10.1016/j.ymeth.2005.04.006

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, 2–10. doi: 10.1186/1754-1611-4-8

Siegal-Gaskins, D., Tuza, Z. A., Kim, J., Noireaux, V., and Murray, R. M. (2014). Gene circuit performance characterization and resource usage in a cell-free "breadboard.". *ACS Synth. Biol.* 3, 416–425. doi: 10.1021/sb400203p

Silverman, A. D., Karim, A. S., and Jewett, M. C. (2019). Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* 21, 151–170. doi: 10.1038/s41576-019-0186-3

Spirin, A. S., Baranov, V. I., Ryabova, L. A., Ovodov, S. Y., and Alakhov, Y. B. (1988). A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* 242, 1162–1163.

Stögbauer, T., Windhager, L., Zimmer, R., and Rädler, J. O. (2012). Experiment and mathematical modeling of gene expression dynamics in a cell-free system. *Integr. Biol.* 4, 494–501. doi: 10.1039/c2ib00102k

Sunami, T., Hosoda, K., Suzuki, H., Matsuura, T., and Yomo, T. (2010). Cellular compartment model for exploring the effect of the lipidic membrane on the kinetics of encapsulated biochemical reactions. *Langmuir* 26, 8544–8551. doi: 10.1021/la904569m

Takahashi, M. K., Chappell, J., Hayes, C. A., Sun, Z. Z., Kim, J., Singhal, V., et al. (2015). Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. *ACS Synth. Biol.* 4, 503–515. doi: 10.1021/sb400206c

Vieira, J. P., Racle, J., and Hatzimanikatis, V. (2016). Analysis of translation elongation dynamics in the context of an *Escherichia coli* cell. *Biophys. J.* 110, 2120–2131. doi: 10.1016/j.bpj.2016.04.004

Vilkhovoy, M., Adhikari, A., Vadhin, S., and Varner, J. D. (2020). The evolution of cell free biomanufacturing. *Processes* 8, 1–19. doi: 10.3390/PR8060675

Vilkhovoy, M., Horvath, N., Shih, C. H., Wayman, J. A., Calhoun, K., Swartz, J., et al. (2018). Sequence specific modeling of *E. coli* cell-free protein synthesis. *ACS Synth. Biol.* 7, 1844–1857. doi: 10.1021/acssynbio.7b00465

Wayman, J. A., Sagar, A., and Varner, J. D. (2015). Dynamic modeling of cell-free biochemical networks using effective kinetic models. *Processes* 3, 138–160. doi: 10.3390/pr3010138

Wick, S., Walsh, D. I., Bobrow, J., Hamad-Schifferli, K., Kong, D. S., Thorsen, T., et al. (2019). PERSIA for direct fluorescence measurements of transcription, translation, and enzyme activity in cell-free systems. *ACS Synth. Biol.* 8, 1010–1025. doi: 10.1021/acssynbio.8b00450

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

ORIGINAL RESEARCH
published: 29 October 2020
doi: 10.3389/fbioe.2020.590341

Tuning the Cell-Free Protein Synthesis System for Biomanufacturing of Monomeric Human Filaggrin

*Jeehye Kim[1], Caroline E. Copeland[1], Kosuke Seki[2,3], Bastian Vögeli[2] and Yong-Chan Kwon[1,4]**

[1] Department of Biological and Agricultural Engineering, Louisiana State University, Baton Rouge, LA, United States, [2] Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, United States, [3] Chemistry of Life Processes Institute, Northwestern University, Evanston, IL, United States, [4] Louisiana State University Agricultural Center, Baton Rouge, LA, United States

The modern cell-free protein synthesis (CFPS) system is expanding the opportunity of cell-free biomanufacturing as a versatile platform for synthesizing various therapeutic proteins. However, synthesizing human protein in the bacterial CFPS system remains challenging due to the low expression level, protein misfolding, inactivity, and more. These challenges limit the use of a bacterial CFPS system for human therapeutic protein synthesis. In this study, we demonstrated the improved performance of a customized CFPS platform for human therapeutic protein production by investigating the factors that limit cell-free transcription–translation. The improvement of the CFPS platform has been made in three ways. First, the cell extract was prepared from the rare tRNA expressed host strain, and CFPS was performed with a codon-optimized gene for *Escherichia coli* codon usage bias. The soluble protein yield was 15.2 times greater with the rare tRNA overexpressing host strain as cell extract and codon-optimized gene in the CFPS system. Next, we identify and prioritize the critical biomanufacturing factors for highly active crude cell lysate for human protein synthesis. Lastly, we engineer the CFPS reaction conditions to enhance protein yield. In this model, the therapeutic protein filaggrin expression was significantly improved by up to 23-fold, presenting 28 $\pm$ 5 μM of soluble protein yield. The customized CFPS system for filaggrin biomanufacturing described here demonstrates the potential of the CFPS system to be adapted for studying therapeutic proteins.

Keywords: cell-free synthetic biology, cell-free protein synthesis, filaggrin, protein therapeutics, bacteria codon bias, *Escherichia coli* cell lysate

OPEN ACCESS

Edited by:
Tao Chen,
Tianjin University, China

Reviewed by:
Pasquale Stano,
University of Salento, Italy
Dong-Myung Kim,
Chungnam National University, South Korea

***Correspondence:**
Yong-Chan Kwon
yckwon@lsu.edu

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

Received: *01 August 2020*
Accepted: *05 October 2020*
Published: *29 October 2020*

Citation:
Kim J, Copeland CE, Seki K, Vögeli B and Kwon Y-C (2020) Tuning the Cell-Free Protein Synthesis System for Biomanufacturing of Monomeric Human Filaggrin. Front. Bioeng. Biotechnol. 8:590341. doi: 10.3389/fbioe.2020.590341

INTRODUCTION

Owing to the open nature of the cell-free biology, the cell-free protein synthesis (CFPS) system provides practicable opportunities to design and re-design the cellular processes outside the cell (Silverman et al., 2019). In the past decades, the CFPS system has received the spotlight as a versatile multipurpose toolkit of synthetic biology research because the core machinery of transcription and

translation can be precisely tuned in an open environment of the *in vitro* platform, which is hard to achieve in the living system (Lu, 2017). The distinctive perspective of the CFPS system brings the technological advantages over traditional fermentation methods such as decoupling of protein synthesis from cell growth and reproduction, which allows engineers to solely utilize the catalyst from cells in an *in vitro* platform (Perez et al., 2016) and in artificial cell-like systems (Stano, 2019), enabling the study of protein synthesis process by examining the role of the supplements such as chaperones, elongation factors, ribosomes in a cell-free system (Li et al., 2014), and changing a paradigm of traditional logistic and storage practice by transforming the wet system to the dry format (Pardee et al., 2016; Stark et al., 2019; Wilding et al., 2019). In addition, continued efforts to develop a scalable CFPS system accelerate technology transfer and industrial adaptation (Zawada et al., 2011; Yin et al., 2012). Hence, the CFPS system takes its strong position in the race for the next-generation biomanufacturing platform as well as opens a new avenue for the precise synthesis of protein therapeutics.

Since protein therapeutics was first introduced a few decades ago, it possesses a substantial position in a modern pharmacotherapy market (Ram et al., 2016). The eukaryotic cell system has been widely accepted for protein therapeutics production because of supporting various post-translational modifications (Sanchez-Garcia et al., 2016). In contrast, the bacterial system, which has about 30% of biopharmaceuticals production, has not been considered a practical platform due to the lack of inherent modification capability (Chen, 2012). However, recent advances in the bacterial CFPS systems have overcome this technological barrier and equipped the synthetic capability of the protein therapeutics. i) The modern CFPS system has opened the opportunity of cell-free biomanufacturing of vaccines and personalized protein therapeutics by enabling cell-free protein glycosylation (Guarino and DeLisa, 2012; Jaroentomeechai et al., 2018; Kightlinger et al., 2018, 2020). ii) The synthesis of membrane proteins, which typically are a target of many drugs in clinical studies, has also been improved by the addition of membrane mimics such as peptide surfactants (Wang et al., 2011), membrane fragments (Shinoda et al., 2016), and liposomes (Matthies et al., 2011). iii) The on-demand vaccine productions have been achieved in the CFPS system as well. The CFPS platform was used in a portable medicines-on-demand system for a glycosylated protein vaccine (Adiga et al., 2018) and bioconjugate vaccine products (Stark et al., 2019). iv) The development and production of antibodies and antibody fragments by utilizing the CFPS system are available (Yin et al., 2012). High-throughput expression of IgG has achieved by disulfide bond isomerase (DsbC) addition, adjustment of redox potential (oxidized and reduced glutathione), chaperone and cofactor addition, and the tuning of the incubation time (Murakami et al., 2019). Also, the cytosol-penetrating antibodies have been synthesized in the CFPS system (Min et al., 2016). v) There have also been improvements in the CFPS systems devoted to the cancer therapeutic proteins such as onconase (Salehi et al., 2016) and crisantaspase (Hunt et al., 2019), demonstrating the CFPS platform to be advanced and on-demand technology for future cancer therapeutics.

Filaggrin (FLG) is a filament associated protein monomer that plays a crucial role in keratinization and maintaining the skin barrier function. The reduction of FLG in human epidermal skin increases the risk for asthma, immune sensitization, and other severe skin disorders such as atopic dermatitis and ichthyosis vulgaris (Cabanillas and Novak, 2016). In the human epidermis, the monomeric FLG is processed by FLG proteolysis from a large precursor profilaggrin. The skin aspartic protease (SASPase) is expressed exclusively in the stratum granulosum layer and cleaves the linker sequence of profilaggrin, leading to profilaggrin-to-FLG processing (Matsui et al., 2011). The free monomeric FLG protein is then cross-linked to keratin filaments by transglutaminases to form intracellular filaments that contribute to the compaction and mechanical strength of the cells (Cabanillas and Novak, 2016) and subsequently deiminated by peptidylarginine deiminases (PADs) 1 and 3 (O'Regan et al., 2008). During the deamination (namely, citrullination), the arginine residue in FLG is converted into citrulline, which drives a drastic charge loss and FLG dissociation from keratin filaments to further metabolism (Nachat et al., 2005). In the outer layers of the stratum corneum (SC), FLG monomers are degraded by proteases such as caspase-14, calpain-1, and bleomycin hydrolase into free amino acids (Egawa and Kabashima, 2018). These FLG breakdown products constitute the natural moisturizing factor (NMF) with the function of hydration or UV protection (Thyssen and Kezic, 2014).

The cell-free synthesized monomeric FLG can be used to restore impaired skin barrier function. There have been studies associated with FLG in regulating skin disorders by using gene therapy and protein therapeutics, including gene therapy targeting the upregulation of the FLG expression in the human immortalized keratinocyte cell line and mice for remedying atopic dermatitis (Otsuka et al., 2014). Another therapeutics study is a direct protein supplement. This approach required the precision delivery of monomeric FLG protein across the SC barrier, so that the reduced size of the FLG monomer was preferentially used rather than large profilaggrin for the therapy (Stout et al., 2014). Although the study by Stout et al. has shown the FLG synthesis using *Escherichia coli* host cell, the problem of lower production yield remains to be overcome to move on to future biomanufacturing of protein therapeutics. The overall protein synthesis rate can be improved by using an alternative *in vitro* biomanufacturing strategy with systemic optimization to set up the optimal production condition for the target protein of interest. Another perspective about FLG that makes this work meaningful, other than considering it just as a treatment of skin disorder, is that FLG is one of the many types of tandem repeat proteins that have a greater chance of mutation, which can render them dysfunctional. Due to their repetitive manner, the tandem repeats have higher mutation rates relative to other genomic loci, which contribute to generate variations that differ in the number of repeated units (Verstrepen et al., 2005). As the variable numbers of tandem repeats (VNTRs) mediate phenotypic plasticity and various polygenic disorders, a better understanding of tandem repeat polymorphisms (TRPs) and disease-associated mutations involving repeat instability needs to be accurately studied

(Hannan, 2018). For example, there are three most common-sized human FLG genes encoding 10, 11, or 12 FLG tandem repeats due to the intragenic copy number variation in the gene (Brown and McLean, 2012). Synthesizing FLG repeat variants along with loss-of-function mutations and documenting the heterogeneity of amino acids in repeats will provide an insight on tandem repeat instability and disease-associated mutation. Many tandem repeats (including FLG) are present in coding and regulatory regions of the human genomes, and the mutations in the repeated units are often associated with many genetic diseases (Duitama et al., 2014). Besides, repetitive DNA sequences are common, consisting of almost half of the human genome (Jelinek et al., 1980). Hence, synthesizing and characterizing monomeric FLG will provide a better understanding of repeat variants associated genetics and pathogenesis of a range of human genetic diseases.

Here, in this study, we demonstrated that high-yield FLG synthesis (1.0 ± 0.2 mg/ml or equivalently 28 ± 6 μM) by streamlining the CFPS optimization process. First, we have chosen the eighth repeat unit of human profilaggrin (GenBank ID 2312) for CFPS as the monomeric FLG repeat does not possess a loss-of-function mutation site in the map of the profilaggrin gene reported previously, allowing a starting point to be set to study the variations of FLG subunit heterogeneity and mutations (Sandilands et al., 2007). The coding sequence [wild type (WT)] was then modified for the *E. coli* codon preference (codon optimization). The WT DNA fragment and *E. coli* codon-optimized DNA fragment were assembled into two separate plasmid backbones. Next, highly active crude cell extracts were prepared from BL21 Star (DE3) and Rosetta-gami B (DE3) pLacI *E. coli* strains to attain the highest FLG production rate. Finally, we assessed the CFPS reaction conditions not only for the synthesis of FLG but also for overlooking the potential for cell-free biomanufacturing of therapeutic proteins.

MATERIALS AND METHODS

Plasmids and Strains

The bacterial plasmid pETBlue-1 (Novagen, St. Louis, MO) and pJL1-sfGFP were used for protein expression vectors. *Escherichia coli* strain Subcloning Efficiency™ DH5α [Genotype F^- Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) U169 *rec*A1 *end*A1 *hsd*R17(r_K^-, m_K^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ-] (Invitrogen, Waltham, MA) was used for subcloning. *E. coli* strains BL21 Star (DE3) [Genotype F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm rne131* (DE3)] (Invitrogen, Waltham, MA) and Rosetta-gami B (DE3) pLacI [Genotype F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm lacY1 ahpC* (DE3) *gor522*:Tn*10 trxB* pLacIRARE (Cam^R, Kan^R, Tet^R)] (Novagen, St. Louis, MO) were used for protein expression.

Growth Media

E. coli cells were grown in Luria–Bertani (LB) media or 2xYTPG media (16 g/L of tryptone, 10 g/L of yeast extract, 5 g/L of sodium chloride, 7 g/L of potassium phosphate dibasic, 3 g/L of potassium phosphate monobasic, pH 7.2, and 0.1 M of glucose) for cell biomass and protein expression.

Preparation of Expression Vector

The human eighth FLG repeat unit (324 amino acids) was selected for this study (**Figure 1A**). The WT nucleotide sequence was optimized for *E. coli* codon usage bias (strain K12) using a codon optimization tool (Integrated DNA Technologies) (**Supplementary Figure 1**). The WT and *E. coli* codon-optimized FLG coding sequences (so-called "Codon Opt" further on) were synthesized (Integrated DNA Technologies) and cloned into plasmids pJL1 and pETBlue-1 by Gibson Assembly (Gibson et al., 2009) for CFPS and *in vivo* protein synthesis, respectively. 6xhistidine tag was added to the C-terminal end of FLG during PCR. *E. coli* DH5α competent cells were used for the cloning host. The sequences have confirmed by DNA Sanger-Sequencing using the 3130xl Genetic Analyzer (Applied Biosystems). The recombinant plasmids were isolated by plasmid maxiprep kit (Invitrogen, Waltham, MA). A schematic cloning workflow is described in the **Supplementary Figure 2**.

Preparation of Cell Extracts

Cell Culture and Harvest

E. coli cell extract was prepared as described previously (**Figure 1B**; Kwon and Jewett, 2015; Kim et al., 2019). *E. coli* BL21 Star (DE3) and Rosett-gami B (DE3) pLacI were grown in 2.5-L baffled Tunair shake flasks (IBI Scientific, Peosta, IA) at 37°C with vigorous shaking at 250 rpm. LB and 2xYTPG media were used for overnight seed culture (20 ml) and main culture (1 L), respectively. Cell cultures were monitored by measuring optical density at wavelength 600 nm (OD_{600}) by UV-Vis spectrophotometer (Genesys 6, Thermo Fisher Scientific, Waltham, WA). The cultured cells were induced by 1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) at OD_{600} 0.6 to regulate T7 RNA polymerase. The cultured cells were collected two times by centrifugation [5,000 relative centrifugal force (RCF) at 4°C for 15 min] at different phases: 0.5 L of cells was harvested at the exponential growth phase (OD_{600} 2.0), and the remaining cells were grown until they reached the stationary growth phase (OD_{600} 2.8–3.0) and harvested. Cell pellets were washed three times in Buffer A, weighed, flash-frozen in liquid nitrogen, and stored at −80°C freezer until use.

Cell Lysis, Runoff, and Cell Extract Clarification

The cell pellets were thawed on ice and resuspended in 1.3 ml of Buffer A per 1 g of wet cell mass. The resuspended cells were then transferred to clean microtube in ice for sonication (Qsonica Q125, Newtown, CT). The sonicator (model Q125) was set at 20-kHz frequency and 50% amplitude with 1/8" diameter probe. To lyse the cells, the pulse was on for 10 s and off for 10 s to minimize overheat protein degradation. Total energy input (joules) was varied in three levels (270, 537, and 1,074 J for 1 ml of cell suspension); 3 μl of 1 M dithiothreitol (DTT) was added to 1 ml of cell lysate, and the microtube was gently inverted several times. The cell lysates were then centrifuged at 12,000 RCF 4°C for 10 min. The supernatant (clarified cell lysate) was transferred to new microtubes. To compare the extract performance between non-runoff reaction

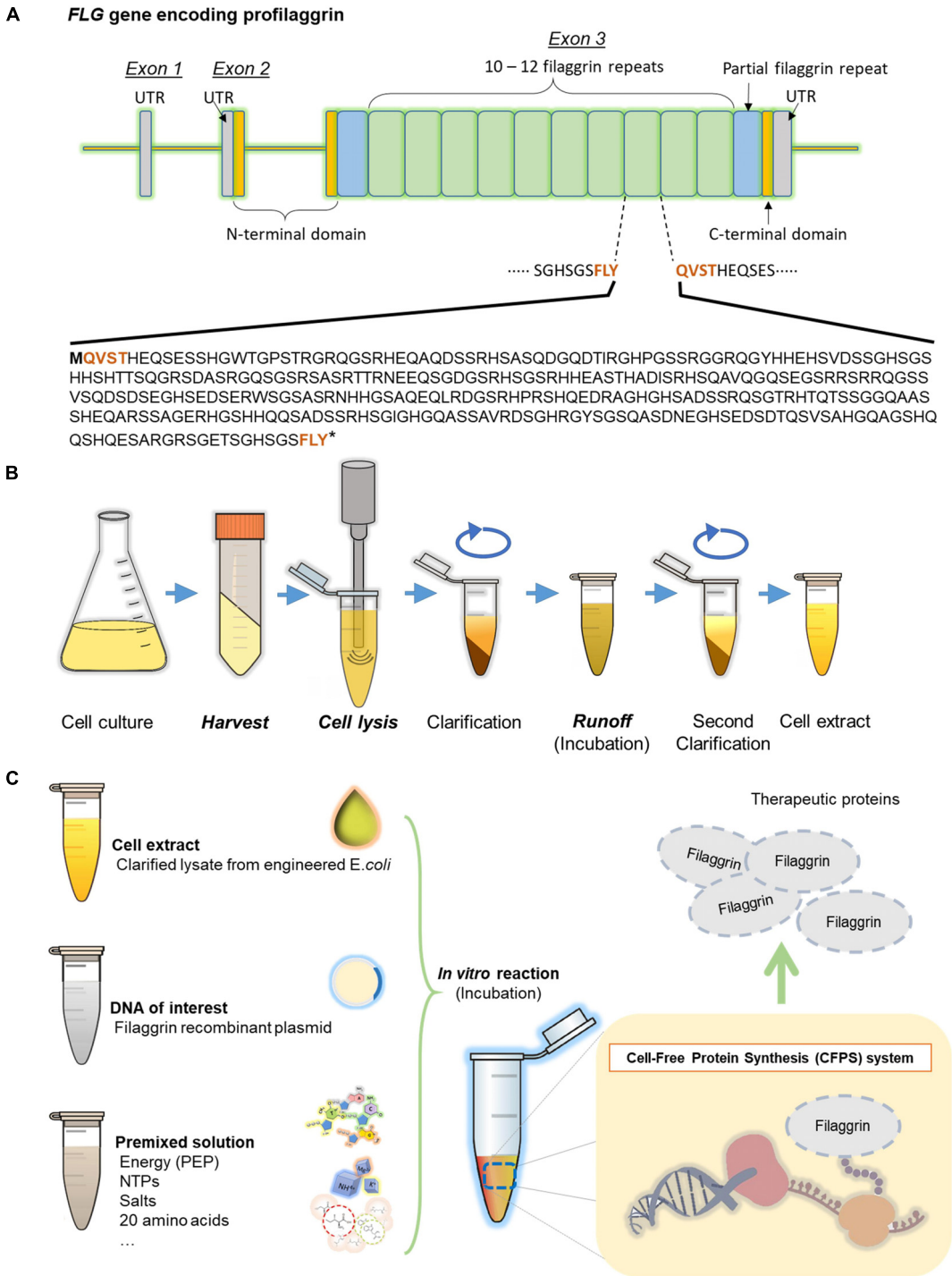


FIGURE 1 | Filaggrin (FLG) gene to protein. **(A)** Human filaggrin tandem repeat sequences in the profilaggrin gene. The gene comprises three exons and two introns. Functional filaggrin repeat units (324 amino acids long, 10 to 12 repeats) are encoded in exon 3. The cleavage site (FLYQVST) is indicated in bold brown. **(B)** Cell extract preparation procedure. **(C)** The cell-free protein synthesis system.

and runoff reaction, half of the clarified cell lysate (cell extract) was aliquoted in a small volume (non-runoff reaction), and the second half was incubated at 37°C for 1 h (runoff reaction) and clarified by centrifugation (12,000 RCF at 4°C for 10 min). The supernatant was transferred to new tubes and aliquoted in a small volume. The aliquoted cell extracts were flash-frozen in liquid nitrogen and stored at −80°C freezer until use.

Cell-Free Protein Synthesis

The CFPS was carried out for FLG expression as represented in **Figure 1C**. The standard 15 μl of CFPS reaction contains 4 μl of cell extract in addition to 12 mM of magnesium glutamate; 10 mM of ammonium glutamate; 130 mM of potassium glutamate; 1.2 mM of ATP; 0.85 mM each of GTP, UTP, and CTP; 34 μg/ml of L-5-formyl-5,6,7,8-tetrahydrofolic acid (folinic acid); 170 μg/ml of *E. coli* total tRNA, 57 mM of HEPES buffer (pH 7.2); 0.4 mM of nicotinamide adenine dinucleotide (NAD); 0.27 mM of coenzyme A; 4 mM of sodium oxalate; 1 mM of putrescine; 1.5 mM of spermidine; 2 mM of each of 20 amino acids; 33 mM of phosphoenolpyruvate (PEP); and 13.3 μg/ml of DNA template (plasmid). For the radioactivity measurement of cell-free synthesized protein, 10 μM of L-[U-^{14}C]-leucine [300 mCi (11.1 GBq)/mmole, GE Life Sciences, Marlborough, MA] was added to the standard CFPS reaction. *E. coli* cell extract contains endogenous T7 RNA polymerase, which was synthesized from the IPTG-induced cells. The CFPS reactions were carried out at 30°C for 20 h.

In vivo Protein Expression

The recombinant plasmid pJL1-FLG^{His} was transformed into Rosetta-gami B (DE3) pLacI [reactive oxygen species (ROS)] for *in vivo* FLG expression study (ROS:pJL1-FLG^{His}). WT ROS and ROS:pJL1-FLG^{His} were grown in two 300-ml Tunair flasks with 100 ml of 2xYTPG media at 37°C with the vigorous shaking at 250 rpm. ROS was induced by 1 mM of IPTG when the cells' optical density reached at 0.8 to inactivate lac repressor. Uninduced cell culture was prepared separately. The cultured cells were harvested by centrifugation. In order to lyse the cells, resuspended cells were disrupted by sonication (energy input 537 J for 1 ml of resuspended cells). The lysate was centrifuged at 14,000 RCF for 10 min to isolate a soluble fraction. The total cell protein and a soluble fraction were mixed with sample loading dye and denaturing solution (DTT), denatured by heating at 80°C for 5 min, and stored at −20°C until sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis.

Protein Analysis

Trichloroacetic Acid-Insoluble Radioactivity Calculation

The amounts of cell-free synthesized proteins were quantified by measuring trichloroacetic acid (TCA)-insoluble radioactivity using a liquid scintillation counter (MicroBeta2, PerkinElmer, Waltham, MA). Radioactive L-[U-^{14}C]-leucine added to the CFPS reaction mixture along with the 20 amino acids and carried out the standard CFPS reaction at 30°C for 20 h. The CFPS reaction was quenched by adding an equal volume of 0.5 M of potassium hydroxide and incubated at 37°C for 20 min. The same amounts of samples were spotted on two separate fiberglass paper sheets (Filtermat A, PerkinElmer, Waltham, MA) and dried at room temperature for 30 min. One of the fiberglass paper sheets was subjected to washing three times with 5% TCA solution at 4°C for 10 min. The sheet was rinsed in 100% ethanol at room temperature for 10 min to remove non-precipitated samples and was dried at room temperature for 40 min. The meltable cocktail (MeltiLex A, PerkinElmer, Waltham, MA) was heated to 95°C, and then radioactivity was measured using the liquid scintillation counter. The total and soluble yields were determined as described previously (Swartz et al., 2004).

Fluorescence Intensity and Quantification

The fluorescence intensity of active sfGFP was measured by multi-well plate fluorometer (Synergy HTX, BioTek, Winooski, VT); 2 μl of cell-free synthesized sfGFP was diluted with 48 μl of diethylpyrocarbonate (DEPC)-treated water (Invitrogen, Waltham, MA) in a 96-well half-area black flat-bottom plate (Corning, Glendale, AZ). The excitation and emission wavelengths for the sfGFP fluorescence were 485 and 528 nm, respectively. The cell-free synthesized sfGFP was purified by affinity chromatography cartridge (Strep-Tactin resin) following the manufacturer's instruction (Qiagen, Germantown, MD). The soluble fraction of sfGFP was filtered by using a 0.2-μm syringe filter to remove remaining precipitants before passing through the column. Lysis buffer (50 mM of sodium phosphate monobasic, 300 mM of sodium chloride, pH 8.0) and elution buffer (50 mM of sodium phosphate monobasic, 300 mM of sodium chloride, 2.5 mM of desthiobiotin, and protease inhibitor) were filtered before use. The column was then washed three times with lysis buffer. Elution fractions were collected in the fresh tubes and concentrated first two fractions using an Amicon Ultra centrifugal filter unit [molecular weight cutoff (MWCO) (10 kDa)] (Millipore, Burlington, MA). The conversion factor was then calculated by Bradford assay using the purified sfGFP. The concentration of the sfGFP (μM) was determined by applying the conversion factor to the fluorescence intensity of the cell-free synthesized sfGFP.

Protein Electrophoresis

Five microliters of cell-free synthesized protein was mixed with 5 μl of sample buffer [NuPAGE LDS (4×), Invitrogen, Waltham, MA] and 10 μl of 200 mM of DTT and then mixed well. The samples were denatured by incubation at 80°C for 5 min before loading; 10 μl of heat-denatured protein samples was loaded on pre-casted gradient gel (NuPage 4-12% Bis-Tris Protein Gel, Invitrogen, Waltham, MA). The protein electrophoresis was carried out at 150 V for 75 min, and the proteins were visualized by Coomassie blue staining.

Filaggrin Purification

The cell-free synthesized FLG was isolated by immobilized-metal affinity chromatography (IMAC). The nitrilotriacetic acid (NTA) agarose matrices (Qiagen, Germantown, MD) were charged with Ni^{2+} and then washed with deionized water (Direct-Q system, Millipore, Burlington, MA) 10 times. All buffers contain the same composition of 50 mM of monosodium phosphate and 300 mM of sodium chloride with the different imidazole concentrations for each step of purification (10 mM for binding, 20 mM for wash, and 250 mM for elution). The CFPS reaction was transferred to a centrifugal filter unit [Amicon Ultra, MWCO (10 kDa)] and centrifuged at 14,000 × *g* RCF for 15 min. The filtrate was discarded, and an equal volume of lysis buffer containing protease inhibitors (Thermo Fisher, Waltham, MA) was added to

the concentrate and subjected to repeated centrifuge to remove salts and other components in CFPS that possibly interfere with the histidine binding. The tagged proteins were then loaded onto the Ni-NTA matrices and immobilized, washed twice, and eluted by adding an elution buffer containing protease inhibitors. The imidazole in elute was removed by buffer exchange to Tris buffer (pH 8.0) containing protease inhibitors. The purified FLG was stored at −20°C until use.

Statistical Analysis

All effect estimation was presented at the 5% significance level. Analyses were conducted using Graphpad Prism 8.4.3 (GraphPad Software). For the parametric analysis of data from quantification of the synthesized protein, two-way ANOVA followed by the Dunnett test was used.

RESULTS AND DISCUSSION

Host Selection and Codon Optimization

Improving Protein Synthesis Using Rare tRNA Expressed Cell Extract

Since the transcription and translation apparatus is well conserved in the cell extract, host strain selection can revamp the overall protein synthesis rate in the CFPS system and likewise in *in vivo* recombinant protein synthesis platforms. Although the genetic code is universally shared across species, different codon usage biases strongly correlate with protein synthesis between the host strain and the origin of a gene of interest (Gingold and Pilpel, 2011). This codon usage impacts protein expression levels and folding efficiency during the translation (Quax et al., 2015). As the abundance of tRNA at the cellular level speeds up ribosomes to decode a codon (Novoa and de Pouplana, 2012), codon optimization and rare tRNA supplementation can improve transcription and translation efficiency without a shortage of tRNA availability. Therefore, choosing an appropriate *Escherichia coli* host that supports sufficient rare tRNA expression provides a decisive improvement in protein production yields, particularly in a prokaryotic heterologous gene expression system for a codon biased exotic gene such as FLG.

We first selected two *E. coli* strains BL21 Star (DE3) and Rosetta-gami B (DE3) pLacI (denoted as "BL21" and "ROS") as hosts to validate and compare initial FLG synthesis level. ROS is a BL21 strain derivative. It is more versatile in producing eukaryotic proteins, including therapeutic proteins, as it supports six-rare-tRNA translation as well as disulfide bond formation. Both *E. coli* strains were cultured and lysed into cell extracts for the CFPS reactions. The separate reactions were conducted at 30°C for 20 h using the cell extracts from BL21 and ROS. The CFPS reaction with ROS cell extract showed a significantly higher yield than the synthesis using BL21 ($p < 0.001$) in both total and soluble fractions (increased yields by 5.3-fold and 9.1-fold, respectively) (**Figure 2A**). There were no significant differences between the BL21 and ROS extracts for sfGFP synthesis, indicating that two *E. coli* extracts are comparable for producing non-codon biased genes such as sfGFP. The SDS-PAGE results confirmed an increase in FLG production (blue arrows) using the ROS cell extract (**Figures 2B,C**). These results indicate that the additional supplement of rare tRNA in the CFPS reaction can improve the production yield of human proteins, which are considered one category of the hard-to-express proteins due to the highly biased codon usage. In other words, the gene encoding a FLG monomer (a functional repeat unit) contains 59 biased codons of its total 324 codons, which are not favorable in *E. coli* translational machinery due to lacking tRNA corresponding to the six rarely used codons in *E. coli* such as AGG, AGA for arginine, GGA for glycine, AUA for isoleucine, CUA for leucine, and CCC for proline. In this work, ROS cell extracts supplied the rare tRNAs to the CFPS system by utilizing the pLacIRARE plasmid containing a set of rare tRNA genes.

Another implication of the lower FLG monomer synthesis result here is that the expression level of the FLG monomer is more likely dependent on the CFPS reaction condition compared

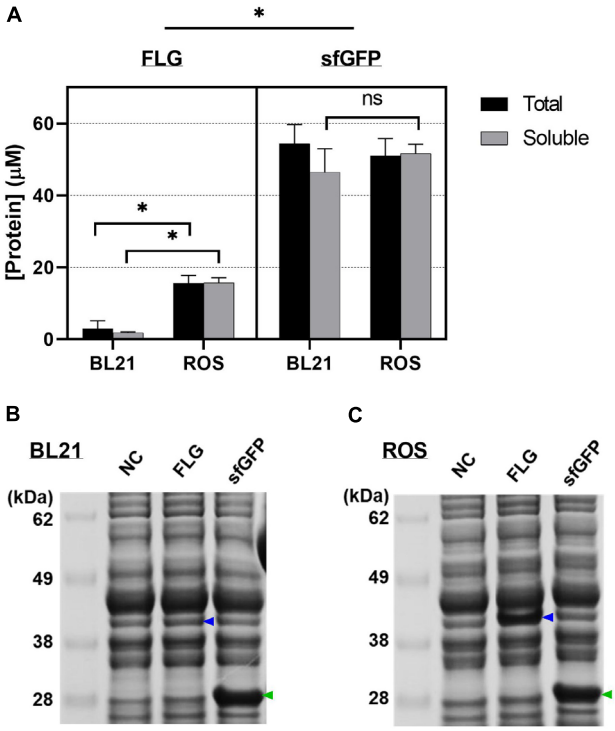


FIGURE 2 | Cell-free protein synthesis (CFPS) using cell extracts from different *Escherichia coli* strains. **(A)** The amount of the cell-free synthesized protein was quantified by radioisotope analysis. sfGFP was used for the protein synthesis indicator (positive control). Data represented as the mean ± SD ($N = 3$). Tukey's multiple comparisons test was performed with the two-way ANOVA. * indicates that the difference exists between the extracts from BL21 and reactive oxygen species (ROS) in terms of protein productivity ($p < 0.001$); ns indicates no significant difference between the two groups at the 0.05 level of significance. **(B)** The soluble portion of cell-free synthesized proteins using BL21 Star (DE3) cell extract. **(C)** The soluble portion of cell-free synthesized proteins using Rosetta-gami B (DE3) pLacI cell extract. NC represents a negative control that used no plasmid for the CFPS reaction. Filaggrin (FLG) and sfGFP represent that pJL1-FLG wild-type plasmid and pJL1-sfGFP plasmid were used for the CFPS reaction, respectively. The blue and green arrows indicate the cell-free synthesized FLG and sfGFP. All CFPS reactions were performed at 30°C for 20 h.

with the non-human-originated gene. Then, the protein synthesis yields would be improved by further optimizing the CFPS reaction condition to recreate a favorable environment for FLG monomer synthesis. On the basis of the assumption, we next modified the major contributors of the CFPS system systematically, as follows: DNA template, cell extract, and the CFPS reaction mixture.

Bypassing the Codon Usage via Codon Optimization

Codon optimization is an alternative method for addressing the biased codon usage in the CFPS system. When the canonical 20 amino acids are encoded within one to six synonymous codons, the frequencies of different codons vary across organisms (Gustafsson et al., 2004). Codon usage has been recognized as one of the most vital factors to regulate gene expression in the prokaryotic system (Lithwick and Margalit, 2003). Therefore, the complete reassignment strategy of human gene sequences for matching *E. coli* codon usage facilitates to maximize the chance of efficient protein expression.

To demonstrate this assumption, we synthesized the WT and codon-optimized FLG using IDT's Codon Optimization Tool (organism set for *E. coli* K12). DNA properties, in particular, guanine–cytosine (GC) contents (%), were not changed significantly (WT, 57.03%; and codon optimized, 58.21%). We then cloned both sequences into pJL1 and pETBlue1 vectors to obtain four recombinant plasmids (pJL1-FLG WT, pJL1-FLG Codon Opt, pETBlue1-FLG WT, and pETBlue1-Codon Opt) by Gibson Assembly. The coding sequences were placed between the T7 promoter and T7 terminator in the plasmid. sfGFP sequence in the plasmid pJL1-sfGFP was substituted to the FLG sequence. pETBlue-1 vector is a compatible expression vector for the *E. coli* host strain Rosetta-gami B (DE3) pLacI. The yields of cell-free synthesized proteins were determined by radioisotope analysis and compared head to head (**Figure 3**).

Coupled with BL21 cell extract, the codon-optimized gene in pJL1 vector (pJL1-FLG Codon Opt) notably improved protein expression level than the WT gene (pJL1-FLG WT) in both total and soluble fractions (increased yields by 5.2-fold and 7.4-fold, respectively) (**Figure 3A**) ($^*p < 0.001$). In contrast, the codon-optimized gene in the pET vector did not show improved protein expression. Interestingly, pJL1-FLG Codon Opt paired with ROS cell extract increased protein yield only by 1.5-fold (total protein) and 1.8-fold (soluble protein) than did pJL1-FLG WT (**Figure 3B**) ($^*p < 0.001$). Regarding the low expression level shown with the pET vector, extra investigation showed that IPTG (0 to 1.0 mM) during the CFPS reaction could be supplemented. We expected efficient transcription initiation on the T7lac promoter located upstream of the FLG gene by IPTG inactivation of the lac repressor. However, the protein expression was still at a low level from uninduced (0 mM) to fully induced (1.0 mM) IPTG level (**Supplementary Figure 3**), implying that T7 RNA polymerase was not the limiting factor in the translation. The cell-free synthesized proteins visualized on SDS-PAGE indicate comparable results with those of the quantitative protein analysis (**Figures 3C,D**). Although the enhancement effect of codon optimization was lower when coupled with ROS extract (1.8-fold, **Figure 3B**, right panel) than with BL21 extract (7.4-fold, **Figure 3A**, right panel), the absolute protein yield increment was 12.2 μM (soluble protein) using ROS extract. This is not a negligible increase considering the entire production yield from pJL1-FLG Codon Opt paired with BL21 extract, which was 13.5 ± 2.7 μM. Conclusively, the codon optimization markedly enhanced FLG production in both BL21 and ROS cell extract-based CFPS system. Furthermore, the rare tRNA supplement by ROS extract combined with the *E. coli* codon-optimized gene has shown the combinational effect of the overall FLG synthesis yield up to 27.9 ± 5.0 (μM) in the soluble fraction (equivalently, 961 ± 172 μg/ml). This is 15.2-fold greater than that of BL21 extract combined with the WT gene. This result clearly indicated the synergic effect of changes in the plasmid construct and host strain.

Tuning the Cell-Free Protein Synthesis System for Filaggrin Synthesis

Cell Extract Optimization

Multiple factors were explored in this study for preparing cell extract to improve FLG synthesis in the CFPS system. First, we tuned the cell culture and the time for cell harvest. Harvesting the cultured cells at the exponential phase of growth is beneficial for the performance of cell extracts because such cell extracts contain more active ribosomes and other translational machinery than other growth phases (Swartz, 2018). The time for harvest cells at high growth rate (OD_{600} = 2.0) was set to optimal harvest time (Harvest Opt), while the time for harvest cells at the stationary phase (2 h after reaching to maximum OD_{600} = 2.6∼3.0) was set to late harvest time (Harvest Late). Second, the level of cell lysis energy input (joules) is critical to generating highly active cell extract for the CFPS (Kwon and Jewett, 2015; Kim et al., 2019). Insufficient sonication energy lowers the total *E. coli* protein mass, which includes the translation apparatus and brings a higher chance of contamination by residual *E. coli* cells in the cell extract, while excessive energy input causes overheating degradation and deactivation of the proteins and enzymes in the cell extract (Kwon and Jewett, 2015). We tested the three different levels of energy input values representing insufficient energy (Low), optimal energy (Opt), and excessive energy (High). The optimal lysis energy was set to 537 J per processing volume (ml) for BL21 and ROS cell lysis; 269 and 1,074 J was used as Low and High lysis energy, respectively. Lastly, the runoff reaction was chosen as a critical post-lysis processing factor. Runoff reaction is a post-incubation procedure after cell lysis for decoupling ribosomal subunits from the paused translation on the endogenous mRNAs (polysome). In our previous work (Kim et al., 2019), the significant increase in protein yield was observed in the CFPS when using the cell extract prepared with runoff reaction. Dialysis was excluded in this study because we observed a decreased protein yield with dialyzed cell extract previously (Kim et al., 2019).

The combinatorial effect of the three critical cell extract processing factors on cell extract performance was studied by examining the protein yield of the 12 cell extract variations at four different scenarios: sfGFP expression using BL21 extracts

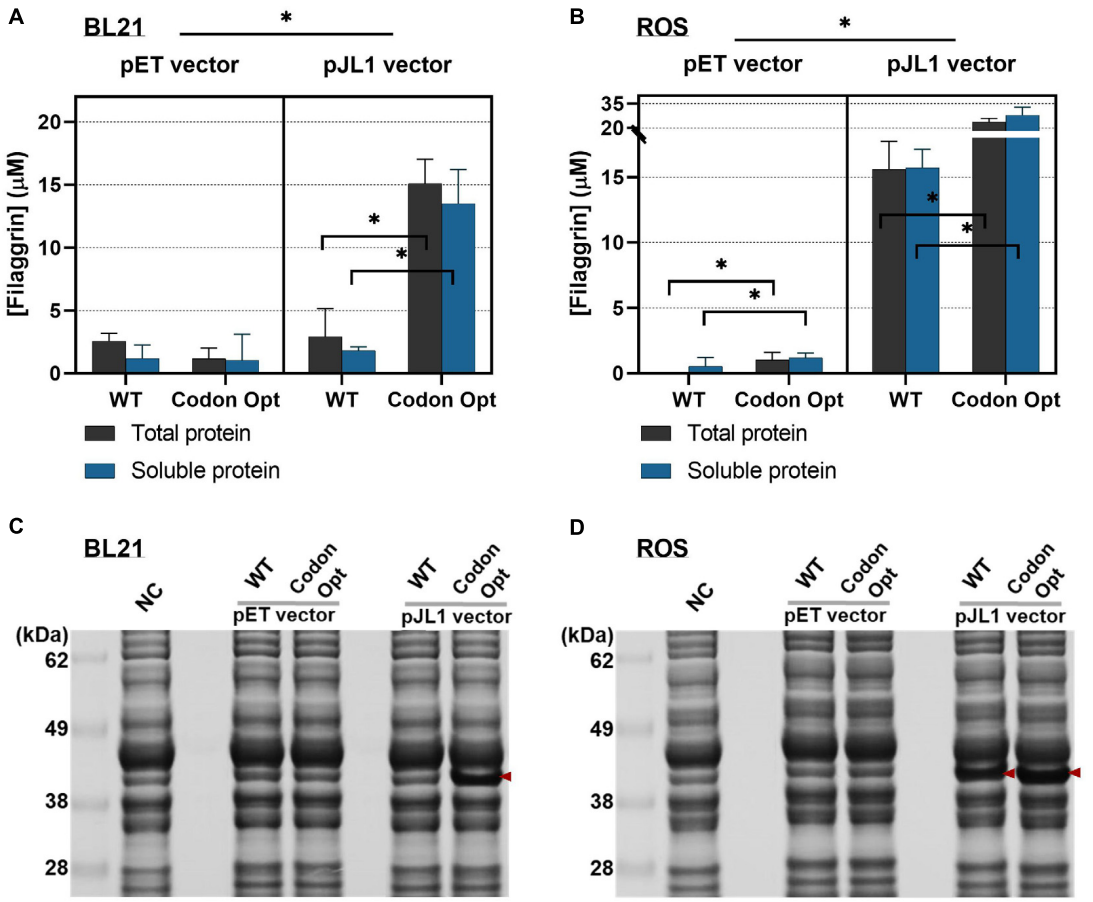


FIGURE 3 | Cell-free synthesized protein yields from different filaggrin (FLG) expression vectors in the cell-free protein synthesis (CFPS). **(A)** FLG synthesis in BL21 extract. From the left to right, pETBlue-1-FLG WT, pETBlue1-FLG Codon Opt, pJL1-FLG WT, and pJL1-FLG Codon Opt. **(B)** FLG synthesis in reactive oxygen species (ROS) extract. From the left to the right, pETBlue-1-FLG WT, pETBlue1-FLG Codon Opt, pJL1-FLG WT, and pJL1-FLG Codon Opt. Data represented as the mean ± SD ($N = 3$). Tukey's multiple comparisons test was used for *post hoc* analysis with two-way ANOVA. *$p < 0.001$. **(C)** and **(D)** Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) images of cell-free synthesized proteins with BL21 and ROS cell extract. The arrows indicate synthesized FLG. All CFPS reactions were performed at 30°C for 20 h.

(**Figure 4A**), sfGFP expression using ROS extracts (**Figure 4B**), FLG expression using BL21 extracts (**Figure 4C**), and FLG expression using ROS extracts (**Figure 4D**) (note that blue bars in each scenario indicate the optimal processing condition giving the highest protein yield). First, we compared the cell-free synthesized sfGFP expression levels using BL21/ROS extracts. As we expected, in **Figure 4A**, the sfGFP synthesis between standard CFPS reaction condition (Harvest Opt:Lysis Opt:No Runoff) and non-runoff condition (Harvest Opt:Lysis Opt:Runoff) in BL21 extract showed the same pattern reported previously (Kwon and Jewett, 2015). However, all the ROS extracts with runoff reactions showed improved sfGFP synthesis levels compared with those in non-runoff conditions (**Figure 4B**, right panel). Next, we studied the cell-free FLG synthesis in BL21/ROS extracts. The FLG synthesis in BL21-based CFPS has shown the highest yield in the combinatorial condition of Harvest Late:Lysis Opt:No Runoff in **Figure 4C** (note that this optimal processing condition significantly improved FLG expression compared with that in the groups in * and **). The highest FLG yields for ROS-based CFPS were obtained in a combinational condition of Harvest Opt:Lysis Opt:No Runoff (**Figure 4D**).

Interestingly, the runoff reaction was able to help to regain the protein synthesis capacity of the cell extract processed with excessive energy (**Figure 4A**, both panels, right bars). The increased pattern was observed in all scenarios (**Figure 4**, Harvest Opt or Late:Lysis High:Runoff) as well. This result indicated the runoff reaction could be an essential step for generating cell extracts that show homogeneous protein synthesis capacity batch by batch. We assume that the excessive sonication condition provided sufficient energy to rupture the most *E. coli* cell wall and release the cytoplasmic components, maximizing the level of endogenous mRNA, genomic DNA, other small molecules, and cell debris that potentially interferes the exogenous gene expression during the CFPS. The runoff reaction can assist in degrading such molecules, which results in reactivating translational machinery in the cell extract processed by the excessive sonication. On the other hand, without the runoff reaction, optimal sonication energy input is considered

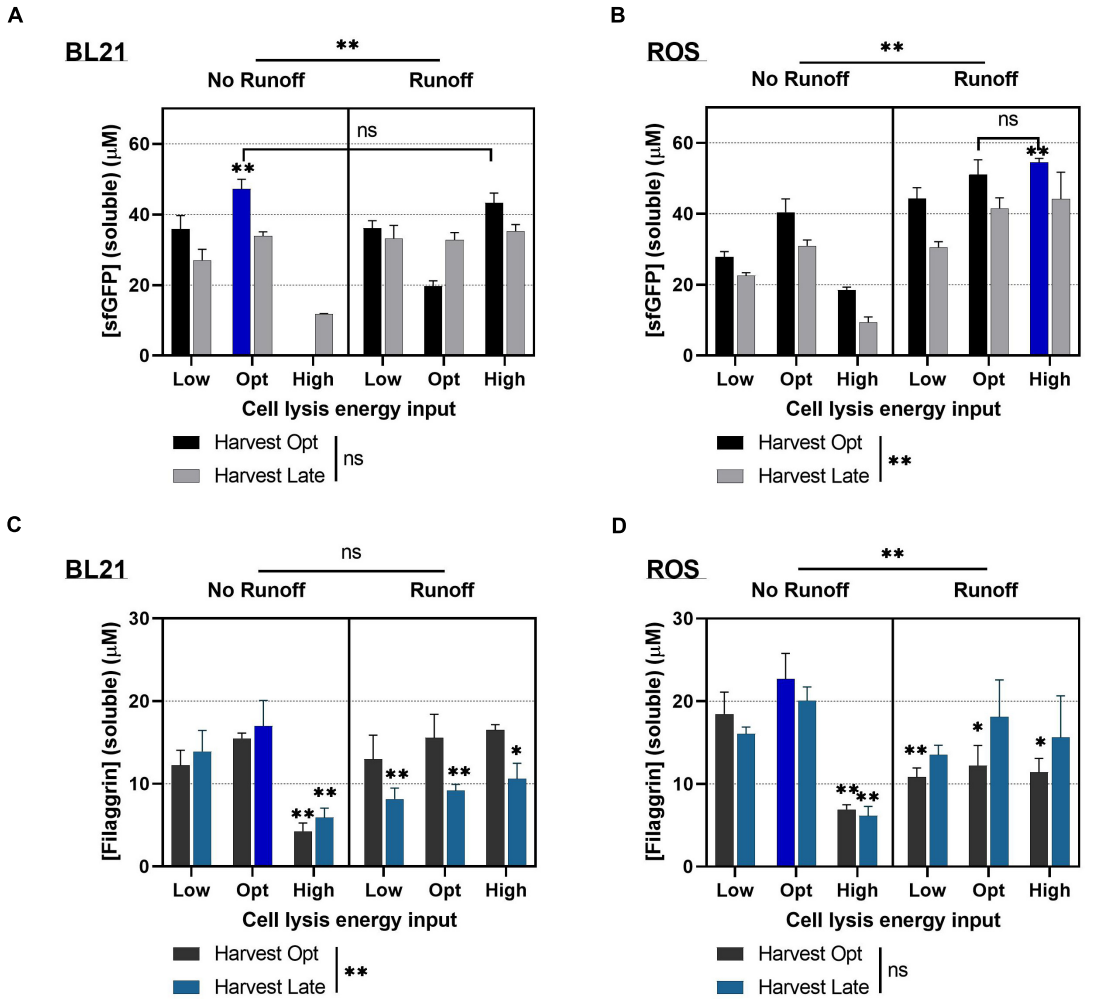


FIGURE 4 | The effect of the combinatorial cell extract processing optimization for the cell-free protein synthesis (CFPS). A set of 12 different cell extracts was prepared for each strain by varying processing conditions with three nominal variables (harvest time, cell lysis energy input, and runoff). **(A)** The impact of processing conditions for BL21 Star (DE3) cell extract (BL21) for the sfGFP expression. **(B)** The impact of processing conditions for Rosetta-gami B (DE3) pLacI [reactive oxygen species (ROS)] cell extract for sfGFP expression. **(C)** The impact of processing conditions for BL21 cell extract for filaggrin (FLG) expression. **(D)** The impact of processing conditions for ROS cell extract for FLG expression. Data represented as the mean ± SD ($N = 3$). Followed by three-way ANOVA along with normality test, the multiple comparison test was performed for *post hoc* analysis. The Dunnett multiple comparisons were used by comparing the mean of a group showing the highest expression with the other group means (11 comparisons). $*p < 0.05$ and $**p < 0.001$ were presented when the difference exists between the control group and another group. All CFPS reactions were performed at 30°C for 20 h.

the major affecter to gain the highest extract performance (**Figure 4**, left panels, middle bars). The overall performance of the CFPS system in terms of protein expression was shown to be significantly affected by cell extract processing conditions. With the combinatorial processing condition for the cell extract, the soluble protein yield in the CFPS increased up to 54 ± 1 μM (1,463 ± 32 μg/ml) for the sfGFP and 23 ± 3 μM (799 ± 109 μg/ml). Total FLG proteins showed the same expression pattern as the soluble portion (**Supplementary Figure 4**).

Even though the cell extract from processing combination of Harvest Opt:Lysis Opt:No Runoff provided a high protein yield in general, there were slight variations in optimal processing condition, which depends on the type of proteins as well as the type of expression hosts. Thus, it is necessary to customize the processing condition to obtain the cell extract specialized for a certain protein cell-free biomanufacturing.

Modifying Cell-Free Protein Synthesis Reaction Conditions

The optimal CFPS conditions vary from protein to protein because of the inherent complexity of protein, gene transcription and translation, and post-translational modification. In the CFPS system, the gene transcription and translation that initiated by the addition of DNA templates are processed simultaneously in both prokaryotic and eukaryotic systems (Zemella et al., 2015). Modifying the CFPS conditions favorable to the specific target protein synthesis can increase the system's

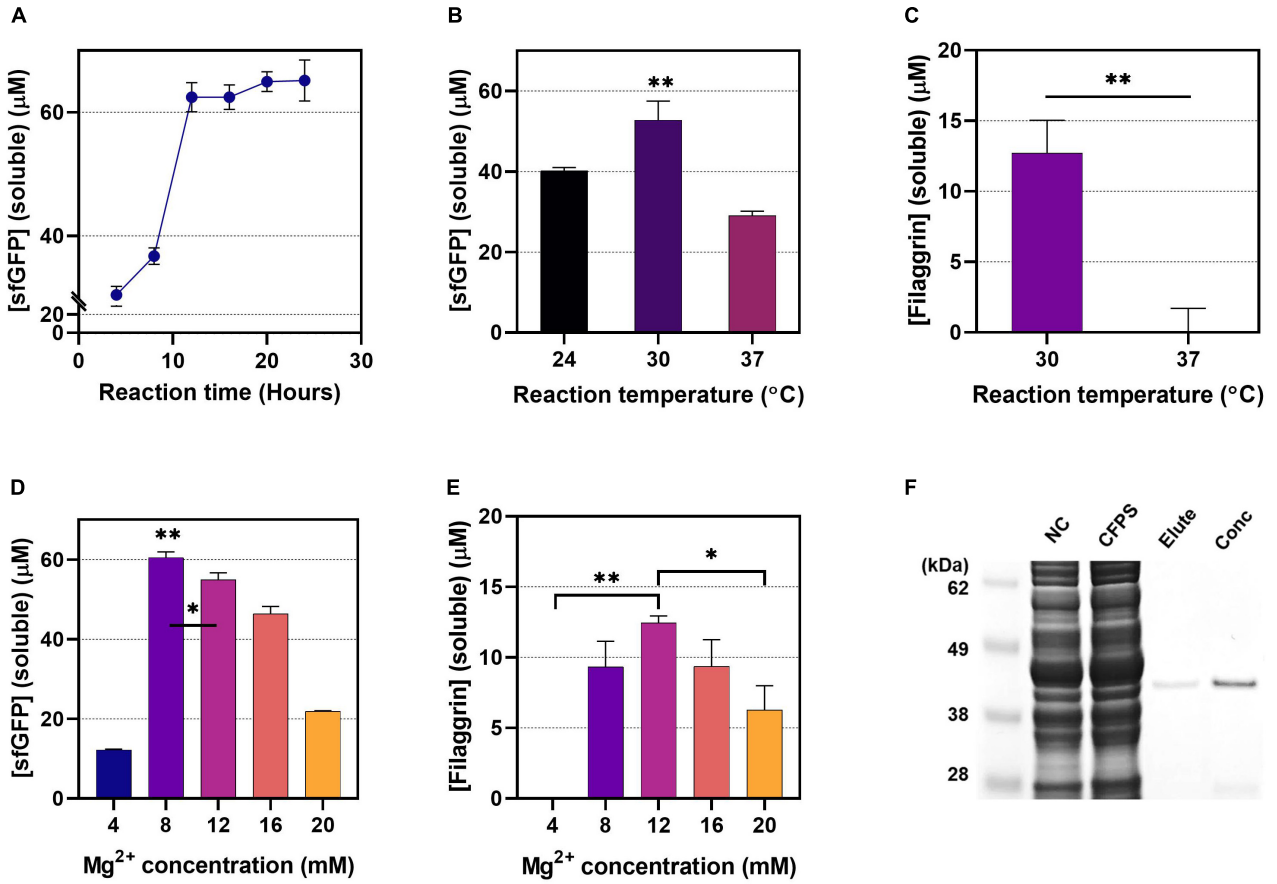


FIGURE 5 | The effect of the cell-free protein synthesis (CFPS) reaction conditions. **(A)** 24-h monitoring of sfGFP expression (CFPS reaction: 30°C, 8 mM of Mg^{2+}). **(B)** The incubation temperature for sfGFP expression. The CFPS reaction was carried out at different temperatures (24°C, 30°C, and 37°C) for 20 h using 12 mM of Mg^{2+}. **(C)** The incubation temperature for filaggrin (FLG) expression. The CFPS reaction was carried out at different temperatures (30°C and 37°C) for 20 h using 12 mM of Mg^{2+}. **(D)** The effect of the concentrations of Mg^{2+} for sfGFP synthesis (CFPS reaction: 30°C, 20 h). **(E)** The effect of the concentrations of Mg^{2+} for FLG synthesis (CFPS reaction: 30°C, 20 h). Data represented as the mean ± SD ($N = 3$). The unpaired t-test and the one-way ANOVA with the Dunnett test were performed. The Dunnett multiple comparisons were used by comparing the mean of a group showing the highest expression with the other group means $^{*}p < 0.05$, $^{**}p < 0.001$. **(F)** Gel electrophoresis of cell-free synthesized and purified FLG. Lane 1, protein size marker; lane 2, CFPS without FLG plasmid; lane 3, soluble FLG in CFPS; lane 4, isolated FLG protein (elute from Ni-NTA IMAC); lane 5, concentrated FLG. The pJL1-FLG^{His} plasmid was used as the expression vector in the reactive oxygen species (ROS) cell extract-based CFPS system.

protein synthesis capability. The incubation temperature is the important factor of protein translation dynamics and directly influences the speed of the transcription and translation-elongation rate, protein folding, and solubility (Yu et al., 2015). Recent researches have shown the synonymous codon substitution affects the translation-elongation rate as well (Spencer et al., 2012; Supek et al., 2014). In addition, magnesium ions (Mg^{2+}) play a critical role in protein synthesis. Lacking or having a shortage of magnesium ions causes incorrect folding of ribosomal RNA and disassembly of the ribosome, which results in early translation termination (Pontes et al., 2016). In this work, we assessed the effect of incubation temperature and the concentration of magnesium ions to explore the optimal condition for FLG synthesis during the CFPS reaction.

The sfGFP expression level was measured at 4-h intervals to gauge the appropriateness of the 20 h of incubation to set as default reaction time when evaluating the effect of the CFPS conditions on protein yield (**Figure 5A**). The sfGFP synthesis increased quickly for the first 10 h of reaction, and then the curve plateaued after 12–16 h of incubation, implying that 20 h is sufficient CFPS reaction time to complete the CFPS reaction at maximum yield when comparing yields at different CFPS conditions. Since we observed that the soluble proteins are comparable with the total protein level in FLG and sfGFP expression (**Figure 2A**), we measured the soluble protein in different temperatures [24°C, 30°C (default), and 37°C] to find out the optimal CFPS reaction. The CFPS reaction was set for 20 h. Both sfGFP and FLG showed the highest soluble protein yield in the 30°C reaction (**Figures 5B,C**). Besides that, the CFPS at 30°C produces a highly soluble form of FLG, which was observed little in the CFPS at 37°C (**Supplementary Figure 5A**), implying that a reaction at 30°C helps the proper protein folding in the *E. coli* CFPS system.

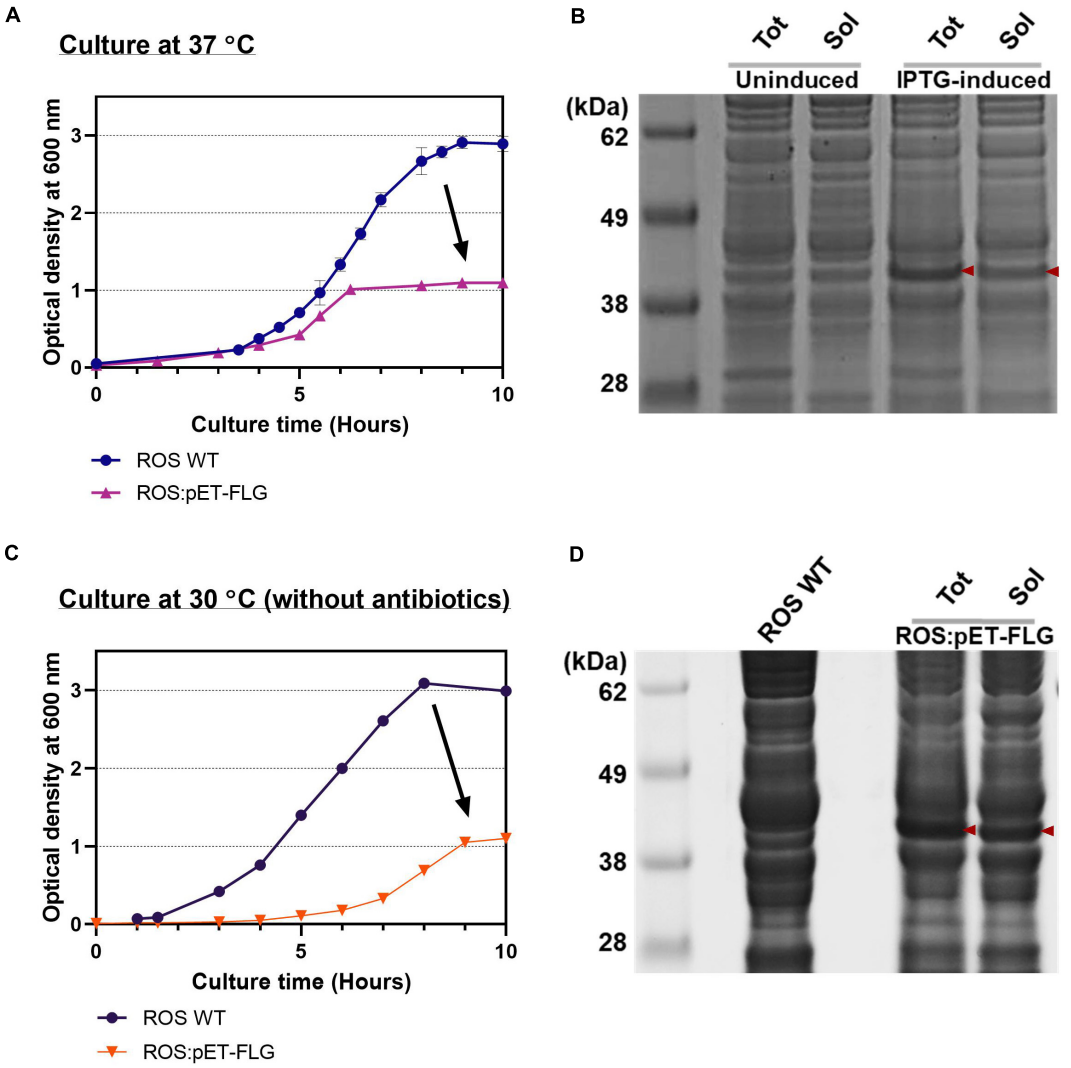


FIGURE 6 | Filaggrin (FLG) expression *in vivo*. **(A)** The cell growth curve of reactive oxygen species (ROS) wild type (WT) and ROS:pET-FLG at 37°C with antibiotics. **(B)** Productivity and solubility of induced and non-induced FLG. **(C)** The cell growth curve of ROS WT and ROS:pET-FLG at 30°C without antibiotics. **(D)** Productivity and solubility of synthesized FLG. Red arrows indicate FLG. Tot, total proteins; Sol, soluble proteins.

We assume that both the FLG gene's codon optimization and the temperature alteration increased the overall protein folding dynamics, particularly the translation-elongation rate of the FLG synthesis. For example, unlike the FLG expression in human, during the *E. coli* CFPS at 37°C, the T7 RNA polymerase-mediated transcription rate and *E. coli* translation-elongation rate would be too fast to have time for protein folding. By lowering the temperature to 30°C, transcription and translation rate would slow down, which provides time for proper protein folding. The optimal concentration of magnesium ion in the CFPS was determined to be 8 mM for sfGFP expression (**Figure 5D**), which was the consistent outcome reported by Dudley et al. (2020). However, the FLG expression was at the highest level in the CFPS system when 12 mM of Mg^{2+} concentration was used (**Figure 5E** and **Supplementary Figures 5B,C**). With the optimized CFPS reaction condition, we successfully isolated and concentrated the cell-free synthesized FLG observed to have a clean single band without noise (**Figure 5F**).

Filaggrin Expression *in vivo*

After achieving a high-yield FLG synthesis via optimizing the CFPS system in this study, we next explore the FLG synthesis *in vivo*. The CFPS system has been compared with *in vivo* recombinant protein synthesis platform in terms of usability, flexibility, scalability, and, more importantly, productivity. One of the features is that the CFPS is accomplished in a single reaction, whereas the *in vivo* system requires multiple procedures for protein expression (Rosenblum and Cooperman, 2014). This allowed the CFPS to become an essential tool to rapidly screen and eliminate factors limiting the production of target proteins. For example, we identify that the codon bias was

the main limiting factor in FLG expression without conducting time-consuming steps of cloning, transformation, and cell cultivation. This highlights the versatility of cell-free platform as a protein production optimization toolkit. Additionally, the CFPS system does not require host strain modification to accommodate the exogenous DNA to express target protein, which causes the poor growth of the host, protein inactivity, and low protein yields (Rosano and Ceccarelli, 2014). Not only the exogenous protein expression but also the presence of the exogenous gene in bacteria can be a burden to the host cell (Andersson et al., 1996).

As the FLG production yield greatly improved through the systemic approach with cell-free platform, the optimized condition was simply applied to the *in vivo* system to stress the suitability of cell-free system as a simulator of FLG production *in vivo*. The ROS cell extract showed promising performance to synthesize FLG *in vitro* due to the endogenous rare tRNA supplement from the plasmid pLacIRARE (**Figure 3B**). The ROS cell extract was prepared following the standard cell extract preparation protocol, as described in *Preparation of Cell Extracts*. We first examined *in vivo* protein synthesis to confirm the FLG synthesis at a higher temperature (37°C) by using the host strain with the plasmid pETBlue1-FLG^{His} [Rosetta-gami B (DE3):pETBlue1-FLG^{His} or simply ROS:pET-FLG]. We observed the slow growth of ROS:pET-FLG compared with WT ROS (ROS WT) **(Figure 6A)**. The cell doubling time (T_d) of ROS WT and ROS:pET-FLG was 46 and 51 min, respectively. IPTG induction was successfully regulated FLG gene translation compared with non-induced cells. However, ROS:pET-FLG cell growth reached at the stationary phase earlier, implying the presence of inhibiting effect. The same trend appeared in IPTG-induced (at OD_{600} = 0.6) and uninduced cell culture. This indicates that the presence of plasmid significantly inhibited cell growth and eventually achieved the lower FLG synthesis yield and solubility (**Figure 6B**). This inhibitory effect results in the difficulty of rapid optimization for target protein expression when using *in vivo* system.

Next, the cell culture condition was slightly modified, similar to the CFPS reaction condition, mainly by lowering culture temperature to 30°C to gain soluble protein synthesis. Also, cell culture, protein expression, and cell harvest conditions were slightly modified to mimic the CFPS condition, which was optimized in this study. For example, ROS WT and ROS:pET-FLG cells were grown in 2xYTPG media at 30°C without antibiotics, IPTG induction at OD_{600} = 0.6, and harvest the cell at OD_{600} = 2.0 (OD_{600} = 1.0 for ROS:pET-FLG) (**Figure 6C**); washed with Buffer A; and disrupted the cell via optimized sonication condition. As a result, we observed increased protein productivity (**Figure 6D**). However, the *in vivo* study revealed that the FLG expression causes a negative effect on cell growth and implies the possible low protein yield that resulted from the reduction of total cell biomass. Consequently, the CFPS with the processing optimization can detour this limitation to ensure the protein synthesis for future biomanufacturing of therapeutic proteins.

CONCLUSION

Although the profit of therapeutic proteins in the pharmaceutics market is significantly expanded in the last few decades, the manufacturing capacity and production cost limit the market availability of biological drugs (Matthews et al., 2017). In this work, we proposed the CFPS system to streamline human therapeutic protein production. In this customized CFPS model, the therapeutic protein FLG expression was significantly improved, representing 28 ± 5 μM of soluble protein yield. As the most FLG protein was expressed with a soluble form in the CFPS system, the protein synthesis enhancement by optimization was evaluated using the amount soluble fraction of FLG. The protein yield was 9.1 times greater with the rare tRNA overexpressing host strain (ROS extract) than with the host strain without rare tRNA (BL21 extract) in the CFPS system. The combination of rare tRNA overexpressing host and DNA codon bias optimization enhanced protein synthesis by up to 15.2-fold of protein yield in the CFPS system. To improve the FLG production yield, we applied the combinational optimization of the three key manufacturing processes without the individual factor optimization. We obtained the 3.7-fold (ROS extract) and 4.0-fold (BL21 extract) improvement of protein yields in the customized CFPS system. Taken together, the customized CFPS system for FLG synthesis described here not only provides the alternative protein synthesis platform but also proves the potential of the CFPS system for biomanufacturing of human therapeutic proteins. In this study, we focused on the high-yield soluble FLG production using the CFPS. Solubility would not be the solid indicator for functionality, but it suggested the proper protein folding and an important prerequisite of it. The follow-up characterization and functionality studies will expand therapeutics applications and capabilities of *in vitro* protein synthesis.

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-CK conceived and supervised the project. JK and Y-CK designed and conceptualized experiments. JK, CEC, KS, and BV performed the experiments and acquired the data. JK, CEC, KS, BV, and Y-CK analyzed data. JK wrote the original manuscript. JK, CEC, and Y-CK revised and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Louisiana Board of Regents [RCS, Grant No. LEQSF(2020-23)RD-A-01], USDA National

Institute of Food and Agriculture (HATCH, Accession No. 1021535, Project No. LAB94414), and Open Access Author fund from LSU Libraries.

ACKNOWLEDGMENTS

JK and Y-CK thank Dr. Huili Yao and Prof. Mario Rivera for their discussion on this work. CEC was supported by the Donald W. Clayton Ph.D. Graduate Assistantship at the college of engineering at Louisiana State University. KS was supported by the Chemistry of Life Processes Predoctoral Training Program at Northwestern University. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number T32GM105538. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. BV was supported by the SNSF Early Postdoc.Mobility fellowship P2SKP3_184036. The content is solely the responsibility of the authors and does not necessarily represent the official views of Northwestern University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.590341/full#supplementary-material

REFERENCES

Adiga, R., Al-Adhami, M., Andar, A., Borhani, S., Brown, S., Burgenson, D., et al. (2018). Point-of-care production of therapeutic proteins of good-manufacturing-practice quality. *Nat. Biomed. Eng.* 2, 675–686.

Andersson, L., Yang, S., Neubauer, P., and Enfors, S. (1996). Impact of plasmid presence and induction on cellular responses in fed batch cultures of *Escherichia coli*. *J. Biotechnol.* 46, 255–263. doi: 10.1016/0168-1656(96)00004-1

Brown, S. J., and McLean, W. H. I. I. (2012). One remarkable molecule: filaggrin. *J. Invest. Dermatol.* 132, 751–762. doi: 10.1038/jid.2011.393

Cabanillas, B., and Novak, N. (2016). Atopic dermatitis and filaggrin. *Curr. Opin. Immunol.* 42, 1–8.

Chen, R. (2012). Bacterial expression systems for recombinant protein production: *E. coli* and beyond. *Biotechnol. Adv.* 30, 1102–1107. doi: 10.1016/j.biotechadv.2011.09.013

Dudley, Q. M., Karim, A. S., Nash, C. J., and Jewett, M. C. (2020). In vitro prototyping of limonene biosynthesis using cell-free protein synthesis. *Metab. Eng.* 61, 251–260. doi: 10.1016/j.ymben.2020.05.006

Duitama, J., Zablotskaya, A., Gemayel, R., Jansen, A., Belet, S., Vermeesch, J. R., et al. (2014). Large-scale analysis of tandem repeat variability in the human genome. *Nucleic Acids Res.* 42, 5728–5741. doi: 10.1093/nar/gku212

Egawa, G., and Kabashima, K. (2018). Barrier dysfunction in the skin allergy. *Allergol. Int.* 67, 3–11. doi: 10.1016/j.alit.2017.10.002

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Gingold, H., and Pilpel, Y. (2011). Determinants of translation efficiency and accuracy. *Mol. Syst. Biol.* 7:481. doi: 10.1038/msb.2011.14

Guarino, C., and DeLisa, M. P. (2012). A prokaryote-based cell-free translation system that efficiently synthesizes glycoproteins. *Glycobiology* 22, 596–601. doi: 10.1093/glycob/cwr151

Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends Biotechnol.* 22, 346–353. doi: 10.1016/j.tibtech.2004.04.006

Hannan, A. J. (2018). Tandem repeats mediating genetic plasticity in health and disease. *Nat. Rev. Genet.* 19:286. doi: 10.1038/nrg.2017.115

Hunt, J. P., Zhao, E. L., Soltani, M., Frei, M., Nelson, J. A. D., and Bundy, B. C. (2019). Streamlining the preparation of "endotoxin-free" ClearColi cell extract with autoinduction media for cell-free protein synthesis of the therapeutic protein crisantaspase. *Synth. Syst. Biotechnol.* 4, 220–224. doi: 10.1016/j.synbio.2019.11.003

Jaroentomeechai, T., Stark, J. C., Natarajan, A., Glasscock, C. J., Yates, L. E., Hsu, K. J., et al. (2018). Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nat. Commun.* 9, 1–11.

Jelinek, W. R., Toomey, T. P., Leinwand, L., Duncan, C. H., Biro, P. A., Choudary, P. V., et al. (1980). Ubiquitous, interspersed repeated sequences in mammalian genomes. *Proc. Natl. Acad. Sci. U.S.A.* 77, 1398–1402. doi: 10.1073/pnas.77.3.1398

Kightlinger, W., Lin, L., Rosztoczy, M., Li, W., DeLisa, M. P., Mrksich, M., et al. (2018). Design of glycosylation sites by rapid synthesis and analysis of glycosyltransferases. *Nat. Chem. Biol.* 14, 627–635. doi: 10.1038/s41589-018-0051-2

Kightlinger, W., Warfel, K. F., DeLisa, M. P., and Jewett, M. C. (2020). Synthetic glycobiology: parts, systems, and applications. *ACS Synth. Biol.* 9, 1534–1562. doi: 10.1021/acssynbio.0c00210

Kim, J., Copeland, C. E., Padumane, S. R., and Kwon, Y. (2019). A crude extract preparation and optimization from a genomically engineered *Escherichia coli* for the cell-free protein synthesis system: practical laboratory guideline. *Methods Protoc.* 2:68. doi: 10.3390/mps2030068

Kwon, Y. C., and Jewett, M. C. (2015). High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* 5, 1–8.

Li, J., Gu, L., Aach, J., and Church, G. M. (2014). Improved cell-free RNA and protein synthesis system. *PLoS One* 9:e106232. doi: 10.1371/journal.pone.0106232

Lithwick, G., and Margalit, H. (2003). Hierarchy of sequence-dependent features associated with prokaryotic translation. *Genome Res.* 13, 2665–2673. doi: 10.1101/gr.1485203

Lu, Y. (2017). Cell-free synthetic biology: engineering in an open world. *Synth. Syst. Biotechnol.* 2, 23–27. doi: 10.1016/j.synbio.2017.02.003

Matsui, T., Miyamoto, K., Kubo, A., Kawasaki, H., Ebihara, T., Hata, K., et al. (2011). SASPase regulates stratum corneum hydration through profilaggrin-to-filaggrin processing. *EMBO Mol. Med.* 3, 320–333. doi: 10.1002/emmm.201100140

Matthews, C. B., Wright, C., Kuo, A., Colant, N., Westoby, M., and Love, J. C. (2017). Reexamining opportunities for therapeutic protein production in eukaryotic microorganisms. *Biotechnol. Bioeng.* 114, 2432–2444. doi: 10.1002/bit.26378

Matthies, D., Haberstock, S., Joos, F., Dötsch, V., Vonck, J., Bernhard, F., et al. (2011). Cell-free expression and assembly of ATP synthase. *J. Mol. Biol.* 413, 593–603. doi: 10.1016/j.jmb.2011.08.055

Min, S. E., Lee, K., Park, S., Yoo, T. H., Oh, C. H., Park, J., et al. (2016). Cell-free production and streamlined assay of cytosol-penetrating antibodies. *Biotechnol. Bioeng.* 113, 2107–2112. doi: 10.1002/bit.25985

Murakami, S., Matsumoto, R., and Kanamori, T. (2019). Constructive approach for synthesis of a functional IgG using a reconstituted cell-free protein synthesis system. *Sci. Rep.* 9, 1–13.

Nachat, R., Méchin, M.-C., Takahara, H., Chavanas, S., Charveron, M., Serre, G., et al. (2005). Peptidylarginine deiminase isoforms 1-3 are expressed in the epidermis and involved in the deimination of K1 and filaggrin. *J. Invest. Dermatol.* 124, 384–393. doi: 10.1111/j.0022-202x.2004.23568.x

Novoa, E. M., and de Pouplana, L. R. (2012). Speeding with control: codon usage, tRNAs, and ribosomes. *Trends Genet.* 28, 574–581. doi: 10.1016/j.tig.2012.07.006

O'Regan, G. M., Sandilands, A., McLean, W. H. I., and Irvine, A. D. (2008). Filaggrin in atopic dermatitis. *J. Allergy Clin. Immunol.* 122, 689–693.

Otsuka, A., Doi, H., Egawa, G., Maekawa, A., Fujita, T., Nakamizo, S., et al. (2014). Possible new therapeutic strategy to regulate atopic dermatitis through upregulating filaggrin expression. *J. Allergy Clin. Immunol.* 133, 139–146.e10.

Pardee, K., Slomovic, S., Nguyen, P. Q., Lee, J. W., Donghia, N., Burrill, D., et al. (2016). Portable, on-demand biomolecular manufacturing. *Cell* 167, 248–259. doi: 10.1016/j.cell.2016.09.013

Perez, J. G., Stark, J. C., and Jewett, M. C. (2016). Cell-free synthetic biology: engineering beyond the cell. *Cold Spring Harb. Perspect. Biol.* 8:a023853. doi: 10.1101/cshperspect.a023853

Pontes, M. H., Yeom, J., and Groisman, E. A. (2016). Reducing ribosome biosynthesis promotes translation during low Mg2+ stress. *Mol. Cell* 64, 480–492. doi: 10.1016/j.molcel.2016.05.008

Quax, T. E. F., Claassens, N. J., Söll, D., and van der Oost, J. (2015). Codon bias as a means to fine-tune gene expression. *Mol. Cell* 59, 149–161. doi: 10.1016/j.molcel.2015.05.035

Ram, K., Hatton, D., Ahuja, S., Bender, J., Hunter, A., and Turner, R. (2016). "Protein production in eukaryotic cells," in *Protein Therapeutics. Topics in Medicinal Chemistry*, eds Z. Sauna and C. Kimchi-Sarfaty (Cham: Springer), 1–39.

Rosano, G. L., and Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* 5:172. doi: 10.3389/fmicb.2014.00172

Rosenblum, G., and Cooperman, B. S. (2014). Engine out of the chassis: cell-free protein synthesis and its uses. *FEBS Lett.* 588, 261–268. doi: 10.1016/j.febslet.2013.10.016

Salehi, A. S. M., Smith, M. T., Bennett, A. M., Williams, J. B., Pitt, W. G., and Bundy, B. C. (2016). Cell-free protein synthesis of a cytotoxic cancer therapeutic: onconase production and a just-add-water cell-free system. *Biotechnol. J.* 11, 274–281. doi: 10.1002/biot.201500237

Sanchez-Garcia, L., Martín, L., Mangues, R., Ferrer-Miralles, N., Vázquez, E., and Villaverde, A. (2016). Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microb. Cell Fact.* 15:33.

Sandilands, A., Terron-Kwiatkowski, A., Hull, P. R., O'Regan, G. M., Clayton, T. H., Watson, R. M., et al. (2007). Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat. Genet.* 39, 650–654. doi: 10.1038/ng2020

Shinoda, T., Shinya, N., Ito, K., Ishizuka-Katsura, Y., Ohsawa, N., Terada, T., et al. (2016). Cell-free methods to produce structurally intact mammalian membrane proteins. *Sci. Rep.* 6, 1–15.

Silverman, A. D., Karim, A. S., and Jewett, M. C. (2019). Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* 21, 151–170. doi: 10.1038/s41576-019-0186-3

Spencer, P. S., Siller, E., Anderson, J. F., and Barral, J. M. (2012). Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. *J. Mol. Biol.* 422, 328–335. doi: 10.1016/j.jmb.2012.06.010

Stano, P. (2019). Gene expression inside liposomes: from early studies to current protocols. *Chem. Eur. J.* 25, 7798–7814. doi: 10.1002/chem.201806445

Stark, J. C., Jaroentomeechai, T., Moeller, T. D., Dubner, R. S., Hsu, K. J., Stevenson, T. C., et al. (2019). On-demand, cell-free biomanufacturing of conjugate vaccines at the point-of-care. *bioRxiv* [Preprint], doi: 10.1101/681841

Stout, T. E., McFarland, T., Mitchell, J. C., Appukuttan, B., and Stout, J. T. (2014). Recombinant filaggrin is internalized and processed to correct filaggrin deficiency. *J. Invest. Dermatol.* 134, 423–429. doi: 10.1038/jid.2013.284

Supek, F., Miñana, B., Valcárcel, J., Gabaldón, T., and Lehner, B. (2014). Synonymous mutations frequently act as driver mutations in human cancers. *Cell* 156, 1324–1335. doi: 10.1016/j.cell.2014.01.051

Swartz, J. R. (2018). Expanding biological applications using cell-free metabolic engineering: an overview. *Metab. Eng.* 50, 156–172. doi: 10.1016/j.ymben.2018.09.011

Swartz, J. R., Jewett, M. C., and Woodrow, K. A. (2004). "Cell-free protein synthesis with prokaryotic combined transcription-translation BT - recombinant gene expression: reviews and protocols," in *Methods in Molecular Biology*, eds P. Balbás and A. Lorence (Totowa, NJ: Humana Press), 169–182. doi: 10.1385/1-59259-774-2:169

Thyssen, J. P., and Kezic, S. (2014). Causes of epidermal filaggrin reduction and their role in the pathogenesis of atopic dermatitis. *J. Allergy Clin. Immunol.* 134, 792–799. doi: 10.1016/j.jaci.2014.06.014

Verstrepen, K. J., Jansen, A., Lewitter, F., and Fink, G. R. (2005). Intragenic tandem repeats generate functional variability. *Nat. Genet.* 37, 986–990. doi: 10.1038/ng1618

Wang, X., Corin, K., Baaske, P., Wienken, C. J., Jerabek-Willemsen, M., Duhr, S., et al. (2011). Peptide surfactants for cell-free production of functional G protein-coupled receptors. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9049–9054. doi: 10.1073/pnas.1018185108

Wilding, K. M., Zhao, E. L., Earl, C. C., and Bundy, B. C. (2019). Thermostable lyoprotectant-enhanced cell-free protein synthesis for on-demand endotoxin-free therapeutic production. *N. Biotechnol.* 53, 73–80. doi: 10.1016/j.nbt.2019.07.004

Yin, G., Garces, E. D., Yang, J., Zhang, J., Tran, C., Steiner, A. R., et al. (2012). Aglycosylated antibodies and antibody fragments produced in a scalable in vitro transcription-translation system. *MAbs* 4, 217–225. doi: 10.4161/mabs.4.2.19202

Yu, C.-H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M. S., et al. (2015). Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. *Mol. Cell* 59, 744–754. doi: 10.1016/j.molcel.2015.07.018

Zawada, J. F., Yin, G., Steiner, A. R., Yang, J., Naresh, A., Roy, S. M., et al. (2011). Microscale to manufacturing scale-up of cell-free cytokine production-a new approach for shortening protein production development timelines. *Biotechnol. Bioeng.* 108, 1570–1578. doi: 10.1002/bit.23103

Zemella, A., Thoring, L., Hoffmeister, C., and Kubick, S. (2015). Cell-free protein synthesis: pros and cons of prokaryotic and eukaryotic systems. *Chembiochem* 16, 2420–2431. doi: 10.1002/cbic.201500340

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

METHODS
published: 02 February 2021
doi: 10.3389/fbioe.2020.604091

Design, Development and Optimization of a Functional Mammalian Cell-Free Protein Synthesis Platform

Chiara Heide[1,2,3], *Gizem Buldum*[1], *Ignacio Moya-Ramirez*[1,3], *Oscar Ces*[2,4], *Cleo Kontoravdi*[1]* *and Karen M. Polizzi*[1,3]*

[1] Department of Chemical Engineering, Imperial College London, London, United Kingdom, [2] Department of Chemistry, Imperial College London, London, United Kingdom, [3] Imperial College Center for Synthetic Biology, Imperial College London, London, United Kingdom, [4] Institute of Chemical Biology, Imperial College London, London, United Kingdom

In this paper, we describe the stepwise development of a cell-free protein synthesis (CFPS) platform derived from cultured Chinese hamster ovary (CHO) cells. We provide a retrospective summary of the design challenges we faced, and the optimized methods developed for the cultivation of cells and the preparation of translationally active lysates. To overcome low yields, we developed procedures to supplement two accessory proteins, GADD34 and K3L, into the reaction to prevent deactivation of the translational machinery by phosphorylation. We compared different strategies for implementing these accessory proteins including two variants of the GADD34 protein to understand the potential trade-offs between yield and ease of implementation. Addition of the accessory proteins increased yield of turbo Green Fluorescent Protein (tGFP) by up to 100-fold depending on which workflow was used. Using our optimized protocols as a guideline, users can successfully develop their own functional CHO CFPS system, allowing for broader application of mammalian CFPS.

Keywords: cell-free protein synthesis, Chinese hamster ovary cells, cell extract, synthetic biology, *in vitro* transcription-translation, coupled batch reactions

OPEN ACCESS

Edited by:
Jian Li,
ShanghaiTech University, China

Reviewed by:
Bo Zhu,
Tokyo Institute of Technology, Japan
Lei Kai,
Jiangsu Normal University, China

***Correspondence:**
Cleo Kontoravdi
cleo.kontoravdi98@imperial.ac.uk
Karen M. Polizzi
k.polizzi@imperial.ac.uk

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

Received: *08 September 2020*
Accepted: *29 December 2020*
Published: *02 February 2021*

Citation:
Heide C, Buldum G, Moya-Ramirez I, Ces O, Kontoravdi C and Polizzi KM (2021) Design, Development and Optimization of a Functional Mammalian Cell-Free Protein Synthesis Platform. Front. Bioeng. Biotechnol. 8:604091. doi: 10.3389/fbioe.2020.604091

INTRODUCTION

Cell-free protein synthesis (CFPS) is an emerging research field. Originally developed to decipher the genetic code (Nirenberg and Matthaei, 1961), CFPS has recently become a powerful tool, providing new opportunities for protein expression, metabolic engineering, therapeutic development, education, and more (Gregorio et al., 2019). CFPS has several advantages over traditional cell-based expression including an open reaction environment, which allows addition of enhancers and non-natural compounds to facilitate protein expression, lack of requirement to maintain cells in a living state, and the ability to direct all energy and cellular machinery to translate mRNA encoding the protein of interest.

Among CFPS platforms, eukaryotic systems are of increasing interest due to their ability to produce post-translationally modified proteins (Mikami et al., 2006a; Brödel et al., 2013; Zemella et al., 2018). Systems have been developed based on yeast (Hodgman and Jewett, 2013; Gan and Jewett, 2014; Aw and Polizzi, 2019; Zhang et al., 2020), plants (Madin et al., 2000; Murota et al., 2011; Buntru et al., 2014, 2015), insect (Ezure et al., 2006, 2010; Madono et al., 2011; Richter et al., 2014; Stech et al., 2014), and mammalian cells

(Stavnezer and Huang, 1971; Shields and Blobel, 1978; Jackson and Hunt, 1983; Starr and Hanover, 1990; Katzen and Kudlicki, 2006; Mikami et al., 2006a,b; Mikami et al., 2008; Brödel et al., 2014; Yadavalli and Sam-Yellowe, 2015; Thoring et al., 2017; Burgenson et al., 2018; Thoring and Kubick, 2018). Mammalian systems are of particular interest due to their ability to produce glycoproteins with human-like N-linked glycosylation. This has led to the development of CFPS platforms for human (Mikami et al., 2008; Yadavalli and Sam-Yellowe, 2015; Burgenson et al., 2018) and Chinese Hamster Ovary (CHO) cells (Brödel et al., 2014, 2015; Martin et al., 2017; Thoring et al., 2017; Thoring and Kubick, 2018) among others. Among previous work demonstrating CFPS using CHO, only one group has reported the successful development of a reliable CFPS platform so far (Brödel et al., 2014, 2015; Thoring et al., 2017; Thoring and Kubick, 2018), while others have relied on the use of the commercial "1-step CHO High Yield *In vitro* Translation (IVT)" kit (Thermo Scientific, West Palm Beach, FL, now discontinued) (Martin et al., 2017). Although the cost of commercial kits is substantial, they simplify the protein expression workflow and offer standardized reagents for reproducible CFPS performance. However, since the product details of the commercial kits are not disclosed and reagents are provided as pre-mixed stocks, they have limited flexibility for customizing the reaction for different purposes.

It is therefore of great interest to increase the understanding of mammalian-based CFPS platforms and promote the development of in-house systems with improved protein yields. In this methods paper, we share our scientific insights and technical protocols with the cell-free community. Specifically, we developed a CFPS based on suspension-adapted CHO cells that includes the supplementation of two accessory proteins to increase protein yields. We studied three variations of implementing these accessory proteins and show increased expression of tGFP in all cases.

Batch reactions are well-suited to high-throughput screening as they can easily and quickly be prepared. Yields of CFPS platforms range from 1 μg/mL to 2.3 mg/mL in batch mode depending on the complexity of proteins (Kobs, 2008; Caschera and Noireaux, 2014). In CHO cell-based CFPS systems, total protein yields of 9–51.3 μg/mL have been achieved in batch mode (Brödel et al., 2014; Stech et al., 2017). Our cell-free system generates yields of up to 20 μg/mL of tGFP with accessory protein supplementation, which is in line with previously reported yields from CHO cell-based mammalian CFPS. This technical protocol should enable complete novices to successfully develop their own CHO CFPS system with results comparable to our reported data.

MATERIALS AND METHODS

CHO Cell Culture and Lysate Generation

The CHO cell lysate was prepared from FreeStyle™ CHO-S cells (Invitrogen, UK) cultured in Erlenmeyer flasks (Corning, Netherlands) using chemically defined animal component-free Gibco™ 1X CD CHO medium (Invitrogen, UK), supplemented with 8 mM L-Glutamine (Thermofisher Scientific, UK) and 10 ml/l of 100X hypoxanthine/thymidine supplement (Thermofisher Scientific, UK). Cells were revived by thawing in a water bath at 37°C, followed by centrifugation to remove the freezing medium and resuspension in 10 mL of complete growth medium. Cultures were seeded at an initial density of 3×10^5 cells/ml in 125 ml vented Erlenmeyer flasks and subcultured every 3–4 days in complete growth medium at a seeding density of 2×10^5 cells/ml. Cells were subcultured at least twice before use for extract preparation. Cultures were maintained in a humidified incubator on an orbital shaking platform rotating at 125 rpm at 37°C and 5% CO_2. Approximately 200–250 ml of CHO-S cells at a density of $\sim 4 \times 10^6$ cells/ml were required for the preparation of a few ml of lysate and culture viability of >90% were required for active lysates.

With the exception of the energy content studies, CHO cells were harvested from the batch culture at mid-log phase (day 4). Cells were collected via centrifugation for 5 min at 200 × g at 4°C and washed twice with ice-cold 25 ml HEPES-based buffer (40 mM HEPES-KOH (pH 7.5), 100 mM NaOAc, and 4 mM DTT, 1/per 50 ml complete EDTA-free Protease Inhibitor Tablet). To prevent degradation during the lysate preparation, the cell suspension was mixed gently and the pellets re-suspended by pipetting. For each wash step, the cells were centrifuged for 5 min at 200 × g at 4°C except for the final wash, when the cells were centrifuged for 10–15 min to ensure they were fully pelletized prior to final re-suspension.

Lysates were prepared in a cold room and handling time was kept to a minimum to prevent loss of activity. The pellet was re-suspended in HEPES-based buffer to a final cell density of $\sim 4.0 \times 10^8$ cells/mL, transferred into pre-cooled 3 or 5 ml glass Snap cap vials, and closed with a transparent plastic lid. Two holes were punched into the lid- one to insert the needle and a second to avoid vacuum formation. The cells were mechanically lysed by pulling and pushing the cell suspension through BD Precisionglide® 23-, 25-, 27-gauge, L 1 1/4 in. needles using sterile, disposable 2 ml BD Plastipak™ syringes. The lysis procedure was started with the largest needle size (23-gauge) and when the resistance of the cell suspension noticeably decreased, smaller needle sizes were used to increase the shear forces and disrupt the remaining cells. The progress of lysis was observed through a light microscope and was deemed complete when over 90% of cells were lysed.

The crude lysate was transferred into 1.5 ml Eppendorf tubes and centrifuged at 10,000 × g, 4°C for 10 min to remove the nuclei and cell debris. The supernatant was transferred into fresh 1.5 ml Eppendorf tubes. To degrade residual DNA, the supernatant was treated with micrococcal nuclease at room temperature. 100 mM $CaCl_2$ was added at a final concentration of 1 mM, followed by 10 U/mL S7 nuclease (M0247S, NEB UK) and the lysate incubated for 2 min at room temperature. The reaction was stopped by adding EGTA to a final concentration of 6.7 mM. Finally, the lysate was flash-frozen in 200 μl aliquots using liquid nitrogen to avoid repeated freeze-thaw cycles and stored at −80°C.

ATP and GTP Quantification

The nucleotide content was determined using high-performance ion exchange chromatography (HPAEC) as previously reported

by del Val et al. (2013). In brief, CHO cells were harvested, washed and quenched using ice-cold 0.9% w/v aqueous NaCl solution. One volume of cell culture sample (1.5×10^7 cells) was added to four volumes of quenching solution. The mixture was centrifuged (1,000 g, 1 min, 4°C), and the cell pellet was re-suspended in a second volume of quenching solution (ice-cold 0.9% w/v NaCl). This suspension was centrifuged again (1,000 g, 1 min) to obtain the cell pellet for acetonitrile extraction. To isolate the intracellular nucleotides, the cell pellet was then re-suspended in 3 ml of ice-cold 50% v/v aqueous acetonitrile solution. 2 µl of 20 mM GDP-Glc were added as an internal standard and the resulting suspension was incubated on ice for 10 min. The sample was centrifuged (0°C, 18,000 g, 5 min) and the supernatant was collected and dried at room temperature using a SpeedVac (Savant, USA). The dried extract was re-suspended in 240 µl of deionized water and stored at −80°C until analysis. The sample was thawed at room temperature and filtered using 0.2-µm sterile filtered Sartorius™ Minisart™ Plus Syringe Filters before analysis by HPAEC using an Alliance HPLC system (Waters) equipped with a CarboPac PA-1 column and a PA-1 guard column (Dionex, USA). HPLC-grade GTP and ATP were used to create a standard curve for the quantification of the intracellular GTP and ATP content of the cells.

Cell-Free Protein Synthesis Reactions

The reaction mix contains a large number of reagents and its activity is highly dependent on its preparation speed and method. Due to its complex composition, preparation of the mix is highly susceptible to pipetting errors and reagent degradation. Hence, in order to obtain comparable data for assessing the effect of lysis preparation and accessory proteins, we used the commercial reaction mix. The reactions were performed using the optimized protocol of the "1-Step Human Coupled IVT Kit" (ThermoScientific, West Palm Beach, Fl) as a guideline.

All CFPS reactions were carried out in coupled batch mode in a total reaction volume of 25 µL.The reactions were composed of four different pre-prepared components (template DNA, lysate, reaction mix, and accessory proteins). For each 25 µL reaction, 1 µg template DNA was added. While a reaction mix based on previous work was initially used for the CFPS reactions (Aw and Polizzi, 2019), in the work described here we used the reaction mix supplied with the commercial kit to enable direct comparison. All components were thawed on ice and mixed together at room temperature. The lysate was incubated with the accessory proteins for 10 min and then the reaction mix and the DNA were added. The plates were sealed with BreatheEasy sealing membrane (Sigma Aldrich). CFPS was conducted overnight at 30°C in a 384-well plate (Corning) in a Synergy HT Microtiter Platereader (BioTek, Winooski, VT) with shaking on the medium setting. The model protein tGFP was expressed using the positive control (pT7CFE1-tGFP-CHis) vector included in the 1-Step Human High-Yield IVT Kit (Thermo Scientific, West Palm Beach, FL).

Functional tGFP Quantification

tGFP protein yields were quantified by fluorescence measurement using a Synergy HT Microtiter platereader (BioTek, Winooski, VT) with an excitation wavelength of 485 nm and an emission wavelength of 528 nm. Recombinant tGFP from Evrogen (Cambridge Bioscience, UK) was used to establish a standard curve to convert fluorescence units into concentrations.

Recombinant Expression and Purification of Accessory Proteins

Expression and purification of K3L, as well as full-length and truncated (Δ1–240) GADD34 was conducted as described in Mikami et al. using *E. coli* C41 (DE3) (Lucigen, Middleton, WI, USA) as the expression host (Mikami et al., 2006a, 2010a). The plasmids pGEX-6P-GADD34-FLAG and pTac-His-K3L were a kind gift from Hiroaki Imataka (University of Hyogo). The N-terminally truncated form of GADD34 was amplified by PCR and cloned in to pGEX-6P-1. After purification, K3L and both versions of GADD34 were subsequently buffer exchanged with 20 mM HEPES pH 7.0–7.6 (Sigma-Aldrich), and concentrated using Vivaspin 500 centrifugal concentrators with a molecular weight cut-off (MWCO) of 3 kDa and 30 kDa (Sigma-Aldrich, St. Louis, MO), respectively. For maximal activity of the accessory proteins, the purification and concentration steps were conducted in a cold room.

Transient Expression of Truncated GADD34 in CHO Cells

The truncated GADD34 sequence (Δ1–240) was cloned into the mCherry2-C1 backbone (Addgene plasmid #54563), a kind gift from Michael Davidson, in frame with the mCherry coding sequence to generate the pmCherry-Trunc-GADD34 construct. pmCherry-Trunc-GADD34 was prepared using the Endotoxin-free Maxi kit (Qiagen, UK) following manufacturer's instruction and the purified plasmid was diluted to 1 µg/µL.

FreeStyle™ CHO-S cells (Invitrogen, UK) were used for transient expression of GADD34Δ1–240 as a C-terminal fusion with mCherry. Culturing conditions were the same as previously described. The cells were revived in a 125 mL polycarbonate sterile Erlenmeyer shake flask containing 30 mL of pre-warmed FreeStyle™ CHO Expression Medium (Thermo Scientific, UK) supplemented with 8 mM L-Glutamine (Sigma-Aldrich, UK). Seventy-two hours post-revival, the cells were subcultured at an initial seeding density of 0.3×10^6 viable cells/mL in pre-warmed medium, followed by subsequent subculturing at 1×10^6 viable cells/mL every 3 days. A minimum of five passages occurred before transfection to allow for cell recovery. Forty-eight hours prior to transfection, cells were passaged to obtain a density of $4–8 \times 10^6$ cells/mL on transfection day. Immediately prior to transfection, the culture was diluted with complete medium to a final density of 2×10^6 cells/mL in a 500 mL Erlenmeyer flask containing 80 mL culture medium.

The transfection complex consisting of 8 mL FreeStyle CHO Expression medium, 96 µL plasmid DNA, 80 µL TransIT-PRO Reagent (Cambridge Bioscience, UK), and 40 µl PRO Boost Reagent (Cambridge Bioscience, UK) was prepared. TransIT-PRO and PRO Boost Reagents were incubated at room temperature and gently vortexed prior to use. The plasmid

DNA was diluted with the FreeStyle CHO medium and after gentle mixing, TransIT-PRO Reagent and PRO Boost Reagent were successively added. The prepared transfection mixture was incubated at room temperature for 10 min to allow sufficient time for complexes to form. After adding the transfection complex to the cell culture, the transfected cells were incubated for 3 days under the same culturing conditions previously described. Three days after transfection, the cells were harvested and lysate was prepared as described previously. A culture viability above 85% was required to produce active lysates. Lysates with transiently expressed mCherry-GADD34Δ1–240 were stored and handled in the same way as lysates without transient expression.

To calculate the amount of expressed mCherry- GADD34Δ1–240 fusion protein, 5×10^6 cells were washed twice in cold PBS and pelleted by centrifugation at 2,500 × g for 5 min. 1 mL of RIPA buffer (Thermofisher Scientific, UK) was added to the cell pellet. The re-suspended pellet was incubated on ice with gentle shaking for 15 min. The mixture was centrifuged at ∼14,000 × g for 15 min to pellet the cell debris and supernatant was used to perform fluorescence measurements. An mCherry standard curve was prepared using mCherry protein (Biovision, mCherry Quantification Kit) serially diluted in RIPA buffer with a range of 50 ng to 0.8 μg. The amount of expressed mCherry-GADD34Δ1–240 fusion in the lysate was calculated as 0.18 μg/μL.

Statistical Analysis

Data was analyzed via one-way analysis of variance (ANOVA), followed by Student's *t*-test using a two-tailed distribution. A value of $p < 0.05$ is denoted by *, $p < 0.01$ by **, $p < 0.001$ by ***.

RESULTS

Generating Functional CHO Cell Lysates

Our initial aim was to develop functional CHO cell lysates, which involved the culture and characterization of CHO cells, the determination of optimal cell harvesting time and the identification of a successful method to lyse the cells while retaining the activity of the intracellular machinery. We initially evaluated the growth characteristics of two parental CHO cell lines, CHO-S and CHO-K1. CHO-K1 was used in previous work (Stech et al., 2013), but this cell line is subject to clumping even though it is suspension adapted, making it more difficult to efficiently lyse (**Supplementary Figure 1**). Moreover, CHO-S cells recover more quickly from cryopreservation and have a higher growth rate, leading to a 1.5-fold higher cell density on day 5 of culture, which increases the speed of lysate preparation. Therefore, we chose CHO-S for all further work.

The characterization of the cell culture is a crucial prerequisite to find the best harvesting conditions for generating active lysate. Harvesting the cell at the right time can lead to improved CFPS yields and better cost efficiency. Therefore, we first evaluated the effect of the day of harvest on the activity of the lysate generated. We hypothesized that the nucleotide triphosphate content at the harvest point might affect the translational activity by influencing the amount of available energy for the reaction as well as the phosphorylation state of translational machinery (Jewett et al., 2009). Cells were harvested on days 3, 4, and 5, corresponding to early, mid, and late exponential phase, respectively, and the ATP and GTP content were measured by anion exchange chromatography. Overall, the ATP content was more variable than the GTP content and peaked on day 4 (**Figure 1A**). On the other hand, while the GTP content was statistically lower on day 3 than day 4 or 5, no statistically significant difference between the GTP content on day 4 and day 5 could be observed ($p = 0.37$). Based on the results evaluated, day 4 appeared to be the optimal harvest day with respect to energy content.

We then compared the yields of the model protein tGFP from CFPS using the lysates prepared from cells harvested on days 3, 4 and 5 (**Figure 1B**). We observed the highest tGFP expression in reactions using lysates from cells harvested on day 4, followed by day 3 and subsequently day 5. The results indicate day 4 is the optimal harvest time for preparing lysates with maximal activity. While the reason for this observation requires further study, the ATP content might be an indicator of the overall status of the cell. For example, ATP availability might correlate with the phosphorylation state of translation initiation factors or the expression level of metabolic enzymes. This is in contrast with previous work that reported the highest activity in extracts from cells harvested from late-log phase (Brödel et al., 2015). However, different cell lines were used (CHO-S vs. CHO-K1) and the mode of culture was also different (shake flasks vs. bioreactors), suggesting that the optimal harvest time may vary depending on the cell line and culture conditions. It is hence important to characterize a chosen cell line to find the optimal harvesting time for each growth condition.

Once the optimal harvest day was found, we compared a range of lysis methods to identify the best method for generating active lysates. The ideal cell lysis method balances efficient disruption of the cell membrane with retaining the activity of the cellular machinery. We began by taking the protocol previously published by Brödel et al. (2014, 2015), which recommends using 20-gauge needles, as a starting point. However, the lysis efficiency was lower than expected (**Supplementary Figure 2**). Therefore, different lysis methods (sonication, homogenization, nitrogen cavitation, freeze-thaw cycles, and needles) were studied. Among the different approaches, only the "needle method" proved to be successful in efficiently lysing the cells (**Figure 2**). However, much smaller needle sizes were required (23-gauge, 25-gauge, and 27-gauge used in succession), possibly due to the smaller size of CHO-S cells compared to CHO-K1 cells. Harsher methods like sonication or homogenization, which are traditionally used for the preparation of microbial cell extract did not prove to be successful. In these cases, the cells were lysed efficiently, but the extract was not functional (see **Supplementary Table 1** for a summary). We present the details of our lysate preparation workflow in **Figure 2**.

Supplementing Reactions With Purified GADD34 and/or K3L Leads to an Increase in Expression Yield

Having generated active lysates, we focused on developing our own workflows for the preparation and addition of accessory

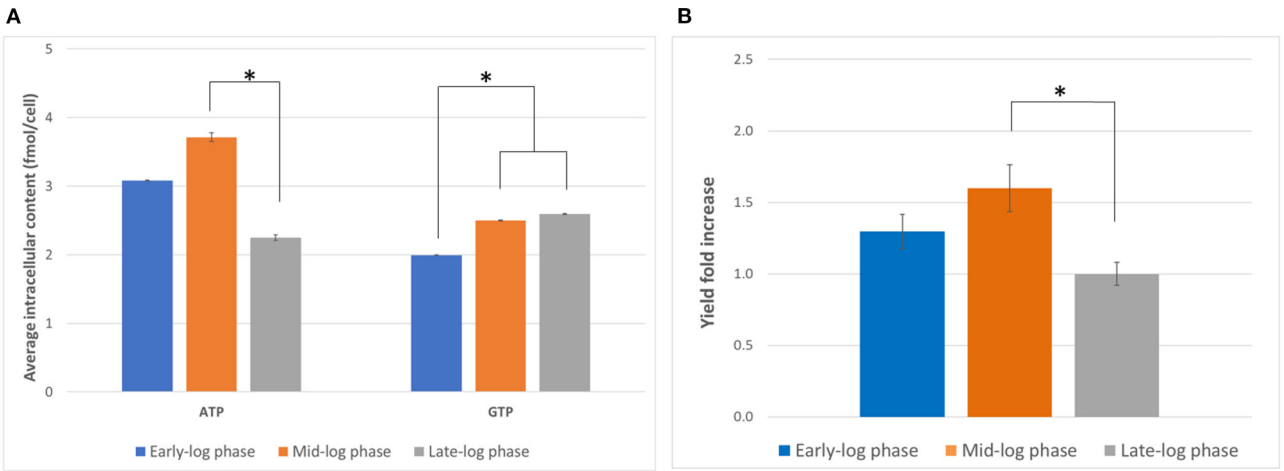


FIGURE 1 | Determination of optimal harvest time. **(A)** Intracellular ATP and GTP content (in fmol/cell) on days 3, 4, and 5 (early-, mid-, and late-log phase, respectively) of culture. Error bars represent the standard deviation of the mean of three samples taken from independent culture flasks. **(B)** GFP expression using lysates harvested and prepared on the three harvesting days. The positive control plasmid, the accessory proteins, and the reaction mixtures from the 1-Step Human Coupled IVT Kit were used to facilitate comparison. Error bars represent the standard deviation of the mean of three independent CFPS reactions (*$p < 0.05$). Early-log phase: day 3, mid-log phase: day 4, late-log phase: day 5. Measurements were corrected for a negative control (NC, no plasmid addition).

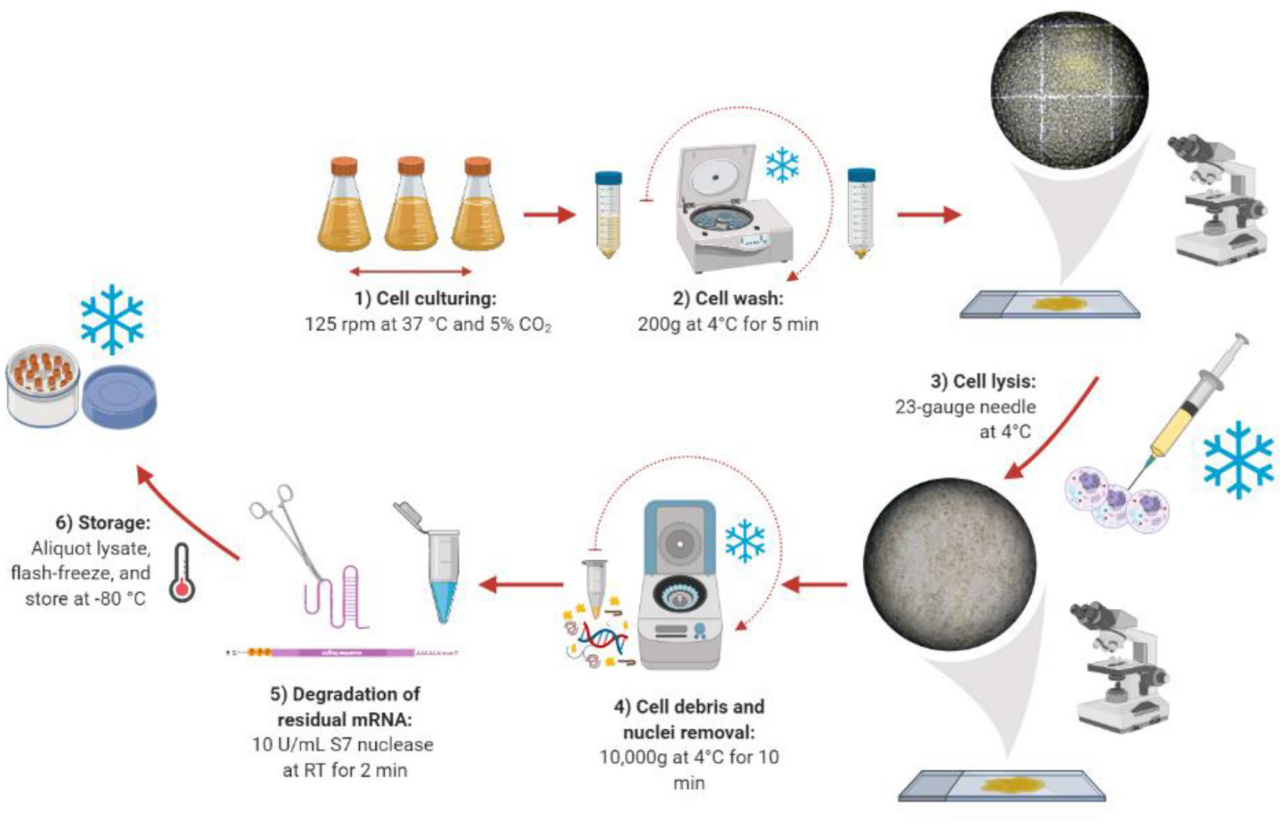


FIGURE 2 | Lysate preparation workflow. Created with Biorender (biorender.com).

proteins to increase CFPS yields. One of the bottlenecks in eukaryotic CFPS is the phosphorylation of the eukaryotic translation initiation factor 2 (eIF2), which leads to reduced translation initiation over time (Zeenko et al., 2008). To counteract this, reactions can be supplemented with two accessory proteins, GADD34 and K3L, which have been shown

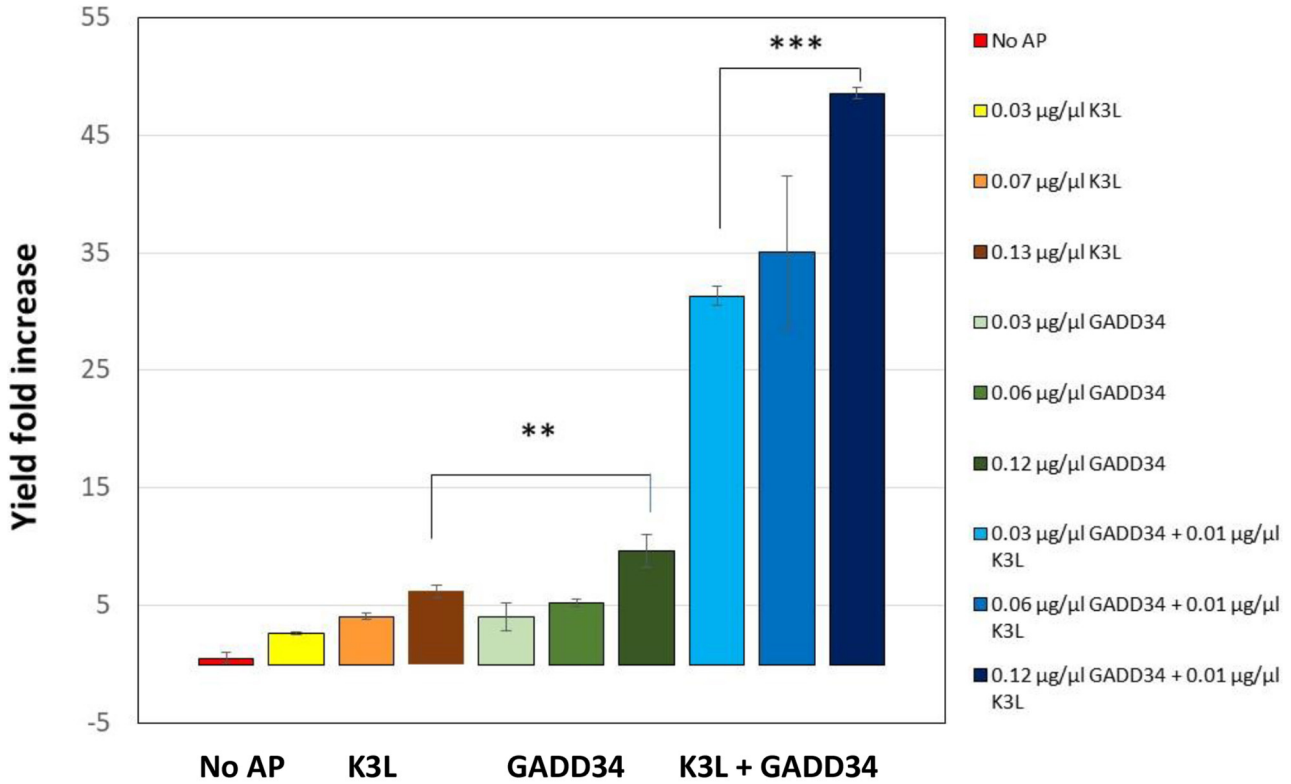


FIGURE 3 | Addition of the accessory proteins GADD34 and K3L results significant increase in expression yields compared to the yield attained in their absence. GADD34 and K3L were separately and added in combination at different concentrations to the CFPS reactions. No AP refers to the CFPS reaction yield in the absence of accessory proteins, which corresponds to an average yield of ∼0.4 μg/ml (range 0.05–1.08 μg/ml). Yields are normalized to individual reactions without accessory proteins run concurrently. Error bars represent the standard deviation of the mean of three independent CFPS reactions using the same lysate (**$p < 0.01$, ***$p < 0.001$).

to increase the activity of human cell-derived CFPS systems, but had not previously been tested in CHO CFPS (Mikami et al., 2006a, 2010b; Burgenson et al., 2018). GADD34 is a regulator that recruits protein phosphatase 1, leading to dephosphorylation of eIF2α (Novoa et al., 2001), and K3L is a viral protein that acts as a pseudo-substrate, reducing phosphorylation of eIF2α (Carroll et al., 1993).

In the first iteration, we tested the addition of different concentrations of full-length GADD34 and K3L to our system. Both proteins were recombinantly expressed in *E.coli* cells and then purified before addition (Hinnebusch, 2000). Active tGFP was produced in CFPS reactions without accessory protein addition, but yields were low and varied widely from reaction-to-reaction (average yields of ∼0.4 μg/mL, range 0.05–1.08 μg/mL). Supplementing CFPS with either accessory protein alone or in combination led to an increase in tGFP yield over the non-supplemented control (**Figure 3**). Increasing the concentration of the accessory proteins, led to further increases in yield, presumably because of a greater degree of de-phosphorylation of the translational machinery leading to enhanced translation initiation. The concentrations tested herein were not sufficiently high to reach a plateau in tGFP expression. However, they could not be further increased due to the poor expression yields of the accessory proteins, which limited the amount of protein available for testing. There was a strong synergistic effect of adding both GADD34 and K3L that resulted in a ∼50-fold increase in tGFP expression levels compared to tGFP yields attained without the addition of accessory proteins. The synergy is most likely due to the different roles that K3L and GADD34 play in reducing EIF2α inactivation. Interestingly, these results differ from those reported for HeLa extracts, where no synergy was noted (Mikami et al., 2006a). Our results further indicate that the effect of addition of GADD34 alone has a larger impact (∼10-fold increase in yield) than the addition of K3L alone (∼six-fold increase). This could be due to the active role of GADD34 in the recruitment of a phosphatase to dephosphorylate the translation initiation machinery, compared to the more passive role of K3L as a pseudo-substrate. With the addition of both accessory proteins, yields of tGFP approached 20 μg/mL.

Truncation of GADD34 Improves Its Recovery Without Compromising Its Effect

Full-length GADD34 was difficult to express and purify in the quantities that were required for CFPS supplementation. However, previous work has shown that truncating membrane domain of GADD34 leads to improved expression in *E. coli*

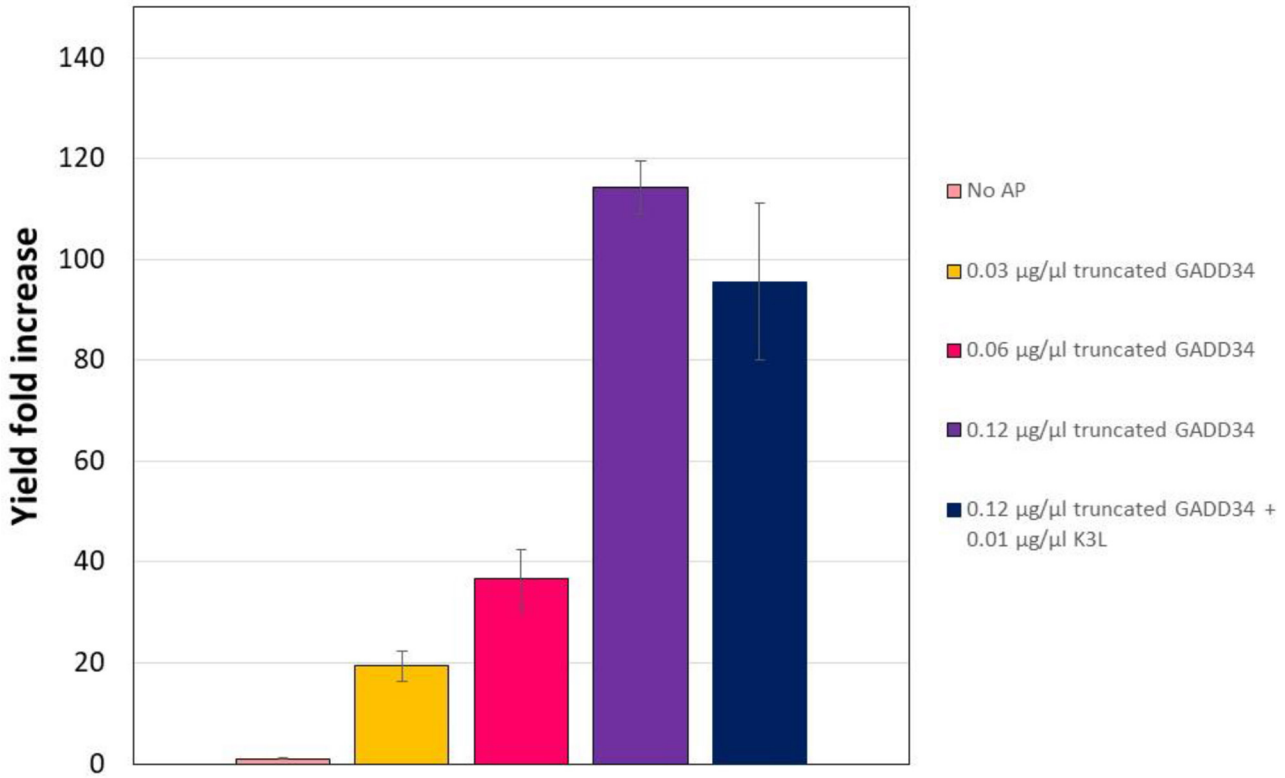


FIGURE 4 | Truncation does not compromise the yield effect of GADD34 on CFPS yield. The negative control refers to the CFPS reaction yield in the absence of accessory proteins, which corresponds to an average yield of 0.13 µg/ml. Error bars represent the standard deviation of the mean of three independent CFPS reactions.

without loss of activity (Mikami et al., 2010a). Therefore, we expressed a truncated version of GADD34 (GADD34Δ1–240) and repeated the supplementation experiments. Overall, *E. coli* expressing GADD34Δ1–240 grew to a higher optical density during the expression, suggesting a reduced expression burden, with a ∼1.4-fold increase in optical density (OD) at harvest (GADD34: 2.11, truncated GADD34: 2.96, **Supplementary Figure 3**). As a result, there was a four-fold increase in the amount of GADD34Δ1–240 obtained after purification.

In the CHO CFPS reaction, supplementation of GADD Δ1–240 had a similar effect on protein yield as the full-length form, with increases in yield of 114-fold at the highest concentration supplemented (**Figure 4**). In order to assess if the truncation compromised the function, GADD Δ1–240 was tested as the same concentrations as the full-length version. However, due to variations in the control reaction without accessory protein, overall yields were lower at ∼14 µg/mL. Furthermore, in contrast to the results with the full-length GADD34, there was no synergistic effect when both GADD34Δ1–240 and K3L were supplemented to the reactions.

The reason for the decreased effect of GADD Δ1–240 and K3L co-addition is not fully understood. However, it could be related to the lower purity level of the GADD34Δ1–240 protein stock produced in *E. coli* (**Supplementary Figure 4**). More impurities were observed for the truncated version than for the non-truncated version, which likely resulted in an overestimation of the truncated GADD34 stock concentration. It has been previously observed that the deletion of the N-terminal region improves the stability of the truncated protein, but increases the amount of co-purified bacterial proteins (Mikami et al., 2010a).

Transient Expression of N-Terminally Truncated GADD34 in CHO-S Cells Increases CFPS Yields

Expression and purification of the accessory proteins requires additional processing steps. Therefore, we next explored the potential of transiently expressing the accessory proteins in CHO cells prior to the generation of the lysate as an alternative method that would reduce handling time and resources. GADD34 showed a greater effect on expression yields than K3L and GADD34Δ1–240 expresses at higher levels. Therefore, we focused on transient expression of GADD34Δ1–240. The GADD34Δ1–240 gene was cloned into the mCherry2-C1 plasmid in-frame with an N-terminal mCherry tag to provide a fluorescent fusion protein that was easy to detect. Cells were

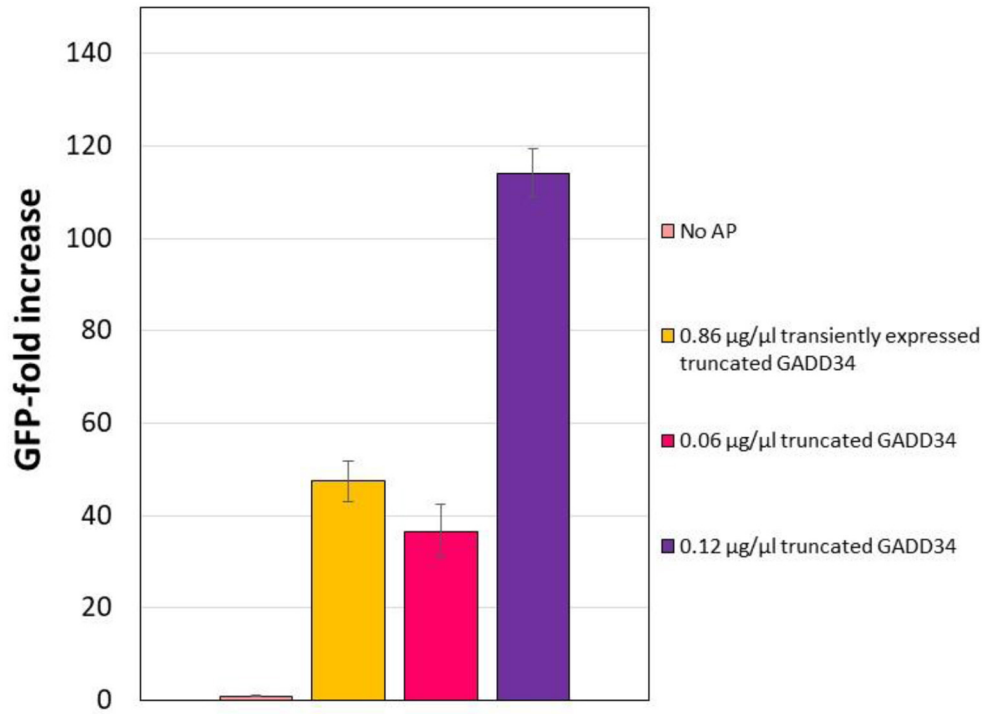


FIGURE 5 | Effect of transiently expressing N-terminally truncated GADD34 in CHO cells before lysate preparation. In all cases, the negative control refers to the CFPS reaction yield in the absence of accessory proteins, which corresponds to an average yield of ∼0.07 μg/ml. Error bars represent the standard deviation of the mean of three independent CFPS reactions.

harvested 3 days after transient transfection and the mCherry-GADD34Δ1–240 fusion protein concentration in the lysate was estimated to be ∼0.18 μg/μL (**Supplementary Figure 5**). Using this lysate in a CFPS reaction results in a final concentration of mCherry-GADD34Δ1–240 of 0.086 μg/μL, which is comparable to the middle concentration added in supplementation experiments (0.06 μg/μL). Overall, expressing GADD34Δ1–240 in the cells prior to the generation of the lysate resulted in a 47.5-fold increase in tGFP expression relative to the negative control in which no GADD34 was expressed (**Figure 5**). However, the total tGFP expression was lower in lysates made from cells transiently expressing GADD34Δ1–240 compared to those where purified GADD34 Δ1–240 was supplemented (3.5 vs. 4.8 μg/ml, respectively). We believe this is likely due to the negative effects of the transient transfection conditions on the energetic state of the cell, although we cannot rule out that the mCherry fusion leads to reduced activity of GADD34. The cell extract from the transient transfection was harvested on day 3, due to lower viability and it is well-known that the polyethyleneimine used for transient transfection reduces culture viability and may also decrease the activity level of protein machinery for CFPS. The negative effect of PEI limits the application of this approach, although stable expression of the accessory proteins in CHO extracts remains to be explored in the future. Furthermore, it may be useful to test other transient transfection methods that may have a lower impact on the physiology of the cells.

DISCUSSION

Mammalian CFPS are useful tools for high throughput screening and production of proteins requiring post-translational modifications. In this paper, we shared our detailed protocols for the design, development and optimization of a functional mammalian cell-free protein synthesis platform derived from cultured CHO cells. Among the required ingredients for a CFPS reaction, the cell lysate is the most labor-intensive component to prepare, requiring careful handling to protect functionality of the protein synthesis machinery. Furthermore, as the productivity of CFPS platforms depends highly on translation initiation, it is an important to enhance this step for developing robust cell-free platforms. Therefore, the focus of this study was to find the optimum method for active cell lysate preparation and to increase the protein synthesis yields via addition of accessory proteins to combat the phosphorylation of the translation machinery. Three options for increasing the expression of proteins by supplementing the reactions with accessory proteins that decrease the phosphorylation of the translational machinery were demonstrated. In general, there is a tradeoff between the effect of expression yields and the ease of implementation. Expressing the accessory proteins in a heterologous host and then supplementing them into the reaction led to higher tGFP yields overall, but is a more cumbersome protocol, which requires extra steps to purify and concentrate the accessory proteins. On the other hand, transiently expressing GADD34Δ1–240 still

increases the expression of tGFP, albeit to a lesser degree due to reduced cell viability, and is much simpler. In the future, generating stable cell lines expressing the accessory proteins may provide a route for the facile generation of lysates with increased activity. By explaining our methods including all details of the developed protocols, we hope to improve the understanding of bottlenecks in CHO CFPS systems. We aim to spur greater participation in mammalian CFPS, both in the cell-free community and in bioprocessing, and hope to offer a powerful and more flexible alternative to using the available commercial kits. We believe a collaborative approach is needed to make mammalian CFPS a more widely used research tool and establish CHO CFPS as a rapid production and screening platform for therapeutic development.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

CH: conceptualization, methodology, investigation, data curation, writing- original draft preparation, visualization. GB: investigation, writing–reviewing and editing. IM-R: investigation, data curation, writing–reviewing and editing. OC: writing–reviewing and editing. CK and KP: conceptualization, supervision, writing–reviewing and editing. All authors read and approved the final manuscript.

FUNDING

CH was funded by the Marit Mohn Westlake Foundation. GB and IM-R were funded by the Engineering and Physical Sciences Research Council UK (EP/K038648/1, EP/R511547/1).

ACKNOWLEDGMENTS

We would like to thank the EPSRC Impact Acceleration Account (EP/R511547/1) for funding. We thank Prof. Hiroaki Imataka for sharing his expression constructs of K3L, GADD34 and truncated GADD34 and Michael Davidson for creating the mCherry2-C1 plasmid. We also thank Dr. Rochelle Aw for sharing her expert knowledge in eukaryotic CFPS systems.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2020.604091/full#supplementary-material

REFERENCES

Aw, R., and Polizzi, K. M. (2019). Biosensor-assisted engineering of a high-yield Pichia pastoris cell-free protein synthesis platform. *Biotechnol. Bioeng.* 116, 656–666. doi: 10.1002/bit.26901

Brödel, A. K., Sonnabend, A., and Kubick, S. (2014). Cell-free protein expression based on extracts from CHO cells. *Biotechnol. Bioeng.* 111, 25–36. doi: 10.1002/bit.25013

Brödel, A. K., Sonnabend, A., Roberts, L. O., Stech, M., Wüstenhagen, D. A., and Kubick, S. (2013). IRES-mediated translation of membrane proteins and glycoproteins in eukaryotic cell-free systems. *PLoS ONE* 8:e82234. doi: 10.1371/journal.pone.0082234

Brödel, A. K., Wustenhagen, D. A., and Kubick, S. (2015). Cell-free protein synthesis systems derived from cultured mammalian cells. *Methods Mol. Biol.* 1261, 129–140. doi: 10.1007/978-1-4939-2230-7_7

Buntru, M., Vogel, S., Spiegel, H., and Schillberg, S. (2014). Tobacco BY-2 cell-free lysate: an alternative and highly-productive plant-based *in vitro* translation system. *BMC Biotechnol.* 14:37. doi: 10.1186/1472-6750-14-37

Buntru, M., Vogel, S., Stoff, K., Spiegel, H., and Schillberg, S. (2015). A versatile coupled cell-free transcription-translation system based on tobacco BY-2 Cell Lysates. *Biotechnol. Bioeng.* 112, 867–878. doi: 10.1002/bit.25502

Burgenson, D., Gurramkonda, C., Pilli, M., Ge, X. D., Ander, A., Kostov, Y., et al. (2018). Rapid recombinant protein expression in cell-free extracts from human blood. *Sci. Rep.* 8:7. doi: 10.1038/s41598-018-27846-8

Carroll, K., Elroy-Stein, O., Moss, B., and Jagus, R. (1993). Recombinant vaccinia virus K3L gene product prevents activation of double-stranded RNA-dependent, initiation factor 2 alpha-specific protein kinase. *J. Biol. Chem.* 268, 12837–12842. doi: 10.1016/S0021-9258(18)31463-7

Caschera, F., and Noireaux, V. (2014). Synthesis of 2.3 mg/ml of protein with an all Escherichia coli cell-free transcription-translation system. *Biochimie* 99, 162–168. doi: 10.1016/j.biochi.2013.11.025

del Val, I. J., Kyriakopoulos, S., Polizzi, K. M., and Kontoravdi, C. (2013). An optimized method for extraction and quantification of nucleotides and nucleotide sugars from mammalian cells. *Anal. Biochem.* 443, 172–180. doi: 10.1016/j.ab.2013.09.005

Ezure, T., Suzuki, T., Higashide, S., Shintani, E., Endo, K., and Kobayashi, S.-,i, et al. (2006). Cell-free protein synthesis system prepared from insect cells by freeze-thawing. *Biotechnol. Prog.* 22, 1570–1577. doi: 10.1021/bp060110v

Ezure, T., Suzuki, T., Shikata, M., Ito, M., Ando, E., Utsumi, T., et al. (2010). Development of an insect cell-free system. *Curr. Pharm. Biotechnol.* 11, 279–284. doi: 10.2174/138920110791111997

Gan, R., and Jewett, M. C. (2014). A combined cell-free transcription-translation system from Saccharomyces cerevisiae for rapid and robust protein synthe. *Biotechnol. J.* 9, 641–651. doi: 10.1002/biot.201300545

Gregorio, N. E., Levine, M. Z., and Oza, J. P. (2019). A user's guide to cell-free protein synthesis. *Methods Protocols* 2:24. doi: 10.3390/mps2010024

Hinnebusch, A. G. (2000). 5 mechanism and regulation of initiator methionyl-tRNA binding to ribosomes. *Cold Spring Harbor Monograph Arch.* 39, 185–243. doi: 10.1101/0.185-243

Hodgman, C. E., and Jewett, M. C. (2013). Optimized extract preparation methods and reaction conditions for improved yeast cell-free protein synthesis. *Biotechnol. Bioeng.* 110, 2643–2654. doi: 10.1002/bit.24942

Jackson, R. J., and Hunt, T. (1983). Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA. *Meth. Enzymol.* 96, 50–74. doi: 10.1016/S0076-6879(83)96008-1

Jewett, M. C., Miller, M. L., Chen, Y., and Swartz, J. R. (2009). Continued protein synthesis at low [ATP] and [GTP] enables cell adaptation during energy limitation. *J. Bacteriol.* 191, 1083–1091. doi: 10.1128/JB.00852-08

Katzen, F., and Kudlicki, W. (2006). Efficient generation of insect-based cell-free translation extracts active in glycosylation and signal sequence processing. *J. Biotechnol.* 125, 194–197. doi: 10.1016/j.jbiotec.2006.03.002

Kobs, G. (2008). Selecting the cell-free protein expression system that meets your experimental goals. *Cell Notes* 21, 6–9. Available online at: https://www.promega.com/-/media/files/resources/cell-notes/cn021/selecting-the-cell-free-protein-expression-system-that-meets-your-experimental-goals.pdf?la=en

Madin, K., Sawasaki, T., Ogasawara, T., and Endo, Y. (2000). A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 559–564. doi: 10.1073/pnas.97.2.559

Madono, M., Sawasaki, T., Morishita, R., and Endo, Y. (2011). Wheat germ cell-free protein production system for post-genomic research. *N. Biotechnol.* 28, 211–217. doi: 10.1016/j.nbt.2010.08.009

Martin, R. W., Majewska, N. I., Chen, C. X., Albanetti, T. E., Jimenez, R. B. C., Schmelzer, A. E., et al. (2017). Development of a CHO-based cell-free platform for synthesis of active monoclonal antibodies. *ACS Synth. Biol.* 6, 1370–1379. doi: 10.1021/acssynbio.7b00001

Mikami, S., Kobayashi, T., and Imataka, H. (2010b). "Cell-free protein synthesis systems with extracts from cultured human cells," in *Cell-Free Protein Production: Methods and Protocols*. eds Y Endo, K. Takai, T. Ueda (Totowa, NJ: Humana Press), 43–52. doi: 10.1007/978-1-60327-331-2_5

Mikami, S., Kobayashi, T., Machida, K., Masutani, M., Yokoyama, S., and Imataka, H. (2010a). N-terminally truncated GADD34 proteins are convenient translation enhancers in a human cell-derived *in vitro* protein synthesis system. *Biotechnol. Lett.* 32, 897–902. doi: 10.1007/s10529-010-0251-7

Mikami, S., Kobayashi, T., Masutani, M., Yokoyama, S., and Imataka, H. (2008). A human cell-derived *in vitro* coupled transcription/translation system optimized for production of recombinant proteins. *Protein Expr. Purif.* 62, 190–198. doi: 10.1016/j.pep.2008.09.002

Mikami, S., Kobayashi, T., Yokoyama, S., and Imataka, H. (2006a). A hybridoma-based *in vitro* translation system that efficiently synthesizes glycoproteins. *J. Biotechnol.* 127, 65–78. doi: 10.1016/j.jbiotec.2006.06.018

Mikami, S., Masutani, M., Sonenberg, N., Yokoyama, S., and Imataka, H. (2006b). An efficient mammalian cell-free translation system supplemented with translation factors. *Protein Expr. Purif.* 46, 348–357. doi: 10.1016/j.pep.2005.09.021

Murota, K., Hagiwara-Komoda, Y., Komoda, K., Onouchi, H., Ishikawa, M., and Naito, S. (2011). Arabidopsis cell-free extract, ACE, a new *in vitro* translation system derived from arabidopsis callus cultures. *Plant Cell Physiol.* 52, 1443–1453. doi: 10.1093/pcp/pcr080

Nirenberg, M. W., and Matthaei, J. H. (1961). The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 47, 1588–1602. doi: 10.1073/pnas.47.10.1588

Novoa, I., Zeng, H., Harding, H. P., and Ron, D. (2001). Feedback inhibition of the unfolded protein response by GADD34-mediated dephosphorylation of eIF2alpha. *J. Cell Biol.* 153, 1011–1022. doi: 10.1083/jcb.153.5.1011

Richter, C., Bickel, F., Osberghaus, A., and Hubbuch, J. (2014). High-throughput characterization of an insect cell-free expression. *Eng. Life Sci.* 14, 409–417. doi: 10.1002/elsc.201300118

Shields, D., and Blobel, G. (1978). Efficient cleavage and segregation of nascent presecretory proteins in a reticulocyte lysate supplemented with microsomal membranes. *J. Biol. Chem.* 253, 3753–3756. doi: 10.1016/S0021-9258(17)34748-8

Starr, C. M., and Hanover, J. A. (1990). Glycosylation of nuclear pore protein p62. Reticulocyte lysate catalyzes O-linked N-acetylglucosamine addition *in vitro*. *J. Biol. Chem.* 265, 6868–6873. doi: 10.1016/S0021-9258(19)39229-4

Stavnezer, J., and Huang, R. C. (1971). Synthesis of a mouse immunoglobulin light chain in a rabbit reticulocyte cell-free system. *Nature New Biol.* 230, 172–176. doi: 10.1038/newbio230172a0

Stech, M., Merk, H., Schenk, J. A., Stöcklein, W. F. M., Wüstenhagen, D. A., Micheel, B., et al. (2013). Production of functional antibody fragments in a vesicle-based eukaryotic cell-free translation system. *J. Biotechnol.* 164, 220–231. doi: 10.1016/j.jbiotec.2012.08.020

Stech, M., Nikolaeva, O., Thoring, L., Stöcklein, W. F. M., Wüstenhagen, D. A., Hust, M., et al. (2017). Cell-free synthesis of functional antibodies using a coupled *in vitro* transcription-translation system based on CHO cell lysates. *Sci. Rep.* 7:12030. doi: 10.1038/s41598-017-12364-w

Stech, M., Quast, R. B., Sachse, R., Schulze, C., Wüstenhagen, D. A., and Kubick, S. (2014). A continuous-exchange cell-free protein synthesis system based on extracts from cultured insect cells. *PLoS ONE* 9:e96635. doi: 10.1371/journal.pone.0096635

Thoring, L., Dondapati, S. K., Stech, M., Wüstenhagen, D. A., and Kubick, S. (2017). High-yield production of "difficult-to-express" proteins in a continuous exchange cell-free system based on CHO cell lysates. *Sci. Rep.* 7:11710. doi: 10.1038/s41598-017-12188-8

Thoring, L., and Kubick, S. (2018). Versatile cell-free protein synthesis systems based on chinese hamster ovary cells. *Methods Mol. Biol.* 1850, 289–308. doi: 10.1007/978-1-4939-8730-6_19

Yadavalli, R., and Sam-Yellowe, T. (2015). HeLa based cell free expression systems for expression of plasmodium rhoptry proteins. *J. Vis. Exp.* 2015:e52772. doi: 10.3791/52772

Zeenko, V. V., Wang, C., Majumder, M., Komar, A. A., Snider, M. D., Merrick, W. C., et al. (2008). An efficient *in vitro* translation system from mammalian cells lacking the translational inhibition caused by eIF2 phosphorylation. *RNA* 14, 593–602. doi: 10.1261/rna.825008

Zemella, A., Thoring, L., Hoffmeister, C., Šamalíková, M., Ehren P, Wüstenhagen, D. A., et al. (2018). Cell-free protein synthesis as a novel tool for directed glycoengineering of active erythropoietin. *Sci. Rep.* 8:8514. doi: 10.1038/s41598-018-26936-x

Zhang, L., Liu, W.-Q., and Li, J. (2020). Establishing a eukaryotic pichia pastoris cell-free protein synthesis system. *Front. Bioeng. Biotechnol.* 8:536. doi: 10.3389/fbioe.2020.00536

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

REVIEW
published: 20 July 2021
doi: 10.3389/fbioe.2021.715328

Effective Use of Linear DNA in Cell-Free Expression Systems

*Megan A. McSweeney and Mark P. Styczynski**

Georgia Institute of Technology, School of Chemical & Biomolecular Engineering, Atlanta, GA, United States

Cell-free expression systems (CFEs) are cutting-edge research tools used in the investigation of biological phenomena and the engineering of novel biotechnologies. While CFEs have many benefits over *in vivo* protein synthesis, one particularly significant advantage is that CFEs allow for gene expression from both plasmid DNA and linear expression templates (LETs). This is an important and impactful advantage because functional LETs can be efficiently synthesized *in vitro* in a few hours without transformation and cloning, thus expediting genetic circuit prototyping and allowing expression of toxic genes that would be difficult to clone through standard approaches. However, native nucleases present in the crude bacterial lysate (the basis for the most affordable form of CFEs) quickly degrade LETs and limit expression yield. Motivated by the significant benefits of using LETs in lieu of plasmid templates, numerous methods to enhance their stability in lysate-based CFEs have been developed. This review describes approaches to LET stabilization used in CFEs, summarizes the advancements that have come from using LETs with these methods, and identifies future applications and development goals that are likely to be impactful to the field. Collectively, continued improvement of LET-based expression and other linear DNA tools in CFEs will help drive scientific discovery and enable a wide range of applications, from diagnostics to synthetic biology research tools.

Keywords: cell-free expression, linear expression template, nuclease inhibition, genetic circuits, rapid prototyping, DNA aptamers

OPEN ACCESS

Edited by:
Yong-Chan Kwon,
Louisiana State University,
United States

Reviewed by:
Cheemeng Tan,
University of California, Davis,
United States
Nigel Reuel,
Iowa State University, United States

****Correspondence:***
Mark P. Styczynski
mark.styczynski@chbe.gatech.edu

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

Received: *26 May 2021*
Accepted: *06 July 2021*
Published: *20 July 2021*

Citation:
McSweeney MA and Styczynski MP (2021) Effective Use of Linear DNA in Cell-Free Expression Systems. Front. Bioeng. Biotechnol. 9:715328. doi: 10.3389/fbioe.2021.715328

INTRODUCTION

Cell-free expression systems (CFEs) are powerful tools for the execution of arbitrary genetic programs or the synthesis of proteins *in vitro*. One of the most common and affordable forms of CFE, the lysate-based system, is composed of a crude cellular extract (typically from *E. coli*, but lysates from other organisms are useful for specific applications) combined with supplemented cofactors and substrates essential for transcription and translation. CFEs offer several advantages over the use of whole-cell *in vivo* systems (Khambhati et al., 2019). Transport limitations inherent to whole-cell systems due to the cell membrane are reduced in CFEs because they have no membrane, yielding improved control over plasmid dosage, pH, and inducer levels (Swartz 2006). CFEs also eliminate cellular toxicity issues that sometimes arise *in vivo* from the expression of certain proteins, with the additional effect of avoiding plasmid instability often caused by toxicity (Katzen et al., 2005). CFEs have expedited research on key biological principles (Nirenberg and Leder 1964) and have been applied in contexts ranging from industrially relevant large-scale protein production (Zawada et al., 2011) to upstream screening for glycoprotein synthesis (Schoborg et al., 2018). CFEs are particularly

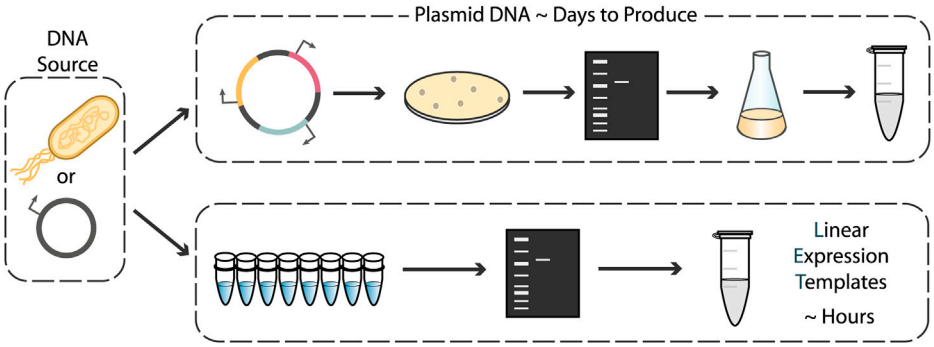


FIGURE 1 | Strategies for preparing DNA templates for use in cell-free expression systems. Traditional plasmid cloning protocols involving construct assembly, transformation, screening, and plasmid purification take days to complete. Linear expression templates can be made via PCR from genomic DNA or plasmid templates and used directly after amplicon verification and purification, drastically reducing DNA template preparation time.

attractive for sensor development, with applications ranging from environmental sensors (Verosloff et al., 2019; Thavarajah et al., 2020) to user-friendly biomedical diagnostics (McNerney et al., 2019).

The configuration of CFEs as a membrane-less solution of expression machinery, compared to the membrane-compartmentalized individual cells of *in vivo* systems, not only reduces the transport limitations that would have been imposed by the membrane, but also makes the use of linear expression templates (LETs) a viable option for CFEs. For *in vivo* systems, the use of plasmids is required for stable expression without genomic integration. Since delivery of DNA into cells happens with low efficiency, successfully transformed cells must be selected for and then expanded, meaning that the DNA vector must replicate during cell growth to avoid growth-associated dilution and loss. Plasmid construction requires cloning, *in vivo* synthesis, and plasmid isolation, a process that takes days for each new construct. However, LETs—which typically consist of a promoter region, gene coding sequence, and transcriptional terminator—can be quickly and easily produced *in vitro* via polymerase chain reaction (PCR) from existing plasmid DNA or genomic DNA (**Figure 1**). With techniques such as Golden Gate assembly, multiple LETs can be rapidly assembled into complex constructs entirely *in vitro*. Using PCR products as expression templates rather than plasmids can decrease the "primers-to-testable-DNA" time from days to only a few hours. This can facilitate high-throughput screening and significantly accelerate the prototyping cycle time of multicomponent genetic circuits to a standard business day (Sun et al., 2013).

Additionally, LETs allow for expression of toxic genes in CFEs that otherwise would be difficult to clone into a plasmid. The LET construct for the toxic gene can be amplified directly from genomic DNA or from a plasmid without a promoter upstream of the gene. Without a promoter, the plasmid containing the toxic gene can be cloned because the gene will not be expressed *in vivo* as normal. This can then be used as a template for a LET, with the promoter sequence added via primers during PCR.

Despite the numerous benefits of LETs, plasmids remain the most widely used DNA template in CFEs due to their resistance to degradation. DNA nucleases native to *E. coli* are present in the crude cellular lysate and remain after lysate purification. These nucleases readily digest linear double-stranded (ds) or single-stranded (ss) DNA fragments in the reaction to an extent not observed for circular plasmids, causing LETs to have a much shorter half-life than plasmids. This leads to much lower protein yields or diminished function of genetic circuits, which has in turn slowed adoption of LETs for CFEs.

However, there exist several approaches for stabilizing linear DNA that lessen the impacts of these phenomena and allow for the effective use of LETs in crude lysate-based CFEs. In this review, we will describe these advances and compare their effectiveness and limitations. Then, we will summarize select applications where LETs have been used to expedite circuit prototyping cycles, rapidly screen synthetic regulators, express toxic proteins, and more. Lastly, we will discuss the restrictions of the existing nuclease inhibition strategies, recognize areas with critical need for improvement, and identify future development goals.

APPROACHES FOR STABILIZING LINEAR DNA IN CELL-FREE EXPRESSION SYSTEMS

While crude bacterial lysate is widely used for CFEs, purified recombinant proteins can also constitute the basis for CFEs. PURExpress (NEB), PUREfrex 2.0, and Magic PURE are three commercially available recombinant systems based on the original PURE system developed by Shimizu et al. (2001) (See also Laohakunakom et al. (2020)). Purchasing these systems from commercial vendors saves valuable laboratory time compared to typical modern protocols for in-house preparation of crude cell extract that typically take three days and over 10 h of active labor (Kwon and Jewett 2015). Perhaps more importantly, these systems have minimal nuclease and protease activity compared to crude lysates, making expression from LETs much easier than in lysates.

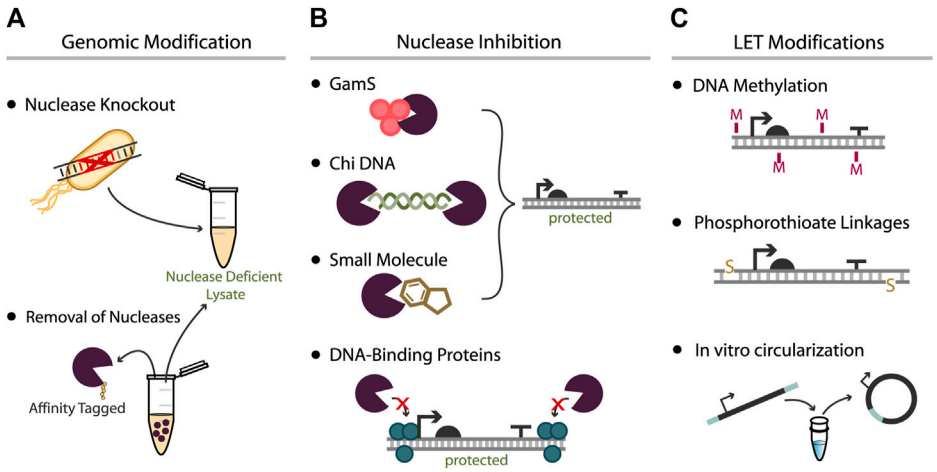


FIGURE 2 | Approaches used to increase expression yield from LETs. **(A)** Bacteria can be genetically modified prior to lysate extraction either by deletion of nuclease genes or by the fusion of affinity peptides to nuclease genes for later removal during lysate processing. Both methods aim to produce a lysate with negligible nuclease activity. **(B)** Nuclease inhibitors can be added to cell-free reactions to mitigate activity of specific nucleases. This effect can be achieved with direct RecBCD inhibitors such as GamS, Chi DNA, and certain small molecules, or via the addition of DNA-binding proteins that interact with the LETs. **(C)** The LET can also be modified to better protect the construct from nuclease degradation: methylation of PCR-generated LETs can mimic the chemistry of native DNA, phosphorothioate linkages can be added to the ends of LETs via modified primers, or amplicons with appropriately designed primers can be recircularized prior to CFE.

TABLE 1 | Effectiveness of different nuclease inhibition strategies quantified by their ability to improve LET-based expression in *E. coli* CFEs. Values with an asterisk were inferred from figures in the corresponding reference.

LET stabilization approach		Improvement	Metric	References
Genomic modifications	*ΔrecCBD::P_{lac}-red-kan-ΔendA*	3–6x	Fold change from WT strain	Michel-Reydellet et al. (2005)
	Affinity tag removal of RecD and PNPase	4x	Fold change from WT strain	Seki et al. (2009)
Nuclease inhibition	GamS	37.6%	Percentage of plasmid expression	Sun et al. (2013)
	Chi DNA	23%*	Percentage of plasmid expression	Marshall et al. (2017)
	Small molecule RecBCD inhibitors	250%* (CID 697851) 300%* (CID 1517823)	Percent increase from no inhibitor	Shrestha et al. (2014)
	ssCro	23%	Percentage of plasmid expression	Zhu et al. (2020)
	Ku	8%*	Percentage of plasmid expression	Yim et al. (2020)
LET modifications and enhanced design	DNA methylation	32% (*dam* methyltransferase) −18% (CpG methyltransferase)	Percent increase from unmethylated LET	Zhu et al. (2020)
	Terminal phosphorothioate (PT) linkages (x2)	36%*	Percent increase from unmodified LET	Sun et al. (2013)
	3′-tail mRNA secondary structures	92%* (poly(G) tail) 265%* (T7 terminator)	Percent increase from LETs lacking 3′ secondary structures	Ahn et al. (2005)

However, crude lysate-based CFEs still dominate the field based in part on their affordability. For example, the PURExpress *in vitro* protein synthesis kit costs $0.35-$0.65 per 1 uL reaction while a crude lysate-based system costs $0.02-$0.04 per 1 uL reaction (Pardee et al., 2014). Recently, robust methods have been reported for in-house production of recombinant protein CFEs by coculturing and inducing expression of all 36 proteins present in the commercial systems (Lavickova and Maerkl 2019). This "OnePot" method can achieve high protein yields at a cost closer to $0.09 per μL. While this is significantly lower than the cost of the PURExpress system, it is still more expensive than crude lysate CFEs and laborious in its own way; it remains to be seen whether this approach will be adopted widely.

The financial benefits of lysate-based CFEs, among other advantages, have motivated substantial effort toward addressing one of its major shortcomings: the effects of nucleases in crude extracts. Significant linear DNA degradation is attributed to exonuclease V, the product of the *recBCD* operon. As a result, previous efforts have attempted to remove, inhibit, or deter RecBCD activity on LETs (**Figure 2**), with varying success. The relative effectiveness of each approach is listed in **Table 1**. The most common approaches to achieve

these goals include genomic removal of nucleases, nuclease inhibition, and protective linear DNA modifications.

Genomic Modifications

In *E. coli*, endonucleases and exonucleases are critical for proper cell function and are essential enzymes involved in double-stranded break repair and recombination events. The predominant endonuclease in *E. coli* is endonuclease I, encoded by the *endA* gene. The dominant source of exonuclease activity in *E. coli* comes from exonuclease V, the product of the *recBCD* operon (Kuzminov and Stahl 1997). In this complex, RecB has both helicase and nuclease activity, and RecD serves only as a helicase. Attempts have been made to engineer *E. coli* mutants that lack the activity of one or several of these subunits (**Figure 2A**). For example, a cell extract made from an A19Δ*recD*Δ*endA* mutant was used for cell-free protein synthesis and tested with LETs for nuclease activity (Michel-Reydellet et al., 2005). Plasmid-based expression was slightly lower in the mutant extract compared to the wild type, suggesting that the *endA* deletion did not improve plasmid stability. However, extracts from both strains possessed significant linear DNA degradation as evidenced by low protein yields from the LETs. And while these results showed almost equal expression from plasmids between strains, others have reported that extracts of exonuclease-deficient strains exhibit significant loss in translational activity (Ahn et al., 2005).

Knockout strains that lack the entire *recBCD* operon have also been explored. Complete *recBCD* knockouts create strains that are extremely slow-growing, cannot support recombination, and do not allow for replication of many plasmids (Yu et al., 2000). To address this, *recBCD* can be replaced with the Red recombination system from bacteriophage λ to remove exonuclease activity while allowing recombination and adequate growth rates (Murphy 1991; Michel-Reydellet et al., 2005). Lysates made from this mutant yielded over three times more protein from PCR-generated LETs compared to the wild-type extract.

Since the growth defects of these knockout strains make the culture steps in lysate preparation more time-consuming, removing nucleases from the lysate post-harvest has been explored as an alternative option. Insertion of streptavidin-binding peptide tag sequences in the 3' termini of RecD and the gene encoding polynucleotide phosphorylase (PNPase) was shown not to effect *E. coli* growth (Seki et al., 2009). Lysates made from these modified strains were treated with an affinity purification resin to remove tagged proteins, leading to approximately four times as much GFP production from LETs compared to wild-type BL21 (Seki et al., 2009). Expanded to multiple other nucleases simultaneously, this approach could potentially have a substantial impact on linear DNA stability. That being said, genes that overlap on the chromosome cannot be tagged (and thus, removed), and physical limitations of the resin such as its binding capacity and specificity can limit the success of this approach. Additionally, the resin-treated lysate with the greatest improvement in LET protein yield exhibited a 15% loss in plasmid-driven protein expression, which reduces the overall potential yield of the system.

Nuclease Inhibition

Inhibiting nucleases present in the lysate instead of removing them completely is another promising approach to improving expression from LETs (**Figure 2B**). This strategy is potentially simpler and more easily generalizable than genetic modifications, since the addition of supplements to CFEs can be applied to lysates of any strain background. Multiple strategies for nuclease inhibition have been reported.

One widely used method for nuclease inhibition is inclusion of the bacteriophage λ protein GamS in CFEs. GamS is an inhibitor of RecBCD that can protect linear DNA from degradation *in vivo* and in crude extracts (Murphy 1991; Sitaraman et al., 2004). GamS has been used in CFEs to allow expression from LETs containing various promoters, including T7, natural and synthetic σ^{70}, and well-characterized inducible promoters (Sun et al., 2013). With GamS supplementation, LETs with a strong σ^{70} promoter yielded 37.6% of that from a plasmid, a large improvement from the 2% LET yield in the absence of GamS. The GamS protein can be purified with affinity tagging or purchased commercially specifically for use in CFEs. Enhanced extracts can also be made from strains that express GamS to improve LET performance without the need to exogenously add protein. These GamS-containing extracts doubled LET protein yield compared to extracts without GamS (Contreras-Llano et al., 2020). While this method is more cost-effective and avoids any negative effects caused by the protein storage buffer, the observed benefits in LET performance are far less than those seen when purified GamS is added. Additionally, enriched lysates made from plasmid-carrying strains can suffer from overall decreases in efficiency, a consequence of the burdens of plasmid maintenance (Zawada and Swartz 2006).

The use of Chi DNA for RecBCD inhibition also avoids the costly purchase or reagent-intensive purification of proteins. Chi DNA are dsDNA oligos containing the short DNA motif (5′-GCTGGTGG-3′) referred to as a Chi (crossover hotspot instigator) site. These sites can be found throughout the genome of *E. coli* and are known to be involved in recombination events. Chi sites act as recognition sequences for RecBCD and other proteins involved in recombination (Smith et al., 1981). Addition of dsDNA composed of six repeated Chi sequences (with spacers) to the cell-free reaction has been shown to stabilize LETs and improve protein yield from undetectable levels to approximately 23% of that from a plasmid (Marshall et al., 2017).

RecBCD activity can also be inhibited by organic small molecules. From an *in vivo* screen of 326,100 organic small molecules, several compounds were identified to inhibit the nuclease activity of RecBCD and AddAB, a RecBCD analog native to *B. subtilis* and other bacteria (Amundsen et al., 2012). Many of these compounds showed reduced helicase and Chi-cutting activity from RecBCD as well. Two of these small molecules, CID 697851 and CID 1517823, improved expression from LETs in CFEs as much as 200% at some concentrations compared to reactions with no RecBCD inhibitors (Shrestha et al., 2014). However, expression remained far less than that from a plasmid.

GamS, Chi DNA, and small molecule RecBCD inhibitors all function by binding directly to RecBCD, but nuclease inhibitors can also bind directly to the LET to protect it from degradation. For example, adding single-chain Cro (ssCro) to a cell-free reaction protects LETs containing the ssCro operator recognition sequences from RecBCD degradation (Jana et al., 1998; Zhu et al., 2020). These ssCro operators can be easily added to LETs through PCR. Addition of ssCro improved LET expression over 6-fold, resulting in about 23% of plasmid-based expression (Zhu et al., 2020). DNA-binding proteins can also inhibit degradation from other exonucleases. Ku, for example, is a DNA-binding protein with homologs found in eukaryotes and prokaryotes that aids in nonhomologous end joining (Aravind and Koonin 2001). Ku binds directly to dsDNA termini and is proven to reduce degradation by AdnAB, a helicase/nuclease in mycobacteria (Sinha et al., 2009). Ku has been used recently in *E. coli* CFEs to protect LETs and improved transcription significantly, but those yields were only about a third of those transcribed when using GamS or Chi DNA oligos as inhibitors (Yim et al., 2020). However, because of Ku's different mechanism for LET protection, it was proven to be the most effective in lysates made from diverse bacteria. In *B. subtilis* and *C. glutamicum* extracts, Ku-protected LETs improved transcription 4.43- and 1.58-fold, respectively, compared to the no inhibitor control (Yim et al., 2020). This is notable improvement, considering GamS and Chi DNA addition to *C. glutamicum* CFEs actually reduced RNA yield slightly. While Chi DNA enhanced transcription from LETs in *B. subtilis* CFEs, Ku was about twice as effective.

Linear Expression Template Modifications and Enhanced Design

Chemical modification of LETs can also reduce their enzymatic degradation in CFEs (**Figure 2C**). In *E. coli*, methylation of genomic DNA provides some protection against restriction endonucleases (Marinus and Løbner-Olesen 2014). PCR-generated LETs, by virtue of being synthesized *in vitro*, are unmethylated. Post-synthesis methylation via commercially available enzymes was thus tested as an avenue for extra protection against degradation of LETs. LET methylation via *dam* methyltransferase (adenosine-specific) increased protein yield by 32% while methylation by CpG methyltransferase (cytosine-specific) actually lowered protein yield by 18% (Zhu et al., 2020). This decrease could be because CpG methyltransferase is a eukaryotic enzyme, yielding methylation patterns that might still appear foreign in *E. coli*. This hypothesis would suggest that it's not simply the presence of methyl groups added by *dam* methyltransferase that led to increased yield, but the prokaryotic methylation pattern on the LET disguising it as native DNA. The potential number of methyl groups that can be incorporated throughout the LET is restricted with this method, since the commercial methyltransferases act only at sequence-specific sites. These sites can easily be added at regions flanking the gene, but the concentration of methylation sites within the gene itself will be fixed.

Noncanonical DNA backbones and nucleotides can also be used to decrease LET susceptibility to degradation by lysate nucleases. Phosphorothioate (PT) linkages have been shown to restrict exonuclease digestion *in vivo,* specifically from Lambda exonuclease (Exo) (Putney et al., 1981; Mosberg et al., 2012). Terminal PT linkages can easily be added to LETs during PCR with modified primers. However, these DNA modifications have generally shown insignificant stabilization of LETs in CFEs (Zhu et al., 2020), with the exception of one instance where two PT linkages were added immediately upstream of the σ^{70} promoter (Sun et al., 2013). Even still, these PT linkages only improved LET expression by about 36%. Internal PT linkages might be more effective in protecting dsDNA from Lambda Exo degradation *in vivo,* as suggested by increased recombination frequency (Mosberg et al., 2012). In addition, several other chemical modifications can be made to LETs via PCR primers that might reduce their susceptibility to degradation, including carbon spacers and 2′-O-methoxy-ethyl bases. However, to our knowledge, these modifications have yet to be tested in CFEs.

Different PCR-based techniques for template DNA production have also been explored for generating templates that provide the highest protein yield while still avoiding the lengthy cloning steps required to make plasmids. With appropriately designed primers, PCR amplicons can be purified and ligated to create circularized templates that give protein yields comparable to plasmid DNA without cloning and transformation (Wu et al., 2007). Furthermore, dilute samples of these recircularized PCR amplicons can be amplified overnight via rolling circle amplification and used directly in CFEs at more suitable concentrations (Dopp et al., 2019). These approaches maintain shorter production times, which is desirable for high-throughput testing, but have not gained much popularity in other applications. Also, their use has been restricted to only T7 expression elements.

Finally, designing the LET sequence in ways that specifically prolong the half-life of the messenger RNA (mRNA) transcripts can significantly improve LET-based production. Increased mRNA half-life is essential for LET yield since it allows for continued protein production even after the LET has been compromised. Use of extracts from strains such as BL21 Star—which has reduced RNase activity due to a mutation in the RNase E gene—is one way to accomplish this. Additionally, the inclusion of various 3′-tail mRNA secondary structures can improve protein yield from LETs, with poly(G) tails and T7 terminator sequences increasing yield by nearly 2-fold and 3-fold, respectively (Ahn et al., 2005).

Reduced Culture Temperatures

It is worth noting that simple modifications to the lysate preparation protocol can sometimes improve LET protein yield, a particularly desirable approach based on its simplicity. It has previously been reported that *E. coli* cultivation temperature can have a significant influence on the exonuclease activity of the lysate. Reducing culture temperature to 30°C was shown to approximately triple LET-driven protein yield compared to lysates cultured at 37°C, while plasmid-driven protein yield decreased slightly (Seki et al., 2008).

However, others have reported no increased expression from LETs when using this method (Sun et al., 2013).

APPLICATIONS OF LINEAR EXPRESSION TEMPLATES

While the discussion in *Approaches for Stabilizing Linear DNA in Cell-Free Expression Systems* often used protein yields as a metric to indicate the relative utility of methods to stabilize LETs, producing large quantities of protein is not the most compelling reason to use LETs. In fact, all of the approaches described above still led to protein yields from LETs below those from plasmids. When producing large amounts of protein is the goal, the use of plasmids is preferred not only due to the attendant improved expression, but also due to the relatively low cost of plasmid production. The most prominent benefits of LETs lie in their ability to accelerate synthetic biology prototyping and enable the expression of genes that would otherwise be toxic to produce *in vivo*.

Circuit Prototyping

Rapid prototyping of genetic circuits has been significantly impacted by the use of LETs in CFEs. Genetic circuit design to date has largely been guided by trial-and-error testing, known in the field as the Design-Build-Test-Learn (DBTL) cycle. The length of the DBTL cycle can be substantially shortened by using LETs to avoid time-consuming cloning steps after every iteration. For example, LETs have been used to assemble feedforward loop circuits and quickly facilitate promoter optimization (Guo and Murray 2019). Other complex systems, such as synthetic genetic oscillators (Niederholtmeyer et al., 2015; Yelleswarapu et al., 2018) and four-piece genetic switches have also been be assembled with LETs (Sun et al., 2013). To quantify rates of transcription and translation in real time, fluorescent reporters can be used at the RNA and/or protein level and measured directly. One example of such an assay, PERSIA, uses LETs in CFEs to rapidly characterize several biological phenomena during the reaction (Wick et al., 2019). The activities of T7 promoter mutants have been efficiently characterized through the use of LETs by using the PCR products from an *in vitro* transcription screen directly in CFEs (Komura et al., 2018). Similarly, using a common reverse primer and unique forward primers, a LET library was quickly generated to test the spatial dependence of promoter and operator sequences at low cost in lysate-based CFEs (McManus et al., 2019).

Engineering Synthetic Regulators

LETs can also expedite the lengthy screening time required for engineering synthetic regulators. RNA toehold switches are powerful riboregulators that have been used to detect nucleic acid sequences *in vivo* (Green et al., 2014) and in CFEs, including low-cost user-friendly diagnostics to detect pathogens (Pardee et al., 2016). A toehold-based sensing system consists of two pieces: the RNA toehold switch sequence responsible for translational regulation, and the *trans*-acting nucleic acid trigger sequence that binds to the switch and modulates its activity. The switch and/or the RNA trigger can be encoded on LETs and expressed in the cell-free reaction (Takahashi et al., 2018), which is particularly important because finding a switch with good performance characteristics (leakiness, limits of detection, etc.) requires screening multiple (and sometimes many) candidates. Also, triggers can be added directly as linear ssDNA oligos to quickly test for switch functionality for a given target, avoiding the encoding of RNA triggers on DNA templates entirely (Amalfitano et al., 2021).

While the above and many other applications of toehold switches have used (nuclease-free) PURE systems, LET-based toehold systems can also be used in crude cell lysate systems for a more cost-efficient way to screen switches. Using Chi DNA as a nuclease inhibitor, ssDNA trigger oligos were stabilized and detected with toehold switches in encapsulated CFEs (Garamella et al., 2019). RNA triggers encoded onto LETs were used in developing a novel crude lysate-based multiplexed diagnostic sensor (Zhang et al., 2021). Additionally, in the more general class of synthetic riboregulators, riboswitch development also suffers from inefficient trial-and-error design (Etzel and Mörl 2017) that can be addressed with LETs in CFEs. This approach has previously been used as a platform to model riboregulator kinetics and predict *in vivo* behavior (Senoussi et al., 2018).

CRISPR-Cas Screening

Development and screening of CRISPR systems has also benefited from the rapid prototyping enabled by LETs in CFEs. CRISPR technologies harness the nuclease activities from what is essentially a bacterial "immune system" to execute functions ranging from precise and efficient DNA editing to indiscriminate cleavage of nucleic acid sequences. These functions have been used in a variety of applications from plant and animal genomic manipulation to therapeutics and diagnostics (Barrangou and Doudna 2016; Petri and Pattanayak 2018). However, the laborious process of screening protein function has imposed obstacles to how rapidly advances can be made. This bottleneck has been mitigated in part through the use of CFEs for rapid CRISPR-Cas characterization (Marshall et al., 2018). LETs amplified from genomic fragments containing multiple Cas proteins have been used in CFEs to discover CRISPR-Cas12a inhibitors (Watters et al., 2018). LETs encoding dozens of guide RNA switch candidates (guide RNAs coupled to RNA toehold switches) were used for rapid characterization in CFEs to develop a modular design scheme for regulating Cas12a activity with arbitrary RNA inducers (Collins et al., 2021). LETs have also been used to elucidate the protospacer-adjacent motif (PAM) sequences recognized by Cas nucleases (Maxwell et al., 2018). While LETs can be used to express a Cas protein as well as its guide RNA sequence, it is particularly advantageous to use LETs for Cas expression since certain Cas genes can be extremely difficult to clone into plasmid vectors containing promoters suitable for CFEs (Maxwell et al., 2018).

Expression of Toxic Proteins

Beyond Cas, the expression of many metabolically taxing or toxic proteins has been facilitated by LETs. *In vivo*, some proteins can interfere with metabolic pathways or inhibit cell division,

resulting in cell death. Since CFEs do not have these same requirements as cells, they are a promising system for expression of such proteins. However, use of plasmids as expression templates in CFEs still requires traditional plasmid cloning, which can be challenging for such toxic proteins. LETs eliminate this challenge, as they can be generated by PCR from a DNA template that is much easier to clone: a plasmid with a defective promoter. For example, the human G-protein coupled receptor (GPCR) has been expressed at cytotoxic levels in both lysate-based CFEs and recombinant protein systems using LETs (Haberstock et al., 2012). Antimicrobial peptides (AMPs), which act as a natural infection defense mechanism and are known to severely stunt bacterial growth, have been successfully expressed with CFEs using LETs (Dopp et al., 2019). Other classes of difficult-to-express proteins that have been expressed successfully off plasmids in CFEs include vaccine antigens (Welsh et al., 2012) and antibiotic efflux pumps (Wuu and Swartz 2008). Employing LETs in these applications could simplify template production or translation to other proteins in these classes.

Other Applications

LETs also allow CFEs to produce larger protein libraries more easily. Protein microarrays are systems that screen the activity and interaction of many proteins in parallel, but their use is limited by persisting technical challenges including lengthy protein expression, purification, and immobilization steps that reduce the maximum diversity of proteins for array construction (Schinn et al., 2016). The same challenges do not arise with DNA arrays; thus, *in situ* protein arrays that can use nucleic acid arrays as a template for protein production are advantageous (He et al., 2008b). Two methods that use LETs and CFEs to create protein microarrays include the protein *in situ* array, PISA (He and Taussig 2001), and a DNA array to protein array approach, DAPA (He et al., 2008a). Both methods involve parallel *in situ* protein expression and surface capture using purification tags that can be easily added to LETs with PCR. With PISA, 35 fg of unpurified LET was sufficient for detectable production of GFP in sub-nL volumes (Angenendt et al., 2006; Berrade et al., 2011).

LETs have also been shown to be functional in various non-standard reaction environments for CFEs. Using Chi DNA as a nuclease inhibitor, LETs have been used in polymer microgels to synthesize functional malonyl-CoA synthetase (MatB) (Köhler et al., 2020). LETs have also been spotted on glass slides and aligned to microfluidic devices enclosing each LET into a reaction chamber (Gerber et al., 2009). Microfluidic devices have also been used to encapsulate LET-containing CFEs (with GamS inhibitor) into agarose hydrogel beads, which then withstood lyophilization, storage, and rehydration into a functional sensor for *P. aeruginosa* quorum-sensing molecules (Seto 2019). Moreover, with sufficient protection, LETs can be used in continuous reaction conditions as well. While batch operation is the simplest mode for CFEs, depletion and degradation of reagents can limit long-term productivity. One notable example of a continuous format to address this issue was the use of a multilayer microfluidic device engineered to conduct extended cell-free reactions, where LETs were used to create a screening process that was entirely *in vitro* (Niederholtmeyer et al., 2013).

THE FUTURE OF LINEAR DNA IN CELL-FREE EXPRESSION SYSTEMS

Advances in nuclease inhibition strategies to facilitate LET-based expression have made the applications described above, and many others, possible. While existing nuclease inhibition methods are sufficient for certain applications, there is still both room and a critical need for improvement. Further enhancement of linear DNA stability and protein yield from LETs in CFEs can drive innovation of novel synthetic biology research tools and broaden the scope of viable applications of CFEs.

Improvements in Linear Expression Template-Based Expression

One insufficiently studied aspect of CFEs is the interplay between promoter strength and template DNA structure. Since a common goal for using LETs in CFEs is maximizing protein yield, and since a common concern with using LETs is diminished expression, most scientists designing LETs choose to use strong promoters. As a result, there is little literature characterization of the expression levels from different promoters from LETs. However, use of diverse promoters is often an integral part of engineering genetic circuits and cell-free technologies, meaning that an accurate characterization of promoter strength in LETs is important. Multiple investigations of synthetic σ^{70} constitutive promoters on LETs showed that the relative strengths seen in CFEs have no correlation to the strengths seen when using the same promoters on plasmids in CFEs or *in vivo* (Chappell et al., 2013; Sun et al., 2013).

One hypothesis is that this discrepancy is due to the relationship between transcriptional rate and DNA supercoiling. This is supported by experiments that show stronger correlations of different ribosome binding site (RBS) strengths between linear and plasmid templates, as RBS sequence and resulting mRNA structure are independent of DNA template conformation (Chappell et al., 2013). Others have also explored the effect of supercoiled linear DNA on transcription by testing two different genes in convergent, divergent, and tandem orientations on a single LET (Yeung et al., 2017). Transcription from each configuration was significantly different, most likely because less supercoiling occurs at free DNA ends compared to the center. Recently, a study that focused on the impact of LET length on performance in CFEs verified via atomic force microscopy (AFM) a significant increase in DNA supercoiling as the LET length increased (Nishio et al., 2021). This study also showed that increasing the LET length from 2.8 to 25.7 kbp improved cell-free expression approximately 3-fold, which could be attributed to the positive influence of DNA supercoiling. The benefits of extending LET length have been previously reported while supplementing with GamS, albeit on a

much smaller scale, where a 2.4-fold increase in LET expression was observed after the addition of just five bp on each end (Sun et al., 2013). Nonetheless, DNA supercoiling, the significance of gene orientation, and LET length should be taken into consideration for optimal design of future LET-based technologies.

It is also important to identify whether linear DNA stabilization techniques are generalizable across lysates from different species of bacteria. *E. coli* extracts make up the majority of CFEs because their lysate is the simplest to prepare and provides high protein yields, and because *E. coli* is otherwise a workhorse model organism. However, *E. coli* extracts are not practical for every application, as they cannot support some functionalities easily (e.g., glycosylation). As a result, the use of other chassis organisms as lysate sources is growing, and thus so will the demand for expression from LETs in CFEs made from those lysates. Some stabilization techniques that are shown to be most effective in *E. coli* systems (e.g., GamS) are practically ineffective in other CFEs such as *V. natriegens* (Wiegand et al., 2019). This is perhaps to be expected, as different organisms can possess different nucleases. One example of a nuclease inhibition approach with potential to be effective in CFEs of different species would be the small molecule RecBCD inhibitors mentioned in *Nuclease Inhibition*, as they have been shown to also inhibit AddAB (a helicase/nuclease native to many bacteria) (Amundsen et al., 2012). Interestingly, Chi DNA sites have also been identified in other bacteria, such as *B. subtilis* and *L. lactis*, indicating that Chi DNA oligos could potentially be useful in extracts made from these bacteria (Chedin and Kowalczykowski 2002; Marshall et al., 2017). Chi DNA inhibition has recently been tested in *B. subtilis* CFEs, and only improved LET yield about 2-fold (Yim et al., 2020). At any rate, identifying which techniques are generalizable—and for those that are not generalizable, identifying alternatives or replacements as needed—will be important in moving this aspect of CFEs forward.

However, it may in fact turn out that lysate-based CFEs from some non-model organisms could actually improve the efficiency of LET-based expression. Chinese hamster ovary (CHO) cell extracts have been shown to allow for sufficient LET functionality to be used for production of "difficult-to-express" proteins (Thoring et al., 2016). Wheat germ CFEs are known to possess undetectable nuclease activity compared to other cell-free systems (Endo and Sawasaki 2003). Surprisingly, in yeast CFEs, expression from LETs has been shown to yield 40–60% more protein compared to a plasmid DNA template (Gan and Jewett 2014). These LETs included a Ω leader sequence for initiation, a T7 promoter, and a 3' poly$(A)_{50}$ tail. Lysates from both *S. frugiperda* (Sachse et al., 2013), and *P. putida* (H. Wang et al., 2018) allow efficient expression from LETs, with the latter allowing for LET-based expression at about 70% of plasmid-based yields in the absence of any nuclease inhibitors or DNA protection. However, the economic viability for any downstream applications of CFEs based on these other chassis organisms remains to be seen.

Simultaneous inhibition of multiple nucleases could have a substantial impact on linear DNA stability in CFEs. The majority of approaches to nuclease inhibition have focused on RecBCD as the primary DNA nuclease, but significant activity from other nucleases has been recognized yet typically overlooked (Sun et al., 2013). For example, ExoVII has been identified as a key dsDNA and ssDNA nuclease; its removal has improved the preservation of mutations on dsDNA termini *in vivo* (Mosberg et al., 2012). Fortunately, some of the widely used nuclease inhibitors can inhibit more than one nuclease. For example, GamS can also inhibit SbcCD, another endo/exonuclease in *E. coli* (Kulkarni and Stahl 1989). Other significant native *E. coli* nucleases include RecJ, ExoI, and ExoX (Mosberg et al., 2012). Generalized inhibition methods would further improve linear DNA stability and help bring LET-based gene expression closer to plasmid yields.

Linear Expression Templates to Avoid Confounding Effects in Cell-Free Expression Systems

LETs also have the potential to solve a unique phenomenon sometimes seen in CFEs but infrequently reported. Our group has observed multiple instances where the addition of plasmids expressing unrelated proteins, or even with no gene insert ("empty vectors"), seems to increase protein production from a separate reporter plasmid (Miguez et al., 2021). This is perhaps counterintuitive, as one might expect transcription or translation to be limited by competition for resources, such that adding more plasmids would increase competition for expression resources and thus likely decrease expression from the original plasmid. We hypothesize that this effect could be caused (in certain expression regimes) by RNase competition. Expression from additional plasmids added to the system would increase the total concentration of RNA, which would increase "competition" for RNases and thus in turn yield increased half-life of mRNA transcripts. Even empty vectors contribute to increasing the total RNA concentration in the reaction due to expression from their selectable marker cassettes. Since LETs encode only the gene of interest and no selectable markers, different LETs could be added for each gene to be expressed without the confounding effects typically caused by plasmids.

Aptamers as Sensing Tools in Cell-Free Expression Systems

In synthetic biology, the engineering of novel genetic circuits is limited most significantly by the characteristics and diversity of extant cellular machinery (Purnick and Weiss 2009). Specifically, in the development of biosensors using CFEs, limitations on the "biological parts" available in nature has led to increasing demand for *de novo* regulators and synthetic circuit elements with low crosstalk and high sensitivity (Green et al., 2014). The most prominent examples of success in this space allow for rapid and specific detection of arbitrary nucleic acid sequences, which is useful in developing sensors for pathogens and other microbes. Detection of small molecules, however, has been more

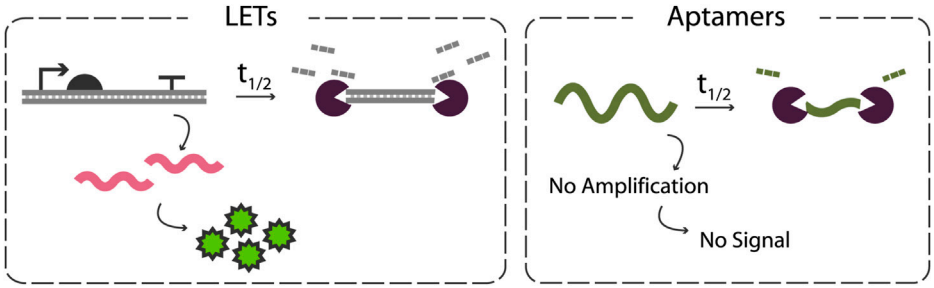


FIGURE 3 | Comparison of the impacts of degradation on LETs vs DNA aptamers in CFEs. LETs benefit from downstream amplification steps, leading to continued protein synthesis even after complete LET degradation. Aptamers, however, must be present and intact to execute their function as they do not serve as templates for amplification.

challenging because there are no *de novo*-designable regulators to facilitate detection of arbitrary small molecules (Silverman et al., 2020). There is thus a critical need for analogous technologies that can be used in synthetic circuits to detect classes of molecules beyond nucleic acid sequences, including small molecules, proteins, and ions.

Aptamers could potentially fill this need. Aptamers are short oligonucleotides that bind to a specific target with high affinity. DNA aptamers are stable at room temperature, can be synthesized at low cost with negligible batch-to-batch variability, and can withstand lyophilization (Zhang et al., 2019). They have previously been used as the basis for robust sensors for the detection of all types of molecular targets. Most importantly, aptamers are evolved *in vitro*, which means new aptamers can be created for virtually any analyte.

Despite all of the strengths of DNA aptamers, there are surprisingly few examples of their use in CFEs, likely due to the instability of linear DNA in bacterial extracts. As discussed in *Approaches for Stabilizing Linear DNA in Cell-Free Expression Systems*, even the best methods for stabilizing LETs still allow significant degradation. DNA aptamers, unlike LETs, are mostly ssDNA oligos, which likely increases their susceptibility to enzymatic degradation in *E. coli* since there are more native exonucleases that act on ssDNA specifically than there are that act solely on dsDNA (Lovett 2011). Also, aptamers are significantly shorter in length than LETs: LETs include entire coding sequences and other transcription elements, while evolved aptamers are normally 20–60 nucleotides long (Lakhin et al., 2013). Because we know that nucleases can cleave DNA at rates as high as 500 bp/s (Spies et al., 2007) and that extending LETs results in improved protein yield (Sun et al., 2013), it is likely that the short lengths of aptamers are a major contributor to their instability. This seems reasonable because on a mass basis, aptamers have more free ends than LETs for exonucleases to bind to, and on a molar basis they may have a higher density of critical sequence without which they cannot function.

Another characteristic that sets aptamers apart from LETs is that aptamers do not undergo downstream amplification steps like LETs do (**Figure 3**). LETs produce numerous mRNA transcripts which are then amplified into protein. These transcripts remain viable templates for protein translation even after the LET has been degraded, and the proteins themselves may also have some function that serves as an amplification of signal in certain applications. Aptamers, however, are not templates for amplification and thus leave no residual function after enzymatic degradation. This means that for aptamers to be used successfully in lysate-based CFEs, nuclease inhibition strategies must be strong enough to prevent aptamer degradation for the entirety of the reaction.

In a few cases, these obstacles have been overcome or circumvented sufficiently to allow DNA aptamers to be successfully used in CFEs. For example, DNA aptamers integrated into plasmids have successfully allowed transcriptional regulation in CFEs (Iyer and Doktycz 2014; Wang et al., 2017). This confirms that the environment of CFEs (in these examples, a commercially available extract) can accommodate proper aptamer secondary structure formation to allow for both target analyte binding and subsequent conformational changes. Integrating the aptamer onto a circular template was done to mitigate exonuclease activity that would otherwise degrade the aptamer. To that end, these reports do not indicate the use of any additional nuclease inhibitors in their experiments.

However, the integration of aptamers onto circular templates can be significantly more labor-intensive and challenging than most of the LET stabilization techniques described above. Since aptamers are single-stranded, incorporation into a circular template typically requires generation of a ssDNA circular template by affinity purification of biotin-labeled ssDNA produced via PCR (Mitchell and Merril 1989). This approach has the advantage of *in vitro* production but often produces insufficient yields for downstream applications. An alternative method that can produce a higher yield of ssDNA entails *in vivo* amplification in *E. coli* with phagemids, which are plasmids that possess both bacteriophage and plasmid properties and thus have all the viral components necessary to enable ssDNA replication (Kuczek et al., 1998; Zhou et al., 2009). However, this approach is even more complex than the *in vitro* strategy. Regardless of the synthesis strategy, constraining DNA aptamers to circular templates limits the design space and could compromise the functionality of some aptamers. Given all of these considerations, the development of simpler, more effective, and more

generalizable methods to enable their stability are a critical need to unleash the potential of DNA aptamers in CFEs.

CONCLUSION

The advances made in nuclease inhibition (specifically of RecBCD) have improved linear DNA stability in lysate-based CFEs sufficiently to allow feasible LET-based production of proteins. The accessibility and ease of implementation of these methods has supported their effective use in driving innovative research in genetic circuit design, efficient screening of *de novo* riboregulators, rapid screening of CRISPR-Cas system parts, and more. However, continued progress is necessary, because while protein yields from LETs have improved significantly, this is not a direct measurement of LET half-life and obscures the difficulties in using other types of linear DNA constructs in CFEs. Keeping linear DNA stable long enough to function on its own, rather than as a template for amplification, is an important challenge to tackle. Some of the potentially most promising applications of linear DNA in CFEs remain infeasible due to nuclease-based degradation, meaning that more effective methods to extend linear DNA half-life could have a dramatic, enabling impact on a myriad of biotechnology applications.

AUTHOR CONTRIBUTIONS

Writing–original draft: MM. Writing–review and editing: MS.

FUNDING

The authors thank the National Institutes of Health (R01-EB022592) for funding support.

ACKNOWLEDGMENTS

The authors thank Alexandra Patterson for insightful discussions regarding DNA methylation.

REFERENCES

Ahn, J.-H., Chu, H.-S., Kim, T.-W., Oh, I.-S., Choi, C.-Y., Hahn, G.-H., et al. (2005). Cell-Free Synthesis of Recombinant Proteins from PCR-Amplified Genes at a Comparable Productivity to that of Plasmid-Based Reactions. *Biochem. Biophysical Res. Commun.* 338 (3), 1346–1352. doi:10.1016/j.bbrc.2005.10.094

Amalfitano, E., Karlikow, M., Norouzi, M., Jaenes, K., Cicek, S., Masum, F., et al. (2021). A Glucose Meter Interface for Point-of-Care Gene Circuit-Based Diagnostics. *Nat. Commun.* 12 (1), 724. doi:10.1038/s41467-020-20639-6

Amundsen, S. K., Spicer, T., Karabulut, A. C., Londoño, L. M., Eberhart, C., Fernandez Vega, V., et al. (2012). Small-Molecule Inhibitors of Bacterial AddAB and RecBCD Helicase-Nuclease DNA Repair Enzymes. *ACS Chem. Biol.* 7 (5), 879–891. doi:10.1021/cb300018x

Angenendt, P., Kreutzberger, J., Glökler, J., and Hoheisel, J. D. (2006). Generation of High Density Protein Microarrays by Cell-free *In Situ* Expression of Unpurified PCR Products. *Mol. Cell Proteomics* 5 (9), 1658–1666. doi:10.1074/mcp.T600024-MCP200

Aravind, L., and Koonin, E. V. (2001). Prokaryotic Homologs of the Eukaryotic DNA-End-Binding Protein Ku, Novel Domains in the Ku Protein and Prediction of a Prokaryotic Double-Strand Break Repair System. *Genome Res.* 11 (8), 1365–1374. doi:10.1101/gr.181001

Barrangou, R., and Doudna, J. A. (2016). Applications of CRISPR Technologies in Research and beyond. *Nat. Biotechnol.* 34 (9), 933–941. doi:10.1038/nbt.3659

Berrade, L., Garcia, A. E., and Camarero, J. A. (2011). Protein Microarrays: Novel Developments and Applications. *Pharm. Res.* 28 (7), 1480–1499. doi:10.1007/s11095-010-0325-1

Chappell, J., Jensen, K., and Freemont, P. S. (2013). Validation of an Entirely *In Vitro* Approach for Rapid Prototyping of DNA Regulatory Elements for Synthetic Biology. *Nucleic Acids Res.* 41 (5), 3471–3481. doi:10.1093/nar/gkt052

Chedin, F., and Kowalczykowski, S. C. (2002). A Novel Family of Regulated Helicases/Nucleases from Gram-Positive Bacteria: Insights into the Initiation of DNA Recombination. *Mol. Microbiol.* 43 (4), 823–834. doi:10.1046/j.1365-2958.2002.02785.x

Collins, S. P., Rostain, W., Liao, C., and Beisel, C. L. (2021). Sequence-independent RNA Sensing and DNA Targeting by a Split Domain CRISPR-Cas12a gRNA Switch. *Nucleic Acids Res.* 49, 2985–2999. doi:10.1093/nar/gkab100

Contreras-Llano, L. E., Meyer, C., Liu, Y., Sarker, M., Lim, S., Longo, M. L., et al. (2020). Holistic Engineering of Cell-free Systems through Proteome-Reprogramming Synthetic Circuits. *Nat. Commun.* 11 (1), 3138. doi:10.1038/s41467-020-16900-7

Dopp, J. L., Rothstein, S. M., Mansell, T. J., and Reuel, N. F. (2019). Rapid Prototyping of Proteins: Mail Order Gene Fragments to Assayable Proteins within 24 hours. *Biotechnol. Bioeng.* 116 (3), 667–676. doi:10.1002/bit.26912

Endo, Y., and Sawasaki, T. (2003). High-Throughput, Genome-Scale Protein Production Method Based on the Wheat Germ Cell-free Expression System. *Biotechnol. Adv.* 21 (8), 695–713. doi:10.1016/S0734-9750(03)00105-8

Etzel, M., and Mörl, M. (2017). Synthetic Riboswitches: From Plug and Pray toward Plug and Play. *Biochemistry* 56 (9), 1181–1198. doi:10.1021/acs.biochem.6b01218

Gan, R., and Jewett, M. C. (2014). A Combined Cell-Free Transcription-Translation System from Saccharomyces Cerevisiaefor Rapid and Robust Protein Synthesis. *Biotechnol. J.* 9 (5), 641–651. doi:10.1002/biot.201300545

Garamella, J., Majumder, S., Liu, A. P., and Noireaux, V. (2019). An Adaptive Synthetic Cell Based on Mechanosensing, Biosensing, and Inducible Gene Circuits. *ACS Synth. Biol.* 8 (8), 1913–1920. doi:10.1021/acssynbio.9b00204

Gerber, D., Maerkl, S. J., and Quake, S. R. (2009). An *In Vitro* Microfluidic Approach to Generating Protein-Interaction Networks. *Nat. Methods* 6 (1), 71–74. doi:10.1038/nmeth.1289

Green, A. A., Silver, P. A. James. J. Collins., Collins, J. J., and Yin, P. (2014). Toehold Switches: De-novo-designed Regulators of Gene Expression. *Cell* 159 (4), 925–939. doi:10.1016/j.cell.2014.10.002

Guo, S., and Murray, R. M. (2019). Construction of Incoherent Feedforward Loop Circuits in a Cell-free System and in Cells. *ACS Synth. Biol.* 8 (3), 606–610. doi:10.1021/acssynbio.8b00493

Haberstock, S., Roos, C., Hoevels, Y., Dötsch, V., Schnapp, G., Pautsch, A., et al. (2012). A Systematic Approach to Increase the Efficiency of Membrane Protein Production in Cell-free Expression Systems. *Protein Expr. Purif.* 82 (2), 308–316. doi:10.1016/j.pep.2012.01.018

He, M., Stoevesandt, O., Palmer, E. A., Khan, F., Ericsson, O., Taussig, M. J., et al. (2008a). Printing Protein Arrays from DNA Arrays. *Nat. Methods* 5 (2), 175–177. doi:10.1038/nmeth.1178

He, M., Stoevesandt, O., and Taussig, M. J. (2008b). *In Situ* Synthesis of Protein Arrays. *Curr. Opin. Biotechnol.* 19 (1), 4–9. doi:10.1016/j.copbio.2007.11.009

He, M., and Taussig, Michael. J. (2001). Single Step Generation of Protein Arrays from DNA by Cell-free Expression and *In Situ* Immobilisation (PISA Method). *Nucleic Acids Res.* 29 (15), 73e–73. doi:10.1093/nar/29.15.e73

Iyer, S., and Doktycz, M. J. (2014). Thrombin-Mediated Transcriptional Regulation Using DNA Aptamers in DNA-Based Cell-free Protein Synthesis. *ACS Synth. Biol.* 3 (6), 340–346. doi:10.1021/sb4000756

Jana, R., Hazbun, T. R., Fields, J. D., and Mossing, M. C. (1998). Single-Chain Lambda Cro Repressors Confirm High Intrinsic Dimer–DNA Affinity†. *Biochemistry* 37 (18), 6446–6455. doi:10.1021/bi980152v

Katzen, F., Chang, G., and Kudlicki, W. (2005). The Past, Present and Future of Cell-free Protein Synthesis. *Trends Biotechnol.* 23 (3), 150–156. doi:10.1016/j.tibtech.2005.01.003

Khambhati, K., Bhattacharjee, G., Gohil, N., Braddick, D., Kulkarni, V., and Singh, V. (2019). Exploring the Potential of Cell-free Protein Synthesis for Extending the Abilities of Biological Systems. *Front. Bioeng. Biotechnol.* 7. doi:10.3389/fbioe.2019.00248

Köhler, T., Heida, T., Hoefgen, S., Weigel, N., Valiante, V., and Thiele, J. (2020). Cell-Free Protein Synthesis and *In Situ* Immobilization of DeGFP-MatB in Polymer Microgels for Malonate-To-Malonyl CoA Conversion. *RSC Adv.* 10 (66), 40588–40596. doi:10.1039/D0RA06702D

Komura, R., Aoki, W., Motone, K., Satomura, A., and Ueda, M. (2018). High-Throughput Evaluation of T7 Promoter Variants Using Biased Randomization and DNA Barcoding. *PLOS ONE* 13 (5), e0196905. doi:10.1371/journal.pone.0196905

Kuczek, K., Kotowska, M., Wiernik, D., and Mordarski, M. (1998). Single-Stranded DNA Production from Phagemids Containing GC-Rich DNA Fragments. *BioTechniques* 24 (2), 214–215. doi:10.2144/98242bm08

Kulkarni, S. K., and Stahl, F. W. (1989). Interaction between the SbcC Gene of Escherichia Coli and the Gam Gene of Phage Lambda. *Genetics* 123 (2), 249–253. doi:10.1093/genetics/123.2.249

Kuzminov, A., and Stahl, F. W. (1997). Stability of Linear DNA in RecA Mutant Escherichia Coli Cells Reflects Ongoing Chromosomal DNA Degradation. *J. Bacteriol.* 179 (3), 880–888. doi:10.1128/jb.179.3.880-888.1997

Kwon, Y.-C., and Jewett, M. C. (2015). High-Throughput Preparation Methods of Crude Extract for Robust Cell-free Protein Synthesis. *Sci. Rep.* 5 (1), 8663. doi:10.1038/srep08663

Lakhin, A. V., Tarantul, V. Z., and Gening, L. V. (2013). Aptamers: Problems, Solutions and Prospects. *Acta Naturae* 5 (4), 34–43. doi:10.32607/20758251-2013-5-4-34-43

Laohakunakom, N., Grasemann, L., Lavickova, B., Michielin, G., Shahein, A., Swank, Z., et al. (2020). Bottom-Up Construction of Complex Biomolecular Systems with Cell-free Synthetic Biology. *Front Bioeng. Biotechnol.* 8, 213. doi:10.3389/fbioe.2020.00213

Lavickova, B., and Maerkl, S. J. (2019). A Simple, Robust, and Low-Cost Method to Produce the PURE Cell-free System. *ACS Synth. Biol.* 8 (2), 455–462. doi:10.1021/acssynbio.8b00427

Lovett, S. T. (2011). The DNA Exonucleases of Escherichia Coli. *EcoSal Plus* 4 (2). doi:10.1128/ecosalplus.4.4.7

Marinus, M. G., and Løbner-Olesen, A. (2014). DNA Methylation. *EcoSal Plus* 6 (1). doi:10.1128/ecosalplus.ESP-0003-2013

Marshall, R., Maxwell, C. S., Collins, S. P., Beisel, C. L., and Noireaux, V. (2017). Short DNA Containing χ Sites Enhances DNA Stability and Gene Expression in E. coli Cell-free Transcription-Translation Systems. *Biotechnol. Bioeng.* 114 (9), 2137–2141. doi:10.1002/bit.26333

Marshall, R., Maxwell, C. S., Collins, S. P., Jacobsen, T., Luo, M. L., Begemann, M. B., et al. (2018). Rapid and Scalable Characterization of CRISPR Technologies Using an E. Coli Cell-free Transcription-Translation System. *Mol. Cell* 69 (1), 146–157. doi:10.1016/j.molcel.2017.12.007

Maxwell, C. S., Jacobsen, T., Marshall, R., Noireaux, V., and Beisel, C. L. (2018). A Detailed Cell-free Transcription-Translation-Based Assay to Decipher CRISPR Protospacer-Adjacent Motifs. *Methods* 143 (July), 48–57. doi:10.1016/j.ymeth.2018.02.016

McManus, J. B., Emanuel, P. A., Murray, R. M., and Lux, M. W. (2019). A Method for Cost-Effective and Rapid Characterization of Engineered T7-Based Transcription Factors by Cell-free Protein Synthesis Reveals Insights into the Regulation of T7 RNA Polymerase-Driven Expression. *Arch. Biochem. Biophys.* 674 (October), 108045. doi:10.1016/j.abb.2019.07.010

McNerney, M. P., Zhang, Y. P. S., Steppe, P.A. D., Jewett, M. C., and Styczynski, M. P. (2019). Point-of-Care Biomarker Quantification Enabled by Sample-specific Calibration. *Sci. Adv.* 5 (9), eaax4473. doi:10.1126/sciadv.aax4473

Michel-Reydellet, N., Woodrow, K., and Swartz, J. (2005). Increasing PCR Fragment Stability and Protein Yields in a Cell-free System with Genetically Modified Escherichia Coli Extracts. *J. Mol. Microbiol. Biotechnol.* 9 (1), 26–34. doi:10.1159/000088143

Miguez, A. M., Zhang, Y., Piorino, F., and Styczynski, M. P. (2021). Metabolic Dynamics in Escherichia Coli-Based Cell-Free Systems. *bioRxiv*. doi:10.1101/2021.05.16.444339

Mitchell, L. G., and Merril, C. R. (1989). Affinity Generation of Single-Stranded DNA for Dideoxy Sequencing Following the Polymerase Chain Reaction. *Anal. Biochem.* 178 (2), 239–242. doi:10.1016/0003-2697(89)90631-3

Mosberg, J. A., Gregg, C. J., Lajoie, M. J., Wang, H. H., and Church, G. M. (2012). Improving Lambda Red Genome Engineering in Escherichia Coli via Rational Removal of Endogenous Nucleases. *PLOS ONE* 7 (9), e44638. doi:10.1371/journal.pone.0044638

Murphy, K. C. (1991). Lambda Gam Protein Inhibits the Helicase and Chi-Stimulated Recombination Activities of Escherichia Coli RecBCD Enzyme. *J. Bacteriol.* 173 (18), 5808–5821. doi:10.1128/jb.173.18.5808-5821.1991

Niederholtmeyer, H., Stepanova, V., and Maerkl, S. J. (2013). Implementation of Cell-free Biological Networks at Steady State. *Proc. Natl. Acad. Sci.* 110 (40), 15985–15990. doi:10.1073/pnas.1311166110

Niederholtmeyer, H., Sun, Z. Z., Hori, Y., Yeung, E., Verpoorte, A., Murray, R. M., et al. (2015). Rapid Cell-free Forward Engineering of Novel Genetic Ring Oscillators. *ELife* 4 (October), e09771. doi:10.7554/eLife.09771

Nirenberg, M., and Leder, P. (1964). RNA Codewords and Protein Synthesis: The Effect of Trinucleotides upon the Binding of SRNA to Ribosomes. *Science* 145 (3639), 1399–1407. doi:10.1126/science.145.3639.1399

Nishio, T., Yoshikawa, Y., Yoshikawa, K., and Sato, S.-i. (2021). Longer DNA Exhibits Greater Potential for Cell-free Gene Expression. *Sci. Rep.* 11 (1), 11739. doi:10.1038/s41598-021-91243-x

Pardee, K., Green, A. A., Ferrante, T., Cameron, D. E., DaleyKeyser, A., Yin, P., et al. (2014). Green, Tom Ferrante, D. Ewen Cameron, Ajay DaleyKeyser, Peng Yin, andPaper-Based Synthetic Gene Networks. *Cell* 159 (4), 940–954. doi:10.1016/j.cell.2014.10.004

Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., et al. (2016). Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* 165 (5), 1255–1266. doi:10.1016/j.cell.2016.04.059

Petri, K., and Pattanayak, V. (2018). SHERLOCK and DETECTR Open a New Frontier in Molecular Diagnostics. *CRISPR J.* 1 (3), 209–211. doi:10.1089/crispr.2018.29018.kpe

Purnick, P. E. M., and Weiss., R. (2009). The Second Wave of Synthetic Biology: From Modules to Systems. *Nat. Rev. Mol. Cell Biol* 10 (6), 410–422. doi:10.1038/nrm2698

Putney, S. D., Benkovic, S. J., and Schimmel, P. R. (1981). "A DNA Fragment with an Alpha-Phosphorothioate Nucleotide at One End is Asymmetrically Blocked From Digestion by Exonuclease III and can be Replicated *In Vivo*. *Proc. Natl. Acad. Sci.* 78 (12), 7350–7354. doi:10.1073/pnas.78.12.7350

Sachse, R., Wüstenhagen, D., Šamalíková, M., Gerrits, M., Bier, F. F., and Kubick, S. (2013). Synthesis of Membrane Proteins in Eukaryotic Cell-free Systems. *Eng. Life Sci.* 13 (1), 39–48. doi:10.1002/elsc.201100235

Schinn, S.-M., Broadbent, A., Bradley, W. T., and Bundy, B. C. (2016). Protein Synthesis Directly from PCR: Progress and Applications of Cell-free Protein Synthesis with Linear DNA. *New Biotechnol.* 33 (4), 480–487. doi:10.1016/j.nbt.2016.04.002

Schoborg, J. A., Hershewe, J. M., Stark, J. C., Kightlinger, W., Kath, J. E., Jaroentomeechai, T., et al. (2018). A Cell-free Platform for Rapid Synthesis and Testing of Active Oligosaccharyltransferases. *Biotechnol. Bioeng.* 115 (3), 739–750. doi:10.1002/bit.26502

Seki, E., Matsuda, N., and Kigawa, T. (2009). Multiple Inhibitory Factor Removal from an Escherichia Coli Cell Extract Improves Cell-free Protein Synthesis. *J. Biosci. Bioeng.* 108 (1), 30–35. doi:10.1016/j.jbiosc.2009.02.011

Seki, E., Matsuda, N., Yokoyama, S., and Kigawa, T. (2008). Cell-Free Protein Synthesis System from Escherichia Coli Cells Cultured at Decreased Temperatures Improves Productivity by Decreasing DNA Template Degradation. *Anal. Biochem.* 377 (2), 156–161. doi:10.1016/j.ab.2008.03.001

Senoussi, A., Lee Tin Wah, J., Shimizu, Y., Robert, J., Jaramillo, A., Findeiss, S., et al. (2018). Quantitative Characterization of Translational Riboregulators Using an *In Vitro* Transcription-Translation System. *ACS Synth. Biol.* 7 (5), 1269–1278. doi:10.1021/acssynbio.7b00387

Seto, J. (2019). On a Robust, Sensitive Cell-free Method for Pseudomonas Sensing and Quantification in Microfluidic Templated Hydrogels. *Micromachines* 10 (8), 506. doi:10.3390/mi10080506

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., et al. (2001). Cell-Free Translation Reconstituted with Purified Components. *Nat. Biotechnol.* 19 (8), 751–755. doi:10.1038/90802

Shrestha, P., Smith, M. T., and Bundy, B. C. (2014). Cell-Free Unnatural Amino Acid Incorporation with Alternative Energy Systems and Linear Expression Templates. *New Biotechnol.* 31 (1), 28–34. doi:10.1016/j.nbt.2013.09.002

Silverman, A. D., Karim, A. S., and Jewett, M. C. (2020). Cell-Free Gene Expression: An Expanded Repertoire of Applications. *Nat. Rev. Genet.* 21 (3), 151–170. doi:10.1038/s41576-019-0186-3

Sinha, K. M., Unciuleac, M.-C., Glickman, M. S., and Shuman, S. (2009). AdnAB: a New DSB-Resecting Motor-Nuclease from Mycobacteria. *Genes Development* 23 (12), 1423–1437. doi:10.1101/gad.1805709

Sitaraman, K., Esposito, D., Klarmann, G., Le Grice, S. F., Hartley, J. L., Chatterjee, D. K., et al. (2004). A Novel Cell-free Protein Synthesis System. *J. Biotechnol.* 110 (3), 257–263. doi:10.1016/j.jbiotec.2004.02.014

Smith, G. R., Kunes, S. M., Schultz, D. W., Taylor, A., and Triman, K. L. (1981). Structure of Chi Hotspots of Generalized Recombination. *Cell* 24 (2), 429–436. doi:10.1016/0092-8674(81)90333-0

Spies, M., Amitani, I., Baskin, R. J., and Kowalczykowski, S. C. (2007). RecBCD Enzyme Switches Lead Motor Subunits in Response to χ Recognition. *Cell* 131 (4), 694–705. doi:10.1016/j.cell.2007.09.023

Sun, Z. Z., Yeung, E., Hayes, C. A., Noireaux, V., and Murray, R. M. (2013). Linear DNA for Rapid Prototyping of Synthetic Biological Circuits in anEscherichia coliBased TX-TL Cell-free System. *ACS Synth. Biol.* 3, 387–397. doi:10.1021/sb400131a

Swartz, J. (2006). Developing Cell-free Biology for Industrial Applications. *J. Ind. Microbiol. Biotechnol.* 33 (7), 476–485. doi:10.1007/s10295-006-0127-y

Takahashi, M. K., Tan, X., Dy, A. J., Braff, D., Akana, R. T., Furuta, Y., et al. (2018). A Low-Cost Paper-Based Synthetic Biology Platform for Analyzing Gut Microbiota and Host Biomarkers. *Nat. Commun.* 9, 1–12. doi:10.1038/s41467-018-05864-4

Thavarajah, W., Silverman, A. D., Verosloff, M. S., Kelley-Loughnane, N., Jewett, M. C., and Lucks, J. B. (2020). Point-of-Use Detection of Environmental Fluoride via a Cell-free Riboswitch-Based Biosensor. *ACS Synth. Biol.* 9 (1), 10–18. doi:10.1021/acssynbio.9b00347

Thoring, L., Wüstenhagen, D. A., Borowiak, M., Stech, M., Sonnabend, A., and Kubick, S. (2016). Cell-Free Systems Based on CHO Cell Lysates: Optimization Strategies, Synthesis of "Difficult-To-Express" Proteins and Future Perspectives. *PLOS ONE* 11 (9), e0163670. doi:10.1371/journal.pone.0163670

Verosloff, M., Chappell, J., Perry, K. L., Thompson, J. R., and Lucks, J. B. (2019). PLANT-dx: A Molecular Diagnostic for Point-of-Use Detection of Plant Pathogens. *ACS Synth. Biol.* 8 (4), 902–905. doi:10.1021/acssynbio.8b00526

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 (1), 1–7. doi:10.1093/synbio/ysy003

Wang, J., Yang, L., Cui, X., Zhang, Z., Dong, L., and Guan, N. (2017). A DNA Bubble-Mediated Gene Regulation System Based on Thrombin-Bound DNA Aptamers. *ACS Synth. Biol.* 6 (5), 758–765. doi:10.1021/acssynbio.6b00391

Watters, K. E., Fellmann, C., Bai, H. B, Ren, S. M., and Doudna, J. A. (2018). Systematic Discovery of Natural CRISPR-Cas12a Inhibitors. *Science* 362 (6411), 236–239. doi:10.1126/science.aau5138

Welsh, J. P., Lu, Y., He, X.-S., Greenberg, H. B., and Swartz, J. R. (2012). Cell-Free Production of Trimeric Influenza Hemagglutinin Head Domain Proteins as Vaccine Antigens. *Biotechnol. Bioeng.* 109 (12), 2962–2969. doi:10.1002/bit.24581

Wick, S., Walsh, D. I., Bobrow, J., Hamad-Schifferli, K., Kong, D. S., Thorsen, T., et al. (2019). PERSIA for Direct Fluorescence Measurements of Transcription, Translation, and Enzyme Activity in Cell-free Systems. *ACS Synth. Biol.* 8 (5), 1010–1025. doi:10.1021/acssynbio.8b00450

Wiegand, D. J., Lee, H. H., Ostrov, N., and Church, G. M. (2019). Cell-Free Protein Expression Using the Rapidly Growing Bacterium Vibrio Natriegens. *JoVE* 145, 1–14. doi:10.3791/59495

Wu, P. S. C., Ozawa, K., Lim, S. P., Vasudevan, S. G., Dixon, N. E., and Otting, G. (2007). Cell-Free Transcription/Translation from PCR-Amplified DNA for High-Throughput NMR Studies. *Angew. Chem. Int. Ed.* 46 (18), 3356–3358. doi:10.1002/anie.200605237

Wuu, J. J., and Swartz, J. R. (2008). High Yield Cell-free Production of Integral Membrane Proteins without Refolding or Detergents. *Biochim. Biophys. Acta (Bba) - Biomembranes* 1778 (5), 1237–1250. doi:10.1016/j.bbamem.2008.01.023

Yelleswarapu, M., van der Lindenvan der Linden, A. J., van Sluijs, B., Pieters, P. A., Dubuc, E., de Greef, T. F. A., et al. (2018). Sigma Factor-Mediated Tuning of Bacterial Cell-free Synthetic Genetic Oscillators. *ACS Synth. Biol.* 7 (12), 2879–2887. doi:10.1021/acssynbio.8b00300

Yeung, E., Dy, A. J., Martin, K. B., Ng, A. H., Del Vecchio, D., Beck, J. L., et al. (2017). Biophysical Constraints Arising from Compositional Context in Synthetic Gene Networks. *Cell Syst.* 5 (1), 11–24.e12. doi:10.1016/j.cels.2017.06.001

Yim, S. S., Johns, N. I., Vincent, N., and Wang, H. H. (2020). Protecting Linear DNA Templates in Cell-free Expression Systems from Diverse Bacteria. *ACS Synth. Biol.* 9 (10), 2851–2855. doi:10.1021/acssynbio.0c00277

Yu, D., Ellis, H. M., Lee, E.-C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2000). An Efficient Recombination System for Chromosome Engineering in Escherichia Coli. *Proc. Natl. Acad. Sci.* 97 (11), 5978–5983. doi:10.1073/pnas.100127597

Zawada, J. F., Yin, G., Steiner, A. R., Yang, J., Naresh, A., Roy, S. M., et al. (2011). Microscale to Manufacturing Scale-up of Cell-free Cytokine Production-A New Approach for Shortening Protein Production Development Timelines. *Biotechnol. Bioeng.* 108 (7), 1570–1578. doi:10.1002/bit.23103

Zawada, J., and Swartz, J. (2006). Effects of Growth Rate on Cell Extract Performance in Cell-free Protein Synthesis. *Biotechnol. Bioeng.* 94 (4), 618–624. doi:10.1002/bit.20831

Zhang, Y., Kojima, T., Kim, G.-A., McNerney, M. P., Takayama, S., and Styczynski, M. P. (2021). Protocell Arrays for Simultaneous Detection of Diverse Analytes. *bioRxiv*. doi:10.1101/2021.02.13.431022

Zhang, Y., Lai, B., and Juhas, M. (2019). Recent Advances in Aptamer Discovery and Applications. *Molecules* 24 (5), 941. doi:10.3390/molecules24050941

Zhou, B., Dong, Q., Ma, R., Chen, Y., Yang, J., Sun, L.-Z., et al. (2009). Rapid Isolation of Highly Pure Single-Stranded DNA from Phagemids. *Anal. Biochem.* 389 (2), 177–179. doi:10.1016/j.ab.2009.03.044

Zhu, B., Gan, R., Cabezas, M. D., Kojima, T., Nicol, R., Jewett, M. C., et al. (2020). Increasing Cell-free Gene Expression Yields from Linear Templates in *Escherichia coli* and Vibrio Natriegens Extracts by Using DNA-binding Proteins. *Biotechnol. Bioeng.* 117 (12), 3849–3857. doi:10.1002/bit.27538

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

ORIGINAL RESEARCH
published: 28 July 2021
doi: 10.3389/fbioe.2021.730663

Designing Modular Cell-free Systems for Tunable Biotransformation of L-phenylalanine to Aromatic Compounds

Chen Yang[1,2,3], *Yushi Liu*[1], *Wan-Qiu Liu*[1], *Changzhu Wu*[4] *and Jian Li*[1]*

[1]School of Physical Science and Technology, ShanghaiTech University, Shanghai, China, [2]Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, China, [3]University of Chinese Academy of Sciences, Beijing, China, [4]Danish Institute for Advanced Study (DIAS) and Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Odense, Denmark

Cell-free systems have been used to synthesize chemicals by reconstitution of *in vitro* expressed enzymes. However, coexpression of multiple enzymes to reconstitute long enzymatic pathways is often problematic due to resource limitation/competition (e.g., energy) in the one-pot cell-free reactions. To address this limitation, here we aim to design a modular, cell-free platform to construct long biosynthetic pathways for tunable synthesis of value-added aromatic compounds, using (*S*)-1-phenyl-1,2-ethanediol ((*S*)-PED) and 2-phenylethanol (2-PE) as models. Initially, all enzymes involved in the biosynthetic pathways were individually expressed by an *E. coli*-based cell-free protein synthesis (CFPS) system and their catalytic activities were confirmed. Then, three sets of enzymes were coexpressed in three cell-free modules and each with the ability to complete a partial pathway. Finally, the full biosynthetic pathways were reconstituted by mixing two related modules to synthesize (*S*)-PED and 2-PE, respectively. After optimization, the final conversion rates for (*S*)-PED and 2-PE reached 100 and 82.5%, respectively, based on the starting substrate of L-phenylalanine. We anticipate that the modular cell-free approach will make a possible efficient and high-yielding biosynthesis of value-added chemicals.

Keywords: cell-free systems, biocatalysis, biotransformation, value-added chemicals, synthetic biology

OPEN ACCESS

Edited by:
Zhiguang Zhu,
Tianjin Institute of Industrial Biotechnology (CAS), China

Reviewed by:
Guo Daoyi,
Gannan Normal University, China
Chun You,
Tianjin Institute of Industrial Biotechnology (CAS), China

***Correspondence:**
Jian Li
lijian@shanghaitech.edu.cn

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

Received: *25 June 2021*
Accepted: *19 July 2021*
Published: *28 July 2021*

Citation:
Yang C, Liu Y, Liu W-Q, Wu C and Li J (2021) Designing Modular Cell-free Systems for Tunable Biotransformation of L-phenylalanine to Aromatic Compounds. Front. Bioeng. Biotechnol. 9:730663. doi: 10.3389/fbioe.2021.730663

INTRODUCTION

Biotransformation is a green and sustainable approach for the production of valuable chemicals, pharmaceuticals, and materials, among others (Bornscheuer et al., 2012; Sheldon and Woodley, 2018; Wu and Li, 2018; Wang et al., 2021; Wu et al., 2021). For decades, microbial organisms have been predominantly engineered to synthesize those compounds of interest from cost-effective carbon sources like renewable biomass (e.g., cellulose and starch, etc.) and one-carbon feedstocks (e.g., formate and CO_2, etc.) (Humphreys and Minton, 2018; Cotton et al., 2020; Ko et al., 2020). While cell-based biotransformation is a promising means for the production, efforts to design and engineer living organisms are constrained by slow design-build-test (DBT) cycles (Nielsen and Keasling, 2016). In addition, overexpression of a set of exogenous enzymes to perform *in vivo* chemical conversions imposes a strong metabolic burden on heterologous microbial hosts. Another key challenge in cellular production is the balance (conflict of resource allocation) between the cell's

survival objectives (i.e., native metabolism and growth) and the engineer's production goals (i.e., synthesis of target products), making the final product yields in cellular systems often low and not satisfactory (Li and Neubauer, 2014; Wu et al., 2016). Therefore, designing and building efficient systems for biotransformation and maximization of product yields is highly desirable.

As a complement to cell-based approaches, cell-free systems have emerged as promising and powerful platforms for biomanufacturing (Bundy et al., 2018; Li et al., 2018; Swartz, 2018; Liu et al., 2019; Bowie et al., 2020; Silverman et al., 2020; Rasor et al., 2021). The unique feature of cell-free systems is that cell growth and product synthesis are separated, thus moving biosynthesis away from living cells and converting input resource/energy into desired products at high yields. Due to the open nature of cell-free systems, the reaction environment can be easily manipulated, directly accessed, and rapidly optimized. Furthermore, cell-free systems without cellular barriers can bypass transfer limitations and are more tolerant of toxic molecules (e.g., substrates, pathway intermediates, and final products) than living microbial cells. Recently, crude extract based cell-free protein synthesis (CFPS) systems were well developed and their applications have been expanded from single protein synthesis to multiple enzyme coexpression (Kwon et al., 2013; Li et al., 2016; Goering et al., 2017; Moore et al., 2021). The CFPS-expressed enzymes without purification can carry out *in situ* cascade biotransformation to synthesize a wide variety of molecules, such as *n*-butanol (Karim and Jewett, 2016), diketopiperazine (Goering et al., 2017), polyhydroxyalkanoate (Kelwick et al., 2018), styrene (Grubbe et al., 2020), 3-hydroxybutyrate (Karim et al., 2020), limonene (Dudley et al., 2020), L-theanine (Feng et al., 2021), and the complex natural product valinomycin (Zhuang et al., 2020). Notably, the yields of the CFPS-based production platform are often higher than *in vivo* production. For example, CFPS systems improved the production of styrene and valinomycin both by more than two times as compared to recombinant cellular systems (Grubbe et al., 2020; Zhuang et al., 2020).

Despite cell-free systems have many advantages and have been applied to reconstitute *in vitro* metabolic pathways for biotransformations, the reported complete pathways in CFPS systems were normally short with less than four enzymes. This is mainly because some full pathways are too long–with many enzymes–to express in a single-pot CFPS. To alleviate the tension of all enzyme coexpression in one pot, partial enzymes of the pathway can be individually expressed *in vivo* to prepare enzyme enriched cell extracts, which then can be used for CFPS to express the rest enzymes to assemble the full biosynthetic pathway (Dudley et al., 2020; Karim et al., 2020; Zhuang et al., 2020). However, this strategy is laborious and time-consuming that includes *in vivo* overexpression of each pathway enzyme, preparation of multiple cell lysates, and *in vitro* mix-and-match of lysate combination. As comparison to *in vivo* expression (days), CFPS systems can significantly reduce the expression time (hours) from DNA (plasmids) to functional enzymes. Therefore, with the rapid property of CFPS reactions, the overall time for enzymatic biotransformation will also be notably reduced if all enzymes of a long pathway can be coexpressed by CFPS.

Recently, we reported the use of CFPS to coexpress three enzymes for the bioconversion of styrene to (*S*)-1-phenyl-1,2-ethanediol with a high conversion rate (Liu et al., 2020). However, the substrate styrene is water-insoluble. Then, an aqueous-organic biphasic system was used to support the enzyme expression and the resulting cascade biotransformation. While this system worked well, an organic solvent like the toxic toluene was required to dissolve styrene, which is not a green or environmentally friendly approach. To avoid using organic solvents, a metabolic pathway is thus needed to synthesize styrene from possible water soluble substrates as an intermediate, which can be enzymatically converted to the downstream products. Such pathway is available by using three enzymes to convert L-phenylalanine (L-Phe) to styrene (Lukito et al., 2019). However, this solution will bring more enzymes to coexpress in one CFPS reaction and as a result coexpression of all enzymes might be problematic. Thus, a rational strategy is needed to solve the issue of multiple enzyme coexpression in a one-pot CFPS reaction.

In this work, we aim to address the challenge by designing modular cell-free systems for tunable biotransformation, which can be used to convert the same precursor L-Phe to two value-added aromatic compounds without using the water-insoluble substrate styrene (**Figure 1**). The first target compound is the chiral drug intermediate (*S*)-1-phenyl-1,2-ethanediol ((*S*)-PED) and the second one is the rose-like fragrance 2-phenylethanol (2-PE) (Peng et al., 2019; Wang et al., 2019). These two aromatic compounds share a common upstream intermediate styrene. To achieve our goal, here we design three CFPS modules. In the first module, an artificial pathway from L-Phe to styrene is constructed by coexpression of three enzymes, namely, phenylalanine ammonia lyase 2 (PAL2) from *Arabidopsis thaliana* (Cochrane et al., 2004) and phenylacrylic acid decarboxylase (PAD, containing two enzymes Fdc1 and Pad1) from *Aspergillus niger* (Payne et al., 2015). The second module uses styrene generated in the first module to synthesize (*S*)-PED, where another three enzymes are coexpressed to perform the cascade reaction. The related enzymes include styrene monooxygenase (SMO, consisting of two subunit enzymes StyA and StyB) from *Pseudomonas* sp. VLB120 (Panke et al., 1998) and epoxide hydrolase (SpEH) from *Sphingomonas* sp. HXN-200 (Wu et al., 2013). In the third module, three enzymes (StyA, StyB, and StyC) involved in the styrene degradation pathway (Panke et al., 1998) are employed to convert styrene to phenylacetaldehyde, which can be reduced to the final product 2-PE by endogenous alcohol dehydrogenases (ADHs). Our data suggest that the designer CFPS modules are efficient for long enzymatic cascade reactions, enabling high conversion rates of L-Phe to (*S*)-PED and 2-PE at 100 and 82.5%, respectively. Looking forward, we anticipate this work will expand the application space of cell-free systems by designing modular CFPS reactions to carry out long cascade biotransformations for efficiently synthesizing valuable chemicals when the whole-cell bioconversion remains difficult.

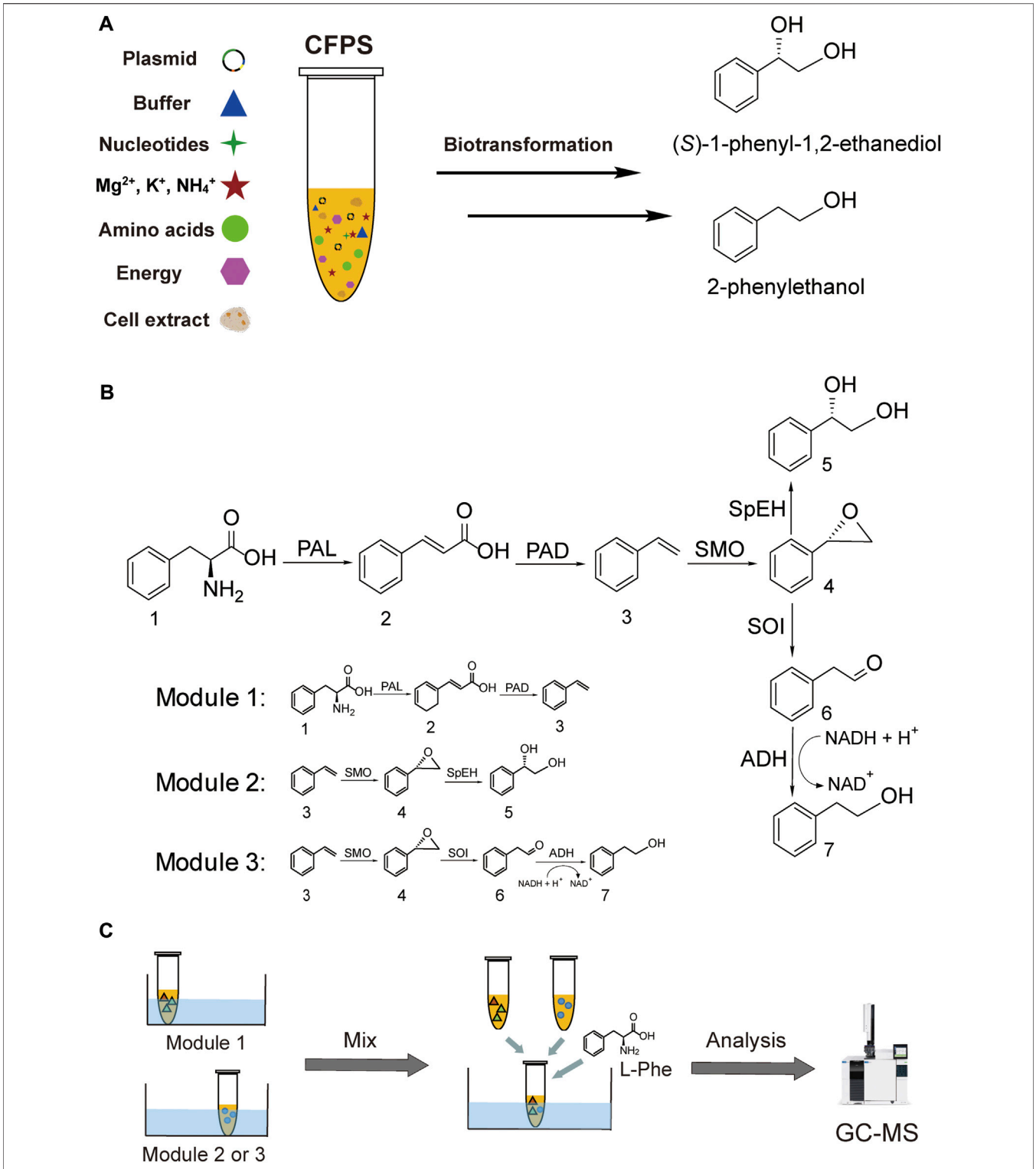


FIGURE 1 | Cell-free synthesis of (*S*)-PED and 2-PE. **(A)** CFPS reaction and the synthesis of target chemicals. **(B)** Overall enzymatic pathways and designing of three modules. **(C)** Modular cell-free systems for the product formation. Abbreviations: CFPS, cell-free protein synthesis; PAL, phenylalanine ammonia lyase; PAD, phenylacrylic acid decarboxylase; SMO, styrene monooxygenase; SpEH, epoxide hydrolase; SOI, styrene oxide isomerase (also called StyC); ADH, alcohol dehydrogenase. Chemicals 1–7: 1, L-phenylalanine (L-Phe); 2, cinnamic acid; 3, styrene; 4, styrene oxide; 5 (*S*)-1-phenyl-1,2-ethanediol ((*S*)-PED); 6, phenylacetaldehyde; 7, 2-phenylethanol (2-PE).

MATERIALS AND METHODS

Chemicals

L-phenylalanine was purchased from Sigma-Aldrich (St. Louis, United States). Styrene, styrene oxide, (*S*)-1-phenyl-1,2-ethanediol, and 2-phenylethanol were purchased from Adamas (Shanghai, China). Cinnamic acid and phenylacetaldehyde were obtained from Aladdin (Shanghai, China). The plasmid miniprep kit was purchased from Sangon Biotech (Shanghai, China). All other chemical reagents were of the highest purity available.

Strains and Cultivation Media

Escherichia coli DH5α was used for molecular cloning and plasmid propagation. *E. coli* BL21 Star (DE3) was grown for cell extract preparation to perform CFPS reactions. General cultivation of *E. coli* strains was carried out in the LB medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl). The 2xYTPG medium (10 g/L yeast extract, 16 g/L tryptone, 5 g/L NaCl, 7 g/L K_2HPO_4, 3 g/L KH_2PO_4, and 18 g/L glucose, pH 7.2) was used to grow *E. coli* cells for cell extract preparation.

Plasmid Construction

The plasmids pET28a-StyA, pET28a-StyB, and pET28a-SpEH used for the expression of StyA, StyB, and SpEH, respectively, were reported in our previous work (Liu et al., 2020). All other genes were codon optimized for *E. coli*, synthesized, and cloned into the plasmid pET28a by GENEWIZ (Suzhou, China), including *PAL2* (GenBank: AY303129), *Fdc1* (GenBank: XM_001390497), *Pad1* (GenBank: XM_001390495), and *StyC* (GenBank: AF031161). The gene *StyC* encodes a styrene oxide isomerase (StyC) in the styrene degradation pathway (Panke et al., 1998). To construct fused StyA and StyB, a flexible linker sequence [AS $(GGGGS)_5$GAS] was inserted between the amino acid sequences of the two enzymes and cloned into pET28a, generating the plasmid pET28a-StyA-linker-StyB. All genes were inserted to the pET28a backbone between two restriction sites NdeI and SalI (see **Supplementary Figure S1** for plasmid maps).

Preparation of Cell Extracts

Cell growth, collection, and extracts were prepared as described previously (Liu et al., 2020). In brief, *E. coli* BL21 Star (DE3) cells were grown in 1L of 2xYTPG media, induced with 1 mM IPTG (OD_{600} = 0.6–0.8) to express T7 RNA polymerase, and harvested by centrifugation at an OD_{600} of around 3.0. Then, cell pellets were washed three times with cold S30 Buffer (10 mM Tris-acetate, 14 mM magnesium acetate, and 60 mM potassium acetate). After the final wash and centrifugation, the pelleted cells were resuspended in S30 Buffer (1 ml per Gram wet cell mass) and lysed by sonication (10 s on/off, 50% of amplitude, input energy ~600 J) on ice. The lysate was then centrifuged twice at 12,000 g and 4°C for 10 min. The resulting supernatant (i.e., cell extracts) was flash frozen in liquid nitrogen and stored at −80°C until use.

Cell-free Protein Synthesis Reactions

Standard CFPS reactions were carried out in 1.5-ml microcentrifuge tubes. Each reaction (15 μL) contains the following components: 12 mM magnesium glutamate, 10 mM ammonium glutamate, 130 mM potassium glutamate, 1.2 mM ATP, 0.85 mM each of GTP, UTP, and CTP, 34 μg/ml folinic acid, 170 μg/ml of *E. coli* tRNA mixture, 2 mM each of 20 standard amino acids, 0.33 mM nicotinamide adenine dinucleotide (NAD), 0.27 mM coenzyme A (CoA), 1.5 mM spermidine, 1 mM putrescine, 4 mM sodium oxalate, 33 mM phosphoenolpyruvate (PEP), 13.3 μg/ml plasmid unless otherwise noted, and 27% (v/v) of cell extracts. T7 RNA polymerase was not added to CFPS reactions since it was induced with IPTG during cell growth and contained in the cell extracts. CFPS expressed proteins were analyzed by SDS-PAGE and Western-blot.

Modular Cell-free Protein Synthesis Reactions

To assemble enzymatic pathways for the synthesis of (*S*)-PED and 2-PE, one upstream module (module 1) and two downstream modules (module 2 and module 3) were constructed in cell-free systems. Each module of the CFPS reaction was performed at 30°C with a total volume of 15 μL in the 1.5-ml tube. In the first module, three enzymes were coexpressed by adding 13.3 μg/ml of pET28a-PAL2, 13.3 μg/ml of pET28a-Fdc1, and 6.7 μg/ml of pET28a-Pad1, respectively. This upstream CFPS module was incubated for 2 h to express enzymes for the conversion of L-Phe to styrene. In module 2, StyA, StyB, and SpEH were coexpressed with plasmid concentrations of 13.3, 3.3, and 13.3 μg/ml, respectively, for 6 h before mixing with module 1 to convert styrene to (*S*)-PED. Similarly, in module 3, StyA and StyB were expressed the same as in module 2; StyC was expressed with 13.3 μg/ml of pET28a-StyC. Coexpression of three enzymes in module 3 was also carried out for 6 h before mixing with module 1. Cell-free expressed StyA, StyB, and StyC together with endogenous ADHs can convert styrene to 2-PE. After mixing module 1 with module 2 or 3, 1 mM of the substrate L-Phe was added to the CFPS mixture and the reaction was further carried out for 16 h at 30°C to synthesize target compounds.

Analytical Methods

All intermediates and final products were extracted with 2 volume of ethyl acetate for twice (as thoroughly as possible), followed by analysis with GC-MS (Trace 1300-ISQ, ThermoFisher Scientific) using a TG-5MS (30 m × 0.25 mm x 0.25 μm) column. For (*S*)-PED and 2-PE analysis, 1 μL of each sample was injected with a split ratio of 33.3:1. Helium was used as the carrier gas at a flow rate of 1.2 ml/min. The initial column temperature was 40°C and then increased at 30°C/min to 270°C (for (*S*)-PED) or 200°C (for 2-PE), which was maintained for 3 min. Afterwards, the column temperature increased at 30°C/min to 300°C with a final 5 min hold. The retention times of (*S*)-PED and 2-PE were 6.04 and 5.07 min, respectively. The concentration of each product was determined by comparison to a linear standard curve generated

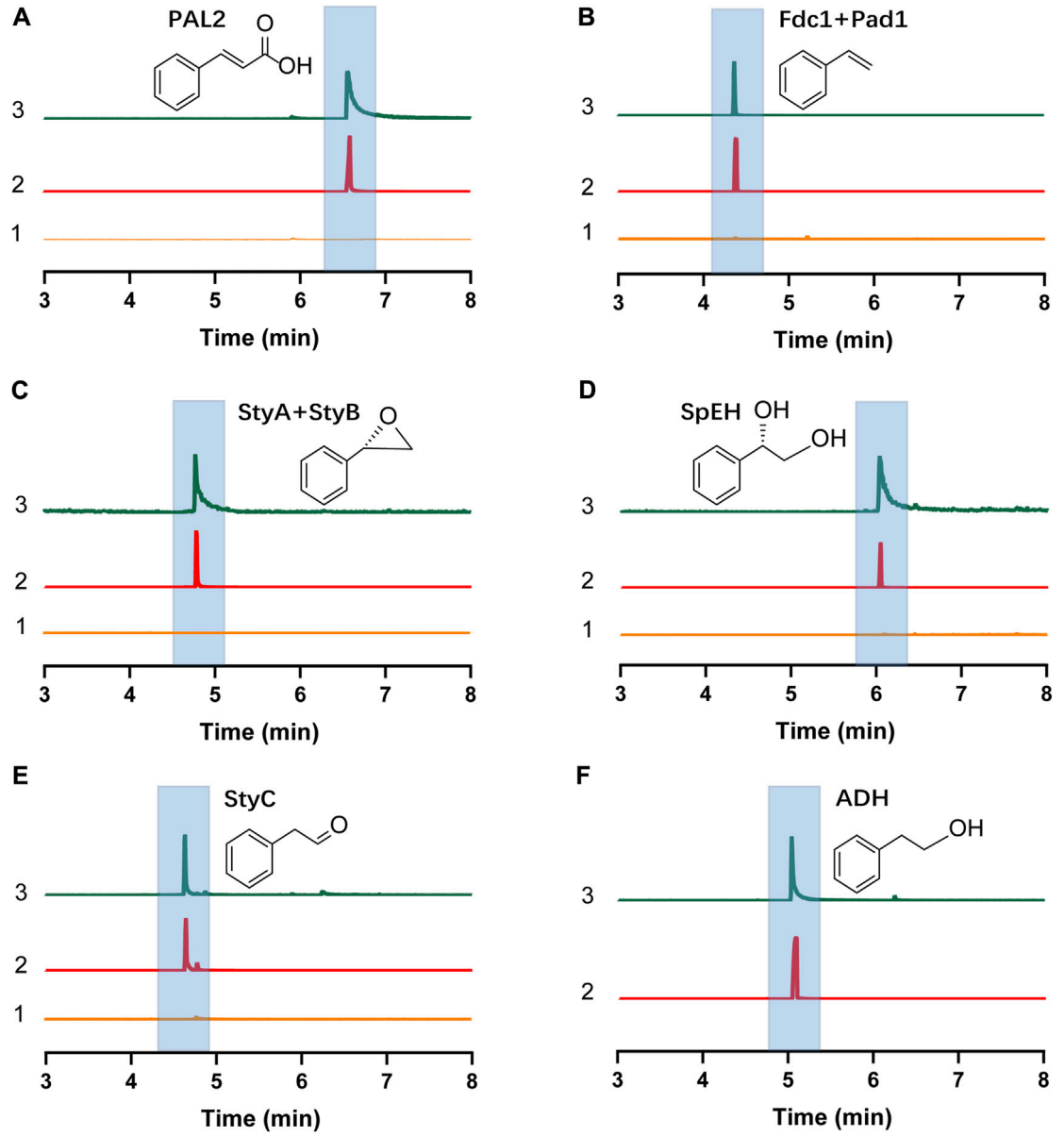


FIGURE 2 | GC-MS detection of the product of each step enzymatic reaction. **(A)** Cinnamic acid. **(B)** Styrene. **(C)** Styrene oxide. **(D)** *(S)*-PED. **(E)** Phenylacetaldehyde. **(F)** 2-PE. In each panel, the numbers 1, 2, and 3 represent negative control without plasmid in the CFPS reaction (substrate was added), chemical standard, and cell-free reaction sample, respectively. Note that in **(F)**, no exogenous enzymes were expressed in CFPS and only *E. coli* endogenous ADHs catalyzed the reaction.

with a commercial standard (**Supplementary Figure S2**). All measurements were performed in triplicate.

RESULTS AND DISCUSSION

Cell-free Expression of Each Enzyme and Demonstration of Their Catalytic Activities

To synthesize (*S*)-PED and 2-PE from the same precursor L-Phe, we chose in total seven enzymes from different organism sources to construct two full enzymatic pathways (**Figure 1B**). First, L-Phe is converted to styrene, which is a shared upstream intermediate, by successive deamination and decarboxylation with three enzymes PAL2, Fdc1, and Pad1 (Cochrane et al., 2004; Payne et al., 2015). Then, styrene undergoes two divergent pathways to form (*S*)-PED and 2-PE, respectively, with enzymes StyA, StyB, StyC, SpEH, and *E. coli* endogenous ADHs (Panke et al., 1998; Wu et al., 2013). Prior to build *in vitro* full pathways, each enzyme except the endogenous ADHs has to be actively expressed in CFPS reactions. To this end, the well-developed *E. coli*-based CFPS system was used to express all

enzymes and test their catalytic activities. Note that the conversion of cinnamic acid to styrene requires two enzymes Fdc1 and Pad1 originated from *A. niger* (Payne et al., 2015); in addition, the styrene monooxygenase (SMO) from *Pseudomonas* sp. VLB120 also consists of two subunit enzymes StyA and StyB to oxidize styrene to styrene oxide (Panke et al., 1998). Therefore, these two enzymatic conversion steps need the coexpression of two related enzymes in one-pot CFPS reactions.

The *E. coli* CFPS system has been used to express various proteins, for example, therapeutic proteins, membrane proteins, and non-natural amino acid modified proteins (Henrich et al., 2015; Martin et al., 2018; Wilding et al., 2019), demonstrating the robustness of *E. coli* CFPS for *in vitro* protein production. To see if all enzymes we selected can be expressed, we used the *E. coli* strain BL21 Star (DE3) to prepare cell extracts for protein expression. We observed that each enzyme was successfully expressed in CFPS as shown by the Western-blot analysis (**Supplementary Figure S3**). Then, the catalytic ability of each enzyme was investigated. To do this, the substrate of each enzyme was added to the reaction mixture when the enzyme expression started. As shown in **Figures 2A–E**, each target compound was detected by GC-MS, demonstrating that all cell-free expressed enzymes were active to convert their substrates to the corresponding products (see **Supplementary Figure S4** for MS spectra of all compounds). In addition, *E. coli* endogenous ADHs were also active to reduce phenylacetaldehyde to 2-PE when no exogenous enzymes were expressed by CFPS (**Figure 2F**). Taken together, the results demonstrated the potential of using CFPS to express functional enzymes to reconstitute (*S*)-PED and 2-PE biosynthetic pathways *in vitro*. As the goal of this work was to establish modular cell-free systems for tunable biotransformation, we next sought to coexpress different sets of enzymes in different single pot CFPS reactions and then mix them for synthesizing final products (**Figure 1C**).

Designing Modular Cell-free Systems for the Synthesis of (S)-PED and 2-PE

After demonstrating the enzyme activity, we next aimed to design CFPS modules for the expression of different enzyme combinations. The key idea was to express short pathways first and then mix them together to assemble the full pathways. Since styrene is the same upstream intermediate of (*S*)-PED and 2-PE, three enzymes (i.e., PAL2, Fdc1, and Pad1) were coexpressed in the first module of CFPS to convert L-Phe to styrene, which can be used by the downstream modules as a substrate. In module 2, styrene undergoes epoxidation and hydrolysis by SMO (i.e., StyA and StyB) and SpEH, respectively, to form (*S*)-PED. While cell-free expressed StyA, StyB, and StyC in module 3 catalyze the cascade conversion of styrene to phenylacetaldehyde, which can be reduced to 2-PE by endogenous ADHs. Initially, cell-free synthesized enzymes in each module were confirmed by Western-blot (**Supplementary Figure S5**). Next, biotransformations in all three modules were tested individually by adding 1 mM of substrate to the CFPS reactions. After the reaction, each product was extracted from the reaction mixture and analyzed by GC-MS. The results showed that all three target compounds of styrene, (*S*)-PED, and 2-PE were detected from the related modules (**Figure 3**), suggesting that our designed cell-free modules are capable of assembling long, full pathways for tunable biotransformation.

We, therefore, set out to build cell-free pathway combinations for (*S*)-PED and 2-PE synthesis. To achieve this goal, we first ran three cell-free modular reactions separately at 30°C to express enzymes and then mixed module 1 with module 2 or 3. While the time for enzyme expression in modules 2 and 3 was 6 h, the upstream module 1 was only incubated for 2 h before mixing because we observed that 2-h CFPS reaction was the best for module 1 to convert L-Phe to styrene (**Supplementary Figure S6**). When mixing two modules, 1 mM of L-Phe was added to the cell-free mixture and the reaction was further carried out for 16 h at 30°C to synthesize (*S*)-PED and 2-PE. The GC-MS analyses suggested that in both cases the final products were successfully detected (**Supplementary Figure S7**), demonstrating the ability of using modular cell-free reactions each with several CFPS-expressed enzymes to assemble long metabolic pathways *in vitro*.

Effect of L-Phe on the Synthesis of (S)-PED and 2-PE

Having validated the ability to mix cell-free modules for biotransformation, we next wanted to investigate the effect of L-Phe on the synthesis of (*S*)-PED and 2-PE, respectively. To do so, we mixed module 1 with module 2 or 3 after cell-free enzyme expression and added L-Phe to the reaction mixture at different final concentrations from 0.5 to 5 mM. Our data indicated that the effect of L-Phe on the synthesis of (*S*)-PED was more obvious than that on the synthesis of 2-PE (**Figure 4**). The highest titer of (*S*)-PED reached 0.82 mM when L-Phe was supplied with 2 mM, which is 2.4-fold higher than the titer achieved with 0.5 mM L-Phe (**Figure 4A**). By contrast, the titers of 2-PE were similar with different L-Phe concentrations, although 2 mM L-Phe gave rise to a bit higher titer of 2-PE as compared to all other concentrations (**Figure 4B**). In order to further optimize our system for enhanced production, we chose to add 2 mM of L-Phe to the reactions in our following experiments.

Optimization of Reaction Temperature and Volume on the Synthesis of (S)-PED and 2-PE

Cell-free reaction temperature is an important factor for optimization because it affects protein expression and enzyme activity. Our cell-free biosynthesis system includes two phases: the initial CFPS reaction phase and the second enzymatic biotransformation phase. While the best temperature for protein expression in *E. coli* CFPS system has been found to be 30°C (Goering et al., 2017; Li et al., 2016; Martin et al., 2018), whether this temperature is optimum for the enzymatic reactions remaining unknown. Therefore, we did not focus on the optimization of the CFPS temperature in this work. Instead, we conducted our optimization by investigating the optimal reaction temperature during the biotransformation stage for producing (*S*)-PED and 2-PE. After mixing two cell-free modules, we incubated the reactions

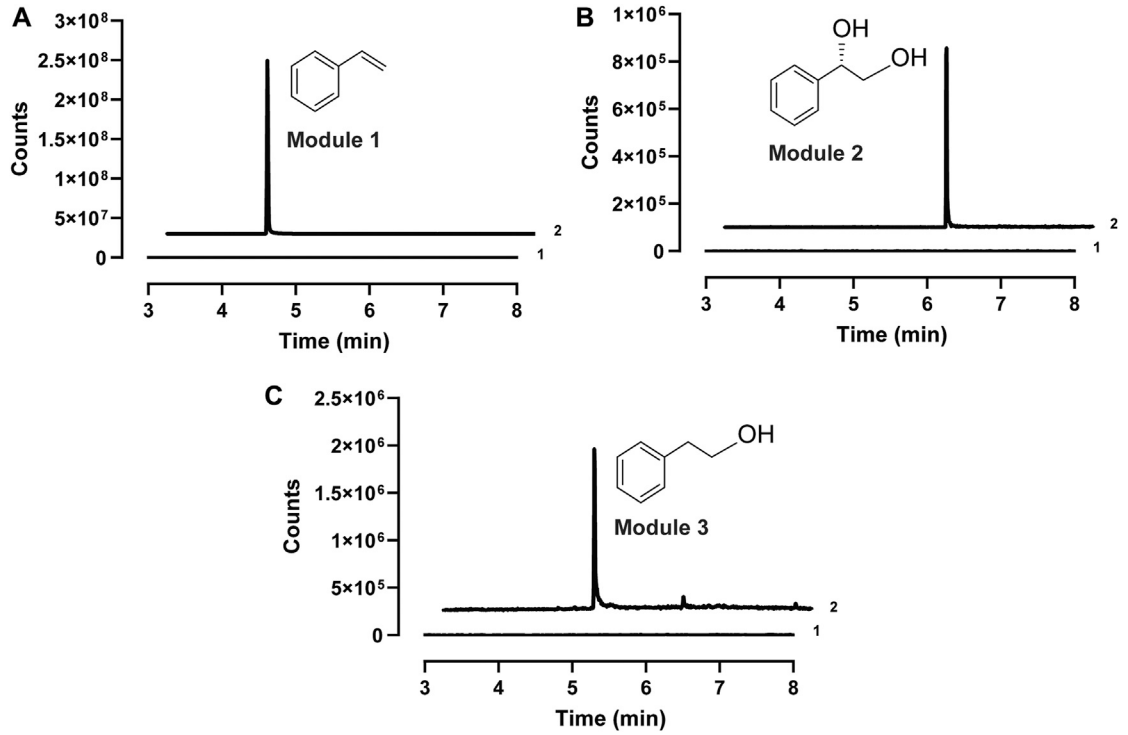


FIGURE 3 | Extracted ion chromatograms (EIC) of the GC-MS detection of **(A)** styrene (m/z 140) in module 1, **(B)** (*S*)-PED (m/z 138) in module 2, and **(C)** 2-PE (m/z 122) in module 3. In each panel, 1, sample at the reaction time of zero; 2, sample from the reaction finished.

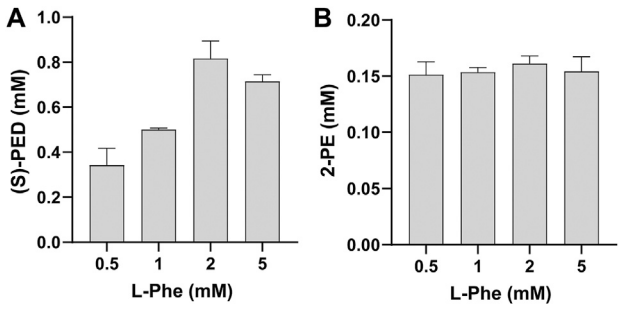


FIGURE 4 | Effect of L-Phe on the synthesis of **(A)** (*S*)-PED and **(B)** 2-PE. Values show means with error bars representing standard deviations (s.d.) of at least three independent experiments.

containing 2 mM of L-Phe at 23, 30, and 37°C. The results indicated that low temperatures favored the product formation and the maximum titers of both products were obtained at 30°C (**Figures 5A,B**). This makes the whole production process easy without the change of incubation temperature from protein expression to product synthesis. Our finding is similar to previous reports that 30°C is the optimal temperature to synthesize enzymes and the resulting chemical compounds (Goering et al., 2017; Grubbe et al., 2020; Liu et al., 2020).

Next, we evaluated the effect of reaction volume on the synthesis of (*S*)-PED and 2-PE. We ran standard CFPS reactions in 15 μL volume and also scaled up the volume to 90 μL per reaction. Thus, the total volume after mixing two modules was 30 μL or 180 μL. For the group of reaction volume with 180 μL, we also compared the reactions without shaking and with shaking at 600 rpm. All enzymatic reactions were carried out at 30°C for 16 h and the final products were extracted and analyzed. Interestingly, we found that the reaction volume has different effects on the product titers (**Figures 5C,D**). For (*S*)-PED, the titer obtained from the reaction volume of 30 μL was higher than that of the volume of 180 μL. By contrast, the titer of 2-PE in the reaction volume of 180 μL was maximum, but shaking was not necessary. In both cases, the product titers in 180 μL volume with shaking were lower than those without shaking. While the impact of reaction volume on product formation was different, our results suggested that 1) the reaction volume should be investigated for each new enzymatic pathway *in vitro* and 2) cell-free system is flexible for manipulation and optimization. With the aim to improve each product titer, we decided to use 30 and 180 μL reaction volumes to synthesize (*S*)-PED and 2-PE, respectively, in the following studies.

Enhancing the Epoxidation Step of Styrene for Efficient Biotransformation

Styrene accumulated in the first cell-free module is the upstream precursor for the synthesis of (*S*)-PED in module 2 and 2-PE in module 3. It is, therefore, important to efficiently

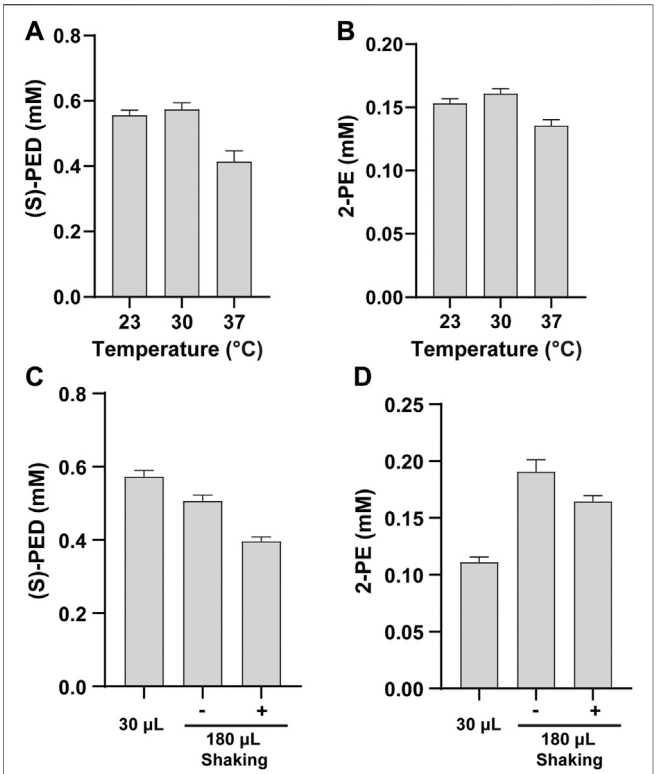


FIGURE 5 | Effects of **(A, B)** reaction temperature and **(C, D)** reaction volume on the synthesis of (*S*)-PED and 2-PE, respectively. Values show means with error bars representing standard deviations (s.d.) of at least three independent experiments.

convert the water-insoluble styrene to styrene oxide, which can be used by the subsequent enzymes for further conversion. The epoxidation of styrene is catalyzed by the enzyme SMO (i.e., StyA and StyB). Previous studies have shown that StyA plays a core role in the biocatalytic activity of SMO and StyB helps maximize the epoxidation of styrene (Panke et al., 1998; Otto et al., 2004). In addition, the highest epoxidation activity can be achieved when the molar ratio of StyA and StyB reaches at about 1:1 (Otto et al., 2004; Corrado et al., 2018). Based on these reports, we were curious to know if tuning StyA and StyB might enhance the epoxidation activity of styrene. Thus, we subsequently sought to test two strategies to see their effects on the efficiency of cell-free biotransformations. First, we overexpressed the key enzyme StyA *in vivo* in *E. coli* rather than in CFPS and prepared StyA-enriched cell extracts to coexpress the rest enzymes, for example, StyB, StyC, and SpEH, in CFPS reactions. Second, we fused StyA and StyB with a flexible linker to construct a fused SMO (StyA-linker-StyB) and, as a result, the molar ratio of the two enzymes should be 1:1 once they are expressed in CFPS.

By using the above-mentioned two strategies, we found that they did impact the final product titers, albeit their effects were different. For the synthesis of (*S*)-PED, StyA-enriched CFPS system enabled the maximum conversion of styrene to (*S*)-PED, giving rise to the titer of 1.79 ± 0.17 mM (**Figure 6A**), which is > 2 times higher than the control reaction by coexpression of StyA and StyB with plasmids. However, cell-free expression of fused SMO (StyA-linker-StyB) led to the lowest production of (*S*)-PED. The results are in agreement with our previous report that cell extracts pre-enriched with StyA notably enhanced the cascade biotransformation of styrene to (*S*)-PED (Liu et al., 2020). In contrast, both StyA-enriched CFPS and fused SMO benefited the 2-PE synthesis and significantly improved the titers, which are >2 and >4 times higher than that of the control reaction (0.20 ± 0.01 mM) (**Figure 6B**). Interestingly, not like (*S*)-PED, fused SMO was better than the StyA-enriched CFPS to produce 2-PE, perhaps as a result of the different downstream enzymes (SpEH and StyC/ADHs in (*S*)-PED and 2-PE pathways, respectively) that influence the whole cell-free metabolic fluxes.

Since the activity of StyB is dependent on NADH as an electron donor (Otto et al., 2004), we next wanted to supplement NADH to the cell-free systems and explore the effect of NADH concentration on product formation. As shown in **Figures 6C,D**, the titers of both products were obviously increased by supplementation of NADH. Specifically, 1.93 ± 0.08 mM of (*S*)-PED (96.5% conversion based on 2 mM L-Phe) and 1.60 ± 0.09 mM of 2-PE (80% conversion based on 2 mM L-Phe) were produced by adding 2 mM of NADH per reaction. Notably, we also observed product formation without NADH supplementation. This is because our cell-free system, which is based on the *E. coli* crude lysate, contains endogenous NADH that can be used to support the activity of StyB. Additionally, in CFPS reactions 0.33 mM of NAD^+ was added that can also be reduced to NADH by the native NADH/NAD^+ recycling systems (Jewett et al., 2008).

Finally, with the optimized cell-free reaction system, we tested the time courses of product formation by measuring (*S*)-PED and 2-PE concentrations at different time points (**Figures 6E,F**). After mixing two cell-free modules, the cascade biotransformation was carried out for 24 h in total. Our data indicated that the biosynthesis of (*S*)-PED occurred with a gradual increase before 12 h, while 2-PE underwent a nearly linear increase manner. Clearly, more than 95% of both products were synthesized during the initial 16 h reaction and no obvious products were accumulated from 16 to 24 h. Eventually, the conversion rates based on the starting substrate (2 mM L-Phe) at 24 h were 100 and 82.5% for (*S*)-PED and 2-PE, respectively. Since styrene is the common upstream intermediate, the standard Gibbs free energy change of the overall reaction from styrene to (*S*)-PED and 2-PE were calculated to be −363.92 kJ/mol and −224.66 kJ/mol (https://www.chemeo.com/), respectively. This demonstrates that the conversion of styrene to (*S*)-PED is more thermodynamically favorable than that to 2-PE, which is in agreement with the experimental data.

In the present work, we demonstrated the robustness and feasibility of using the *E. coli*-based CFPS system to assemble long biosynthetic pathways enabled by designing modular cell-free systems. While the two paradigms were successful, the overall conversion rates from L-Phe to (*S*)-PED and 2-PE were different. One main reason might be the different levels of the concerted activity of each pathway reconstituted with multiple cell-free

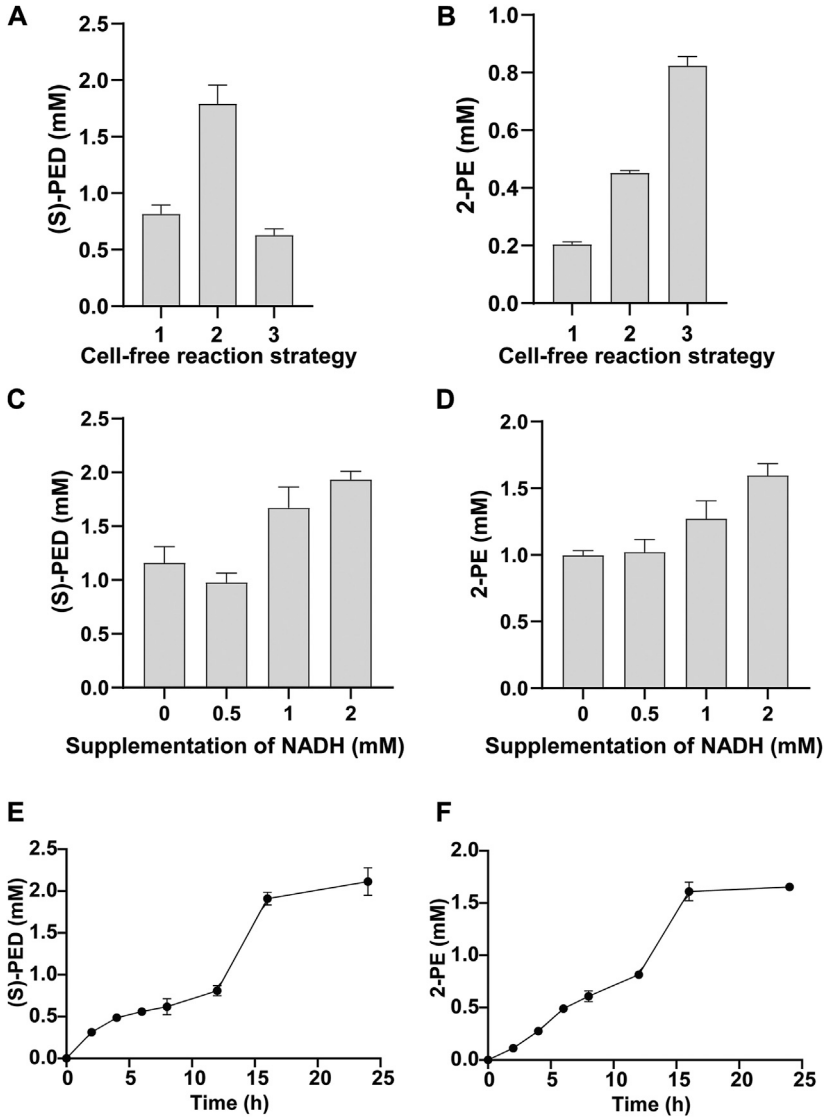


FIGURE 6 | Optimization of StyA and StyB expression for the synthesis of **(A)** (*S*)-PED and **(B)** 2-PE. In panels **(A)** and **(B)**: 1, both StyA and StyB were expressed in CFPS; 2, StyA was expressed *in vivo* to prepare cell extracts and StyA-enriched CFPS was used to express StyB; 3, StyA and StyB was fused and the resulting StyA-linker-StyB was expressed in CFPS. Effect of supplementation of NADH on **(C)** (*S*)-PED and **(D)** 2-PE formation. Time courses of **(E)** (*S*)-PED and **(F)** 2-PE production within 24 h. Values show means with error bars representing standard deviations (s.d.) of at least three independent experiments.

expressed enzymes. For example, the conversion of L-Phe to 2-PE was not complete within 24 h, which is likely due to the low expression of soluble StyC and the resulting low catalytic efficiency. Therefore, given the flexibility of cell-free systems, one might switch from the *E. coli* CFPS system to many other CFPS systems (Kelwick et al., 2016; Li et al., 2017; Des Soye et al., 2018; Wang et al., 2018; Xu et al., 2020; Zhang et al., 2020) to better mimic the physicochemical environment of native hosts for the expression of related enzymes with high (soluble) yields and high activities. In other words, various CFPS systems can be used to express highly active enzymes in different cell-free modules and then mix them for efficient biotranformations, thus giving rise to high productivity.

CONCLUSION

In this study, we demonstrated the design of modular cell-free systems to build *in vitro* long metabolic pathways for tunable conversion of L-Phe to two aromatic compounds (*S*)-PED and 2-PE. The activities of all cell-free expressed enzymes were first confirmed and then mixed modules were able to synthesize (*S*)-PED or 2-PE. After optimization, the final conversion rates based on L-Phe (2 mM) at 24 h were 100% for (*S*)-PED and 82.5% for 2-PE. Looking forward, we envision that modular cell-free systems will provide a feasible approach to construct long enzymatic pathways for the synthesis of valuable chemicals of pharmaceutical and industrial importance.

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

CY, W-QL, and JL designed the experiments. CY and YL performed the experiments. CY and JL analyzed the data and prepared the figures. CW and JL edited the manuscript. JL conceived the project, supervised the research, and wrote the original manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 31971348 and No. 31800720) and the Natural Science Foundation of Shanghai (No. 19ZR1477200). J.L. also acknowledges the starting grant of ShanghaiTech University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.730663/full#supplementary-material

REFERENCES

Bornscheuer, U. T., Huisman, G. W., Kazlauskas, R. J., Lutz, S., Moore, J. C., and Robins, K. (2012). Engineering the Third Wave of Biocatalysis. *Nature* 485, 185–194. doi:10.1038/nature11117

Bowie, J. U., Sherkhanov, S., Korman, T. P., Valliere, M. A., Opgenorth, P. H., and Liu, H. (2020). Synthetic Biochemistry: The Bio-Inspired Cell-free Approach to Commodity Chemical Production. *Trends Biotechnol.* 38, 766–778. doi:10.1016/j.tibtech.2019.12.024

Bundy, B. C., Hunt, J. P., Jewett, M. C., Swartz, J. R., Wood, D. W., Frey, D. D., et al. (2018). Cell-free Biomanufacturing. *Curr. Opin. Chem. Eng.* 22, 177–183. doi:10.1016/j.coche.2018.10.003

Cochrane, F. C., Davin, L. B., and Lewis, N. G. (2004). The *Arabidopsis* Phenylalanine Ammonia Lyase Gene Family: Kinetic Characterization of the Four PAL Isoforms. *Phytochemistry* 65, 1557–1564. doi:10.1016/j.phytochem.2004.05.006

Corrado, M. L., Knaus, T., and Mutti, F. G. (2018). A Chimeric Styrene Monooxygenase with Increased Efficiency in Asymmetric Biocatalytic Epoxidation. *ChemBioChem* 19, 679–686. doi:10.1002/cbic.201700653

Cotton, C. A., Claassens, N. J., Benito-Vaquerizo, S., and Bar-Even, A. (2020). Renewable Methanol and Formate as Microbial Feedstocks. *Curr. Opin. Biotechnol.* 62, 168–180. doi:10.1016/j.copbio.2019.10.002

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 fromVibrio Natriegens. *ACS Synth. Biol.* 7, 2245–2255. doi:10.1021/acssynbio.8b00252

Dudley, Q. M., Karim, A. S., Nash, C. J., and Jewett, M. C. (2020). *In Vitro* prototyping of Limonene Biosynthesis Using Cell-free Protein Synthesis. *Metab. Eng.* 61, 251–260. doi:10.1016/j.ymben.2020.05.006

Feng, J., Yang, C., Zhao, Z., Xu, J., Li, J., and Li, P. (2021). Application of Cell-free Protein Synthesis System for the Biosynthesis of L-Theanine. *ACS Synth. Biol.* 10, 620–631. doi:10.1021/acssynbio.0c00618

Goering, A. W., Li, J., McClure, R. A., Thomson, R. J., Jewett, M. C., and Kelleher, N. L. (2017). *In Vitro* reconstruction of Nonribosomal Peptide Biosynthesis Directly from DNA Using Cell-free Protein Synthesis. *ACS Synth. Biol.* 6, 39–44. doi:10.1021/acssynbio.6b00160

Grubbe, W. S., Rasor, B. J., Krüger, A., Jewett, M. C., and Karim, A. S. (2020). Cell-free Styrene Biosynthesis at High Titers. *Metab. Eng.* 61, 89–95. doi:10.1016/j.ymben.2020.05.009

Henrich, E., Hein, C., Dötsch, V., and Bernhard, F. (2015). Membrane Protein Production inEscherichia Colicell-free Lysates. *FEBS Lett.* 589, 1713–1722. doi:10.1016/j.febslet.2015.04.045

Humphreys, C. M., and Minton, N. P. (2018). Advances in Metabolic Engineering in the Microbial Production of Fuels and Chemicals from C1 Gas. *Curr. Opin. Biotechnol.* 50, 174–181. doi:10.1016/j.copbio.2017.12.023

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. doi:10.1038/msb.2008.57

Karim, A. S., Dudley, Q. M., Juminaga, A., Yuan, Y., Crowe, S. A., Heggestad, J. T., et al. (2020). *In Vitro* prototyping and Rapid Optimization of Biosynthetic Enzymes for Cell Design. *Nat. Chem. Biol.* 16, 912–919. doi:10.1038/s41589-020-0559-0

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. doi:10.1016/j.ymben.2016.03.002

Kelwick, R., Ricci, L., Chee, S. M., Bell, D., Webb, A. J., and Freemont, P. S. (2018). Cell-free Prototyping Strategies for Enhancing the Sustainable Production of Polyhydroxyalkanoates Bioplastics. *Synth. Biol.* 3, ysy016. doi:10.1093/synbio/ysy016

Kelwick, R., Webb, A. J., MacDonald, J. T., and Freemont, P. S. (2016). Development of a *Bacillus Subtilis* Cell-free Transcription-Translation System for Prototyping Regulatory Elements. *Metab. Eng.* 38, 370–381. doi:10.1016/j.ymben.2016.09.008

Ko, Y.-S., Kim, J. W., Lee, J. A., Han, T., Kim, G. B., Park, J. E., et al. (2020). Tools and Strategies of Systems Metabolic Engineering for the Development of Microbial Cell Factories for Chemical Production. *Chem. Soc. Rev.* 49, 4615–4636. doi:10.1039/D0CS00155D

Kwon, Y.-C., Oh, I.-S., Lee, N., Lee, K.-H., Yoon, Y. J., Lee, E. Y., et al. (2013). Integrating Cell-free Biosyntheses of Heme Prosthetic Group and Apoenzyme for the Synthesis of Functional P450 Monooxygenase. *Biotechnol. Bioeng.* 110, 1193–1200. doi:10.1002/bit.24785

Li, J., Lawton, T. J., Kostecki, J. S., Nisthal, A., Fang, J., Mayo, S. L., et al. (2016). Cell-free Protein Synthesis Enables High Yielding Synthesis of an Active Multicopper Oxidase. *Biotechnol. J.* 11, 212–218. doi:10.1002/biot.201500030

Li, J., and Neubauer, P. (2014). *Escherichia coli* as a Cell Factory for Heterologous Production of Nonribosomal Peptides and Polyketides. *New Biotechnol.* 31, 579–585. doi:10.1016/j.nbt.2014.03.006

Li, J., Wang, H., Kwon, Y.-C., and Jewett, M. C. (2017). Establishing a High Yieldingstreptomyces-Based Cell-free Protein Synthesis System. *Biotechnol. Bioeng.* 114, 1343–1353. doi:10.1002/bit.26253

Li, J., Zhang, L., and Liu, W. (2018). Cell-free Synthetic Biology for *In Vitro* Biosynthesis of Pharmaceutical Natural Products. *Synth. Syst. Biotechnol.* 3, 83–89. doi:10.1016/j.synbio.2018.02.002

Liu, W.-Q., Zhang, L., Chen, M., and Li, J. (2019). Cell-free Protein Synthesis: Recent Advances in Bacterial Extract Sources and Expanded Applications. *Biochem. Eng. J.* 141, 182–189. doi:10.1016/j.bej.2018.10.023

Liu, W. Q., Wu, C., Jewett, M. C., and Li, J. (2020). Cell-free Protein Synthesis Enables One-pot cascade Biotransformation in an Aqueous-organic Biphasic System. *Biotechnol. Bioeng.* 117, 4001–4008. doi:10.1002/bit.27541

Lukito, B. R., Wu, S., Saw, H. J. J., and Li, Z. (2019). One-Pot Production of Natural 2-Phenylethanol fromL-Phenylalanine via Cascade Biotransformations. *ChemCatChem* 11, 831–840. doi:10.1002/cctc.201801613

Martin, R. W., Des Soye, B. J., Kwon, Y.-C., Kay, J., Davis, R. G., Thomas, P. M., et al. (2018). Cell-free Protein Synthesis from Genomically Recoded Bacteria Enables Multisite Incorporation of Noncanonical Amino Acids. *Nat. Commun.* 9, 1203. doi:10.1038/s41467-018-03469-5

Moore, S. J., Lai, H.-E., Chee, S.-M., Toh, M., Coode, S., Chengan, K., et al. (2021). A *Streptomyces Venezuelae* Cell-free Toolkit for Synthetic Biology. *ACS Synth. Biol.* 10, 402–411. doi:10.1021/acssynbio.0c00581

Nielsen, J., and Keasling, J. D. (2016). Engineering Cellular Metabolism. *Cell* 164, 1185–1197. doi:10.1016/j.cell.2016.02.004

Otto, K., Hofstetter, K., Röthlisberger, M., Witholt, B., and Schmid, A. (2004). Biochemical Characterization of StyAB from *Pseudomonas* Sp. Strain VLB120 as a Two-Component Flavin-Diffusible Monooxygenase. *J. Bacteriol.* 186, 5292–5302. doi:10.1128/JB.186.16.5292-5302.2004

Panke, S., Witholt, B., Schmid, A., and Wubbolts, M. G. (1998). Towards a Biocatalyst for (S)-Styrene Oxide Production: Characterization of the Styrene Degradation Pathway of Pseudomonas Sp. Strain VLB120. *Appl. Environ. Microbiol.* 64, 2032–2043. doi:10.1128/AEM.64.6.2032-2043.1998

Payne, K. A. P., White, M. D., Fisher, K., Khara, B., Bailey, S. S., Parker, D., et al. (2015). New Cofactor Supports α,β-unsaturated Acid Decarboxylation via 1,3-dipolar Cycloaddition. *Nature* 522, 497–501. doi:10.1038/nature14560

Peng, F., Ou, X. Y., Zhao, Y., Zong, M. H., and Lou, W. Y. (2019). Highly Selective Resolution of Racemic 1-phenyl-1,2-ethanediol by a Novel Strain Kurthia Gibsonii SC 0312. *Lett. Appl. Microbiol.* 68, 446–454. doi:10.1111/lam.13123

Rasor, B. J., Vögeli, B., Landwehr, G. M., Bogart, J. W., Karim, A. S., and Jewett, M. C. (2021). Toward Sustainable, Cell-free Biomanufacturing. *Curr. Opin. Biotechnol.* 69, 136–144. doi:10.1016/j.copbio.2020.12.012

Sheldon, R. A., and Woodley, J. M. (2018). Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* 118, 801–838. doi:10.1021/acs.chemrev.7b00203

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. doi:10.1038/s41576-019-0186-3

Swartz, J. R. (2018). Expanding Biological Applications Using Cell-free Metabolic Engineering: An Overview. *Metab. Eng.* 50, 156–172. doi:10.1016/j.ymben.2018.09.011

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

Wang, Y., Zhang, H., Lu, X., Zong, H., and Zhuge, B. (2019). Advances in 2-phenylethanol Production from Engineered Microorganisms. *Biotechnol. Adv.* 37, 403–409. doi:10.1016/j.biotechadv.2019.02.005

Wang, Z., Sundara Sekar, B., and Li, Z. (2021). Recent Advances in Artificial Enzyme Cascades for the Production of Value-Added Chemicals. *Bioresour. Tech.* 323, 124551–551. doi:10.1016/j.biortech.2020.124551

Wilding, K. M., Hunt, J. P., Wilkerson, J. W., Funk, P. J., Swensen, R. L., Carver, W. C., et al. (2019). Endotoxin-Free E. Coli- Based Cell-free Protein Synthesis: Pre-expression Endotoxin Removal Approaches for On-Demand Cancer Therapeutic Production. *Biotechnol. J.* 14, 1800271. doi:10.1002/biot.201800271

Wu, G., Yan, Q., Jones, J. A., Tang, Y. J., Fong, S. S., and Koffas, M. A. G. (2016). Metabolic burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol.* 34, 652–664. doi:10.1016/j.tibtech.2016.02.010

Wu, S., Li, A., Chin, Y. S., and Li, Z. (2013). Enantioselective Hydrolysis of Racemic and Meso-Epoxides with Recombinant *Escherichia coli* Expressing Epoxide Hydrolase from *Sphingomonas* Sp. HXN-200: Preparation of Epoxides and Vicinal Diols in High *Ee* and High Concentration. *ACS Catal.* 3, 752–759. doi:10.1021/cs300804v

Wu, S., and Li, Z. (2018). Whole-cell cascade Biotransformations for One-Pot Multistep Organic Synthesis. *ChemCatChem* 10, 2164–2178. doi:10.1002/cctc.201701669

Wu, S., Snajdrova, R., Moore, J. C., Baldenius, K., and Bornscheuer, U. T. (2021). Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angew. Chem. Int. Ed.* 60, 88–119. doi:10.1002/ange.202006648

Xu, H., Liu, W.-Q., and Li, J. (2020). Translation Related Factors Improve the Productivity of a *Streptomyces*-Based Cell-free Protein Synthesis System. *ACS Synth. Biol.* 9, 1221–1224. doi:10.1021/acssynbio.0c00140

Zhang, L., Liu, W.-Q., and Li, J. (2020). Establishing a Eukaryotic *Pichia pastoris* Cell-free Protein Synthesis System. *Front. Bioeng. Biotechnol.* 8, 536. doi:10.3389/fbioe.2020.00536

Zhuang, L., Huang, S., Liu, W.-Q., Karim, A. S., Jewett, M. C., and Li, J. (2020). Total *In Vitro* Biosynthesis of the Nonribosomal Macrolactone Peptide Valinomycin. *Metab. Eng.* 60, 37–44. doi:10.1016/j.ymben.2020.03.009

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

ORIGINAL RESEARCH
published: 23 August 2021
doi: 10.3389/fbioe.2021.727584

Decentralizing Cell-Free RNA Sensing With the Use of Low-Cost Cell Extracts

*Anibal Arce[1,2,3], Fernando Guzman Chavez[4], Chiara Gandini[5], Juan Puig[1,2], Tamara Matute[1,2], Jim Haseloff[4], Neil Dalchau[6], Jenny Molloy[5], Keith Pardee[7] and Fernán Federici[1,2,3,8]**

[1]ANID – Millennium Science Initiative Program – Millennium Institute for Integrative Biology (iBio), Santiago, Chile, [2]Schools of Engineering, Institute for Biological and Medical Engineering, Medicine and Biological Sciences, Pontificia Universidad Católica de Chile, Santiago, Chile, [3]Departamento de Genética Molecular y Microbiología, Pontificia Universidad Católica de Chile, Santiago, Chile, [4]Department of Plant Sciences, University of Cambridge, Cambridge, United Kingdom, [5]Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, United Kingdom, [6]Microsoft Research, Cambridge, United Kingdom, [7]Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada, [8]FONDAP Center for Genome Regulation, Santiago, Chile

Cell-free gene expression systems have emerged as a promising platform for field-deployed biosensing and diagnostics. When combined with programmable toehold switch-based RNA sensors, these systems can be used to detect arbitrary RNAs and freeze-dried for room temperature transport to the point-of-need. These sensors, however, have been mainly implemented using reconstituted PURE cell-free protein expression systems that are difficult to source in the Global South due to their high commercial cost and cold-chain shipping requirements. Based on preliminary demonstrations of toehold sensors working on lysates, we describe the fast prototyping of RNA toehold switch-based sensors that can be produced locally and reduce the cost of sensors by two orders of magnitude. We demonstrate that these in-house cell lysates provide sensor performance comparable to commercial PURE cell-free systems. We further optimize these lysates with a CRISPRi strategy to enhance the stability of linear DNAs by knocking-down genes responsible for linear DNA degradation. This enables the direct use of PCR products for fast screening of new designs. As a proof-of-concept, we develop novel toehold sensors for the plant pathogen Potato Virus Y (PVY), which dramatically reduces the yield of this important staple crop. The local implementation of low-cost cell-free toehold sensors could enable biosensing capacity at the regional level and lead to more decentralized models for global surveillance of infectious disease.

Keywords: cell-free, toehold-sensor, low-cost, diagnostics, decentralization

OPEN ACCESS

Edited by:
Simon J. Moore,
University of Kent, United Kingdom

Reviewed by:
Nadanai Laohakunakorn,
University of Edinburgh,
United Kingdom
Amir Pandi,
Max Planck Institute for Terrestrial Microbiology, Germany

***Correspondence:**
Fernán Federici
ffederici@bio.puc.cl

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

Received: *18 June 2021*
Accepted: *06 August 2021*
Published: *23 August 2021*

Citation:
Arce A, Guzman Chavez F, Gandini C, Puig J, Matute T, Haseloff J, Dalchau N, Molloy J, Pardee K and Federici F (2021) Decentralizing Cell-Free RNA Sensing With the Use of Low-Cost Cell Extracts. Front. Bioeng. Biotechnol. 9:727584. doi: 10.3389/fbioe.2021.727584

INTRODUCTION

Cell-free sensors have emerged as a promising technology for fast and field-deployable detection of nucleic acids (Pardee et al., 2014, 2016; Ma et al., 2018; Takahashi et al., 2018; Verosloff et al., 2019) and chemical pollutants (Didovyk et al., 2017; Jung et al., 2020; Liu et al., 2020; Silverman et al., 2020; Thavarajah et al., 2020). Cell-free toehold RNA sensors (CFTS) have been shown to detect viruses such as Ebola (Pardee et al., 2014), norovirus (Ma et al., 2018), and Zika (ZIKV) (Pardee et al., 2016) at femtomolar concentrations of viral RNA when combined with isothermal RNA amplification (Pardee et al., 2016). Toehold RNA sensors are *de novo* designed genetic structures that can be programmed to detect RNAs of arbitrary sequences (Green et al., 2014). A toehold switch relies on

the formation of a hairpin loop at the 5′-end of a reporter gene that structurally prohibits an interaction between the ribosome binding site (RBS) and ribosomes, thereby preventing translation. This structure is released *via* toehold-mediated strand displacement by an unstructured RNA "trigger" with sequence-specific complementarity to the switch (Green et al., 2014), thus enabling ribosomal-binding and subsequent protein translation. These CFTS can be lyophilized and transported at room temperature for in-field detection of viral RNAs (Pardee et al., 2014, 2016; Ma et al., 2018).

CFTS are implemented in a reconstituted cell-free expression system, known commercially as PURExpress, which consists of a mixture of 34 individually purified components involved in transcription, translation and energy generation (Shimizu et al., 2001). Despite the potential of CFTS for point-of-care deployment, constraints in the accessibility to PURE systems in South America impose significant barriers to locally engineer *de novo* sensors. PURE systems are expensive and proprietary (Laohakunakorn et al., 2020). Furthermore, they need to be imported from the Northern hemisphere under cold-chain shipping, which increases their costs and the risk of activity loss. PURE systems can also be prepared in the laboratory by purifying the components individually or as a single step from a bacterial consortium of strains (Shimizu et al., 2001; Horiya et al., 2017; Villarreal et al., 2018; Lavickova and Maerkl, 2019). Alternatively, cell lysate-based protein expression systems, made from E. coli crude extract that do not require purification of particular components, can also be used as a substrate for CFTS (Silverman et al., 2019). Lehr et al. have shown the implementation of CFTS in commercially available cell-free transcription-translation systems (TX-TL) based on cell extracts (Lehr et al., 2019), which are also sourced from the US and shipped under cold chain. Encouragingly, Silverman et al. (2019), McNerney et al. (2019) and Pieters et al. (2021) have recently shown the successful implementation of CFTS in home-made cell extracts, offering a path for local prototyping *de novo* sensors of interest (McNerney et al., 2019; Silverman et al., 2019; Pieters et al., 2021).

Cell extracts are cell lysate-based protein expression systems that have been instrumental to molecular biology and biotechnology for decades (Nirenberg and Matthaei, 1961). Cell extracts have also become a versatile tool for engineering complex biomolecular systems outside living cells (Garenne and Noireaux, 2019; Laohakunakorn et al., 2020). Diverse genetic systems have been implemented in cell extracts, e.g. for sensing small molecules (Salehi et al., 2018; Jung et al., 2020; Silverman et al., 2020; Thavarajah et al., 2020), engineering metabolic pathways (Dudley et al., 2016; Grubbe et al., 2020; Swartz, 2020), producing virus-like particles (Bundy et al., 2008; Rustad et al., 2018), and prototyping genetic devices (Siegal-Gaskins et al., 2014; Moore et al., 2017), including RNA circuits (Takahashi et al., 2015), transcriptional cascades (Shin and Noireaux, 2012; Garamella et al., 2016) and oscillators (Niederholtmeyer et al., 2015; Tayar et al., 2017; Yelleswarapu et al., 2018). Various protocols have been crafted to simplify extract preparation and optimize lysate performance, including methods based on cell homogenization (Kim et al., 2006), sonication (Shrestha et al., 2012; Kwon and Jewett, 2015), bead-beating (Sun et al., 2013), enzymatic lysis or flash freezing (Didovyk et al., 2017).

With the goal of making cell-free toehold sensor technology more accessible and affordable, here, we describe the implementation of CFTS in cell extracts supplemented with maltodextrin as an energy source. These CFTS can detect RNA at similar sensitivity to the PURExpress-based CFTS but for a fraction of the cost. Furthermore, using a CRISPRi strategy for silencing nucleases before harvesting cells, we increased DNA stability, enabling cell-free protein synthesis from linear, PCR-derived templates. These optimized extracts were suitable for fast prototyping of novel toehold sensors directly from linear PCR products using a lacZ gene template. As a proof-of-concept demonstration, we applied our framework to engineer cost-effective CFTS for PVY virus detection, which poses a significant economic threat to local potato farmers and producers of potato seeds.

MATERIALS AND METHODS

Crude Extract Preparation

BL21 DE3 Star cells (Invitrogen, C601003), BL21dLacGold cells (Didovyk et al., 2017) (Agilent, 230,132), or the CRISPR+, CRISPR-derivatives from this work were grown overnight at 37°C on 2xYT agar plates with the appropriate antibiotics. A single colony was grown in a 5 ml 2xYT culture overnight at 37°C. The next day, a 50 ml culture was started with a 1:1,000 dilution overnight at 37°C. On the fourth day, a total of 2 L of culture was started and split in 5 2-L Erlenmeyer, at an initial OD_{600} 0.05 in 2xYT supplemented with 40 mM phosphate dibasic (Sigma, 94,046), 22 mM phosphate monobasic (Sigma, 74,092) and 20 g/L D-glucose. This culture was grown at 37°C with vigorous agitation (230 rpm) until OD_{600} reached 0.6–0.8 (approximately 2.5 h) and cells were induced with 1 mM IPTG for another 3 h at 37°C before harvesting by centrifugation at 5,000 *g* and 4°C. Pellets were washed twice with cold S30B buffer (5 mM tris acetate pH 8.2, 14 mM magnesium acetate, 60 mM potassium glutamate, and 2 mM dithiothreitol) (Sun et al., 2013). The pellet was then weighted, and a relationship of 0.9 ml of S30B and 5 g of 0.1 mm diameter silica beads (Biospec, 11079101) were added per Gram of pellet obtained. 2.5 ml bead-beating tubes and cups were filled with the viscous solution composed of cells and glass beads, without generating bubbles, sealed and beaten for 30 s using a homogenizer (MP Biomedicals, FastPrep-24 5G). To remove the beads from extracts, processed samples were placed on the top of a micro-chromatography column (Biorad, 7326204), which was placed into an empty bead-beating tube. The bead-beaten tube attached to the micro-chromatography column and the empty recipient tube was placed into a 15 ml Falcon tube and centrifuged at 6,000 g for 5 min at 4°C. Properly beat extracts should appear clear, and two distinct layers should be observed. The supernatant from all tubes was pooled and agitated at 37°C inside a 5 ml unsealed tube for 1 hour for runoff. After the run-off reaction, the tubes were placed on ice and re-centrifuged at 6,000 g for 10 min at 4°C to pull down the debris

generated during the run-off reaction. The supernatant from this centrifugation, i.e., the crude extract, was aliquoted, frozen, and stored at −80°C until use. A link to the full protocol of crude extract and cell-free reaction preparation can be found here: https://www.protocols.io/view/preparation-of-cell-free-rnapt7-reactions-kz2cx8e.

Cell-Free Transcription-Translation Reaction

The in-house cell-free reactions were composed of 45% crude lysate and 40% reaction buffer. The remaining 15% included DNA input solution chlorophenol red-β-D-galactopyranoside (Sigma, 59,767 at final concentration of 0.6 mg/ml), and in the case of the RNA sensing reactions, trigger RNA from *in vitro* transcription. A typical reaction is composed of 50 mM HEPES pH 8, 1.5 mM ATP and GTP, 1.4 mM CTP and UTP as triphosphate ribonucleotides, 0.2 mg/ml tRNA (Roche, 01109541001), 0.26 mM Coenzyme A, 0.33 mM NAD, 0.756 mM cAMP, 0.0068 mM folinic acid, 0.112 mg/ml spermidine, 2% PEG-8000, 3.4 mM of each of the 20 amino acids (glutamate is also added in excess in potassium and magnesium salts), and 12 mg/ml maltodextrin and 0.6 mg/ml sodium hexametaphosphate (Sigma, 305,553) or 30 mM 3-PGA as an energy source. Cell-free reactions were mounted on ice in a final volume of 10 μL in 1.5 ml Eppendorf tubes and vortexed for 30 s before sampling 5 μL that were loaded in V-bottom 96 well plates (Corning, CLS3957) or rounded 384 well-plates (Corning, CLS3540). All the reactions were incubated at 29°C in a Synergy HTX plate reader (BioTek).

Input DNA Preparation

All plasmids used in this work correspond to single transcriptional units (Level 1 plasmids) that were prepared using Golden Gate assembly from Level 0 parts (T7 Promoter, T7 terminator, Toehold Sensors, Triggers, etc.) (Pollak et al., 2019). The level 0 parts were previously prepared by Gibson assembly (Gibson et al., 2009) from PCR linear DNA amplified from gblocks (IDT). All plasmids were commercially sequenced before use. Plasmid DNA input was produced by midi prepping an overnight culture of 200 ml LB with the appropriate strain (Promega, A2492) and cleaned again using PCR cleanup (Promega, A6754). Plasmid DNA inputs were used in a set of concentrations ranging from 0.5 to 10 nM. Input linear DNA was produced by PCR and cleaned with PCR cleanup kit (Promega, A6754), quantified and diluted to desired concentrations using ultra-pure water. A list of all the primers and plasmids used in this work is shown in **Supplementary Tables S1, 2**, respectively.

Trigger RNA Preparation

PCR products containing the trigger T7 transcriptional unit were cleaned by PCR cleanup kit (Promega, A6754) and 250 ng were processed by *in vitro* transcription using the HiScribe T7 kit (Promega, E2040S) and incubated at 37°C overnight. The products were then treated with DNAseI for 1 hour at 37°C and the enzyme was inactivated by heat at 70°C for 10 min. Next, standard RNA clean up steps were performed according to the RNeasy MinElute kit (Qiagen, 74,204) before quantification.

Isothermal NASBA Amplification

We used the NASBA lyophilized kit from Life Sciences following the manufacturer's instructions. Negative controls and RNA dilutions in the range 27 mM to 0.2 fM were prepared in ultra-pure water just before use. The lyophilized reaction buffer containing nucleotide mix (Life Sciences, LRB-10) was reconstituted with DRB-10 diluent (Life Sciences, DRB-10), heated at 41°C for 5 min and kept at room temperature. The initial mixture, consisting of 10 μL of the reconstituted reaction buffer, 1 μL RNAsin (Promega, N2111), 380 mM of each DNA primer, 2 μL ultra-pure water and a 1 μL RNA amplicon dilution, were assembled at 4°C and incubated at 65°C for 2 min, followed by a 10 min annealing at 41°C. During annealing, the lyophilized enzyme mix, consistent of AMV RT, RNaseH, T7 RNAP, and high molecular weight sugar mix (Life Sciences, LEM-10), was reconstituted with cold D-PDG (Life Sciences, D-PDG-10) and placed on ice. Immediately after annealing, 5 μL of the dissolved enzyme mix was added to the reaction, to a final volume of 20 μL, and the mixture was incubated at 41°C for 2 h with thermocycler lid at 98°C. A 10 min incubation at 70°C was performed for enzyme inactivation. The amplified product was subsequently cleaned up with the RNeasy MinElute kit (Qiagen, 74,204).

Lyophilization and Storage of Cell-free Reactions

Lyophilization was performed by assembling 9 μL of cell extracts and 8 μL of reaction buffer in a PCR tube. In a separate tube, 3 μL of the corresponding plasmid were mixed with 1 μL of substrate (9 mg/ml), when LacZ reporter was used. Tubes containing the assembled reactions were closed with adhesive aluminum film and punctured with a 16-gauge needle to create a hole. The tubes were placed in a FreeZone 2.5 L Triad Benchtop Freeze Dryer (Labconco), previously pre-chilled at −75°C (shelf temperature), for overnight freeze-drying of two segments. The first segment was at −75°C (shelf temperature) with a temperature collector of −80°C for 12 h and 0.04 mbar of pressure. The second segment was performed at −20°C for 4 h with the same pressure. Samples were stored at room temperature in open plastic bags over 30 g of silica inside a closed Tupperware. Rehydration of cell-free reactions was done with 20 μL of a plasmid solution in ultra-pure water. pT7:sfGFP was used at 6 nM or 8 nM final concentration whereas ZIKV Toehold 27 was used at 10 nM final concentration. Trigger 27 RNA was added at a final concentration of 300 nM.

Design of Novel Toehold Sensors

With the aim of designing novel toehold sensors, we implemented an algorithm developed by Green et al. (Green et al., 2014), as an in silico tool called NupackSensors (http://www.github.com/elanibal/NupackSensors). To identify possible triggers from a target sequence, the algorithm first determines all contiguous 36 nucleotide sub-sequences. Each possible trigger defines one specific toehold sensor. After filtering out the sensors that contained stop

codons, potential structural defects of each design were calculated, along with a set of thermodynamic parameters (**Supplementary Figure S1**; **Supplementary Tables S3, 4**). The following score function was implemented, and one score value was assigned to each of the possible designs (Ma et al., 2018):

$$Score = -71d_{Sensor} - 49.1d_{ActiveSensor} - 22.6d_{BindingSite} + 54.3$$

Here, d_{Sensor}represents the average number of incorrectly paired nucleotides with respect to the ideal toehold sensor structure, defined as ".....((((((((((...(((((.....)))))...)))))))))).....” (Pardee et al., 2016) in dot-bracket notation (**Supplementary Figure S1**). Similarly, the parameter$d_{ActiveSensor}$represents the average number of incorrectly paired nucleotides with respect to the ideal activated sensor structure.$d_{BindingSite}$represents the deviations from the perfect single-stranded binding site in the toehold domain, which is essential for the initial trigger binding and subsequent strand-displacement reaction. After assigning one score to each toehold sensor design, they are ranked.

For the PVY toehold sensor design, two RNA target sequences were evaluated in NupackSensors, corresponding to the long and short sequences found to be conserved in several PVY strains (**Supplementary Tables S5, 6**). For each of them, four of the highest scores were picked for experimental evaluation. Linear DNA encoding the PVY toeholds were prepared by PCR using specific primers that add the T7 promoter and each toehold sensor sequence on their 5' tails along with the UNSXR reverse primer. LacZ full length was used as template (**Supplementary Table S1** for primer sequences).

sgRNA Design

Sequences of each operon of the genes that are part of the RecBCD complex and the EndI endonuclease were identified in E.coli K12 strain genome using the web platform RegulonDB (Salgado et al., 2013). sgRNAs were designed according to (Larson et al., 2013), following these steps: all available 5'->3' PAM sites (CCN) within the nucleotide sequence of the gene of interest were marked, identifying contiguous sites as target candidates; the target candidates were analyzed and selected as the 18–22 bp contiguous to the PAM sequence that are within the first ~156 nt of the CDS (i.e., counting from the ATG); sequences that end in C or T were selected so that the sgRNA begins with A or G (corresponding to nucleotide +1 of the sgRNA transcript) since this increases the transcription of the promoter used (Larson et al., 2013; Qi et al., 2013). The designed sgRNA corresponded to the reverse complementary sequences of the chosen target. These steps were performed in ApE open source software (https://jorgensen.biology.utah.edu/wayned/ape/). We designed three sgRNAs that targeted the cistron of recB and recD, recC and endA genes, respectively.

Stability of PCR-Derived Linear DNA in Cell-free Reactions

Linear DNA fragments of 1.5 kb and 500 bp were prepared *via* PCR for its use as linear DNA input and standards, respectively. Those were purified by wizard gel and PCR clean-up kit (Promega), achieving a final concentration of 200 ng/μL 600 ng of the 1.5 kb linear DNA were added to 20 μL cell-free reactions and incubated at 29°C for 0, 30, 60 or 120 min and quickly submerged in liquid nitrogen to stop the reaction. Frozen samples were put on ice; and immediately, a volume of binding buffer from PCR clean-up kit (Promega, A9282) was added followed by 2 μL (300 ng) of the 500 bp linear DNA that served as an internal standard for the purification. The samples were then purified following standard supplier's protocol and eluted with 30 μL of ultra-pure water. Next, 0.3 volumes of loading buffer were added to the purified samples and 10 μL were run on a 1% agarose gel electrophoresis containing 1X SYBR-Safe for DNA visualization. 1.5 kb and 500 bp DNA fragments were placed in two independent lines along with another line for 1 kb Ladder (NEB, N232). Once bands were clearly identifiable, the gel was photographed on a blue LED light transilluminator with a blue-light filter. The intensity of the 1.5 kb bands in each sample was integrated using ImageJ and normalized with the intensity of the 500 bp band of the same lane.

RESULTS

Implementation of an In-House Cell-free Gene Expression System

We first sought to implement a low-cost and simple protocol for cell-free gene expression. To do this, we combined the use of S12 crude extracts–a simple post lysate processing which does not require ultracentrifugation nor dialysis (Kim et al., 2008)—with a highly concentrated amino acid mixture (Caschera and Noireaux, 2015b) and a cost-effective energy source. This energy source is based on maltodextrin and polyphosphates, which do not require cold-chain transportation (Caschera and Noireaux, 2015a). We characterized our system by measuring the dynamics or endpoints of the pT7:sfGFP constitutive expression (**Figure 1A**). Initially, we prepared four different batches and tested optimum magnesium concentration in the reaction buffer for the cell-free expression of sfGFP. Optimal values ranged between 8.5 and 10 nM for all extracts (**Supplementary Figure S2**). With the goal of reducing the cost of lysate preparation, we compared maltodextrin and polyphosphates (HMP) to the conventional and more expensive 3-phosphoglyceric acid (3-PGA) energy source. In a side-by-side comparison with 3-PGA, maltodextrin produced a similar initial rate in sfGFP fluorescence (**Supplementary Figure S3**). Moreover, the endpoint measurements were significantly higher for the cost-effective maltodextrin-based energy source in each batch tested (**Figure 1B**). Next, we measured the dynamics of constitutive pT7:sfGFP (**Figure 1A**) expression in these extracts compared to two RNAP T7-based commercial kits. In comparison to these commercial cell-free systems, our in-house cell lysates generated higher endpoint fluorescent values, although variability was observed between different batches of cell-free extracts (**Figure 1C**).

A cost-breakdown analysis of lab-scale extract production of this optimized cell-free lysate system was performed, comparing

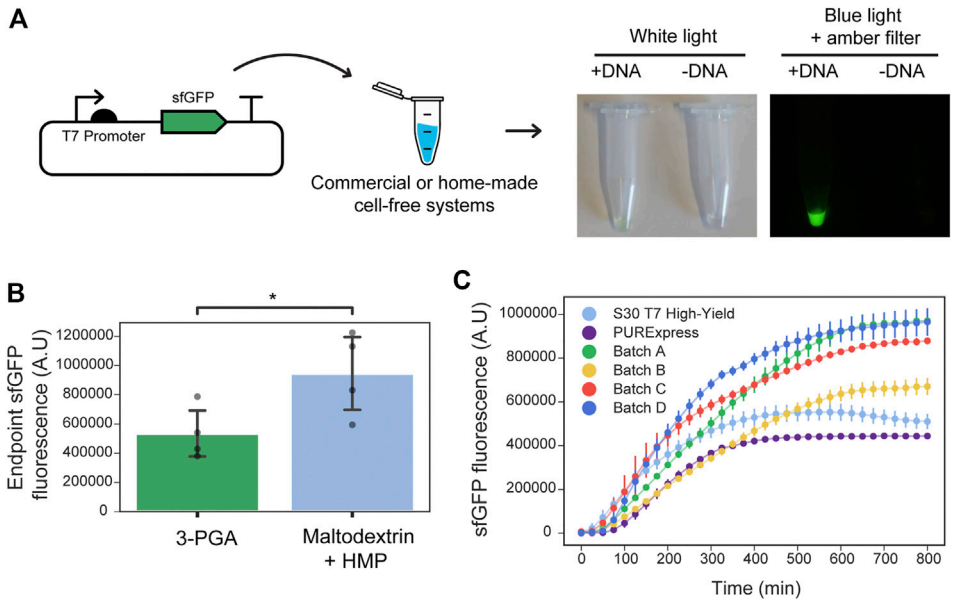


FIGURE 1 | Optimized in-house cell-free reactions compared to commercial alternatives. **(A)** Left: Schematic representation of testing the performance of home-made or commercial cell-free systems using the sfGFP constitutive reporter. Right: Example of the endpoint sfGFP reaction and negative control (without input DNA) in a home-made cell-free system supplemented with maltodextrin energy source. Tubes were photographed under white light or blue light plus an amber filter that allows visualizing the sfGFP fluorescence. **(B)** Endpoint sfGFP fluorescence (plasmid DNA at 9 nM final concentration) was measured in four different cell extracts (Batch A, B, C, D) supplemented with either maltodextrin and polyphosphates (light blue) or 3-PGA (green) as energy source. Grey dots represent the arithmetic mean of three measurements performed on each batch, and error bars represent standard deviations of the means of the four batches tested (N = 4). *t*-test for paired measurements was performed and statistically significance was found between the two groups (*p*-value = 0.03, shown by *). Assumptions of the paired *t*-test were verified using the Shapiro-Wilk test for normality of the differences between energy sources for a given batch (*p*-value = 0.97), and Levene test for homoscedasticity of the 3-PGA and maltodextrin data sets (*p*-value = 0.25). **(C)** sfGFP production dynamics from plasmid DNA (9 nM final concentration) in reactions performed at 29°C using NEB PURExpress and Promega S30 T7 High Yield commercial kits along with four optimized in-house cell-free reactions (Batch A, B, C, D) using maltodextrin and polyphosphates as the energy source. Error bars represent the standard deviations of three independent replicates, dots are centered at the arithmetic mean for each time point.

reagent prices available to the collaborating institutions in Chile and the United Kingdom (**Supplementary Tables S7–13**). The cost per reaction (5 µL) in Chile was $0.069 USD for the maltodextrin-based preparation in comparison to $7.80 USD for the locally purchased commercial PURExpress, representing a reduction of two orders of magnitude. In comparison, the home-made preparations were about two times less expensive in the United Kingdom than in Chile (**Supplementary Figure S4**; **Supplementary Table S7**). Maltodextrin-based reactions were approximately 20% cheaper than the 3-PGA alternative.

Implementation of CFTS in Low-Cost Cell-free Reactions

To investigate whether RNA toehold sensors would be functional in the in-house prepared cell extracts, we used Zika virus (ZIKV) toehold sensors already shown to work well in PURExpress (Pardee et al., 2016). We selected ZIKV toehold sensors 8 and 27 and their respective triggers due to their high dynamic range and orthogonality (Pardee et al., 2016). These toehold sensors, controlling full-length LacZ expression, were incubated with the corresponding RNA trigger (400 nM) independently (**Figure 2A**). For each sensor, we chose the highest DNA concentration that ensures low-background in PURExpress and used this concentration for comparison with in-house cell-free reactions. Sensor 8 (at 0.8 nM plasmid DNA) behaved similarly in both reaction conditions, achieving a near-maximal response about 120 min after induction while maintaining a low signal in the absence of the trigger (**Figure 2B**). Sensor 27 (2 nM input DNA), also displayed comparable performance and sequence-specific activation in both systems used (**Figures 2B,C**). We also compared the performance of the full-length LacZ with LacZ-alpha reporters supplemented with the previously expressed LacZ-omega peptide in commercial and in-house cell-free extracts (**Supplementary Figure S5**). Full-length LacZ reporters exhibited higher ON/OFF endpoint absorbance values in both systems tested (**Supplementary Figure S6**, right). This can be explained by the slightly higher OFF state observed on both toehold sensors when using LacZ-alpha as a reporter (**Supplementary Figure S6**).

Next, we studied the effect of plasmid DNA input concentration on the leaky expression and saturation of sensing reactions. We tested three ZIKV toehold sensors controlling full-length LacZ expression (ZIKV Sensors 8, 27, and 32) (Pardee et al., 2016) in concentrations ranging from

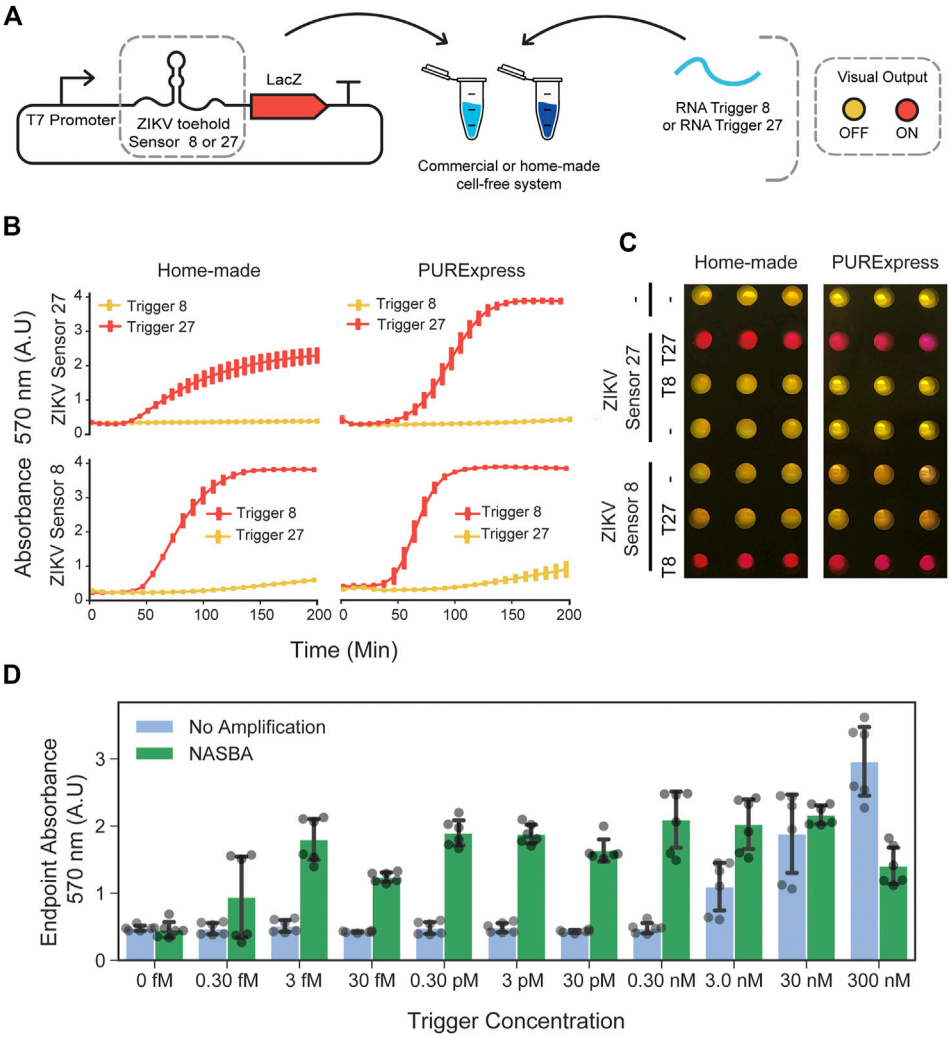


FIGURE 2 | Performance of ZIKV toehold sensors in low-cost cell-free lysate reactions. **(A)** Schematic representation of toehold-mediated RNA sensing. **(B)** Dynamics of the RNA sensing reactions performed with ZIKV toehold sensor 8 (0.7 nM plasmid DNA) and 27 (2 nM plasmid DNA), regulating the expression of the full-length LacZ in home-made cell extracts and PURExpress cell-free reactions. Error bars represent the standard deviations of three independent experiments, dots are centered at the arithmetic mean for each time point. **(C)** Example of the endpoint visualization of the experiments after 4 hours of incubation at 29°C. **(D)** Endpoint measurement of RNA sensing reactions performed with ZIKV sensor 27 and trigger 27 in a range of concentrations with and without NASBA isothermal amplification. Gray dots represent data from six independent measurements performed from two independent NASBA amplifications performed on different days. Black error bars correspond to standard deviations of these six measurements.

0.5 to 10 nM. In all sensors, OFF state signal increased as a function of the input DNA concentration (**Supplementary Figure S7**). These results indicate that input DNA concentration can affect the dynamic range of RNA toehold sensors and that optimal concentrations may be different for each sensor.

To increase the sensitivity of CFTS in home-made cell-free preparations, we performed NASBA amplification of the target RNA before adding it to the reactions containing the sensors (Pardee et al., 2016). While unamplified RNA was detected at concentrations no lower than 3 nM, the NASBA amplification of target RNAs upstream in the workflow permitted the detection of RNA at concentrations as low as three fM (**Figure 2D**), representing a sensitivity increment of six orders of magnitude.

Lyophilization and Shelf-Stability of In-House Cell-free Reactions

Deployment of RNA sensors in point-of-care settings is facilitated by the lyophilization of cell-free reactions, enabling room temperature transportation and use upon rehydration (Pardee et al., 2016). In order to test whether home-made CFTS can also be freeze-dried and stored at room-temperature, we lyophilized reactions supplemented with pT7:sfGFP plasmid. These reactions were sealed immediately after lyophilization and kept at room

temperature for up to 90 days in a closed tupperware containing desiccant and, upon rehydration, compared to a fresh sample for reference. We observed more than 60% recovery of the endpoint sfGFP expression 1 week after lyophilization, which decreased to 17% after 90 days at room temperature (**Supplementary Figure S8**). The ZIKV Sensor 27 was also tested in home-made cell extracts after one and 7 days of lyophilization. Besides the evident decrease in dynamic range, especially after 7 days, these reactions retained the ability to detect its target trigger, while maintaining target specificity (**Supplementary Figure S9**). Taken together, these results indicate that the home-made cell-free system is compatible with lyophilization both for constitutive protein expression and for toehold RNA sensing reactions. The efficiency loss during storage at room temperature indicates that the preservation conditions should be addressed for further optimization and the use of cryoprotectants (Karig et al., 2017).

Increasing PCR-Product Stability in Home-Made Cell-free Reactions

The use of linear DNA for CFTS offers several advantages compared to plasmid DNA that has to be generated by cloning and then amplified using bacterial cultures prior to use (Yang et al., 1980; Bassett and Rawson, 1983; Hoffmann et al., 2001; Michel-Reydellet et al., 2005; Seki et al., 2008; Shrestha et al., 2012; Sun et al., 2014; Marshall et al., 2017). Linear DNA can be produced by PCR amplification from a lacZ gene template, enabling the design of toehold structures as primer tails for rapid screening of CFTS libraries. The use of linear dsDNA in cell-free, however, is affected by nuclease-mediated DNA degradation in crude cell extracts. This effect is primarily due to the action of exonuclease V (Yang et al., 1980; Bassett and Rawson, 1983) (encoded in the RecBCD operon) (Dillingham and Kowalczykowski, 2008) and endonuclease I (encoded by endA), which are the dominant sources of endonuclease activity in E. coli (I. R. Lehman et al., 1962). Previous studies have shown an increase in the efficiency of cell-free protein synthesis from PCR products using several approaches, such as supplementing with exonuclease V inhibitors (Sitaraman et al., 2004; Sun et al., 2014), using modified nucleotides downstream the PCR reaction (Hoffmann et al., 2001), incorporating competitive DNA strands containing χ-sites (Marshall et al., 2017), adding dsDNA binding proteins (Zhu et al., 2020), depleting exonuclease V in the crude extracts (Seki et al., 2008), deleting RecBCD with the lambda phage red recombinase (Michel-Reydellet et al., 2005), and generating strains lacking endA (Michel-Reydellet et al., 2005; Hong et al., 2015).

To increase the stability of PCR-derived linear DNA, we explored a CRISPRi-based strategy (Larson et al., 2013) to control the genes encoding exonuclease V and endonuclease I in cells before harvesting. We designed sgRNAs targeting the 5' ends of these genes and an IPTG-inducible dCas9 (**Supplementary Figure S10A–C**) and transformed *E coli* strain BL21 DE3 STAR (hereafter named CRISPRi+). As a control, an off-target sgRNA that does not have a known target in the E.coli genome (Nuñez et al., 2017) was transformed into *E coli* strain BL21 DE3 STAR and tested in parallel (hereafter named CRISPRi-). A growth comparison was performed between the CRISPRi+, CRISPRi- and the basal genotype strain BL21 DE3 STAR, finding a slightly lower growth rate for the CRISPRi+ (**Supplementary Figure S10D**).

This CRISPRi strategy is cost-effective as it does not require the supplementation with extra components, or *in vitro* modifications of the PCR product. To test this approach, our crude extracts of each genotype (BL21 DE3 STAR/CRISPRi+, BL21 DE3 STAR/CRISPRi- and BL21 DE3 STAR) were prepared and the stability of a 1.5 Kb linear DNA was quantified over a 2-h time course using agarose gel electrophoresis (**Supplementary Figure S11A**). The following exponential degradation model was fitted to the data:

$$DNA\,(t) = DNA_0 \cdot e^{-K_{deg}t}$$

where DNA_0 is the initial concentration of DNA, $DNA(t)$ is the remaining concentration of DNA at a given time t, and K_{deg} the degradation rate of the DNA. Modelling was undertaken assuming the lower plateau equal to zero, and the ordinary least squares approach was used for parameter estimation. Model screening indicated no statistical evidence of the parameters DNA_0 being different across the data sets evaluated (F-test for extra sum of squares, $p = 0.9285$), finding the best fit value at $DNA_0 = 0.9797$ (**Supplementary Figure S11B**). Statistical analysis of the K_{deg} parameters showed significant differences between cell extract genotypes (F-test for extra sum of squares, $p < 0.001$). Degradation rate decreased significantly between the CRISPRi-optimized extracts ($K_{deg} = 0.008/min$) to the two controls ($K_{deg} = 0.022/min$ for the CRISPRi-negative control, and $K_{deg} = 0.040/min$ for the BL21-derived crude extract). Also, the CRISPRi optimized extracts significantly increased protein production capacity from PCR products, leading to a two-fold increase in the sfGFP endpoint signal (**Supplementary Figure S12A**).

Plasmid-derived endpoint expression was also measured in the three genotypes. No statistical differences were found between these values (**Supplementary Figure S12B**). The linear/plasmid ratio of end-point sfGFP fluorescence was also calculated for each batch, revealing statistical differences between CRISPRi+ and each of the other two genotypes (**Supplementary Figure S12C**). Taken together, these results indicate that CRISPRi + strain enhances stability of linear DNA, and increases expression yields of sfGFP from PCR-derived linear templates.

We observed a small increase in variability in the PCR-derived linear DNA expression of sfGFP with respect to the expression from plasmid DNA. This was observed across the three genotypes, with coefficients of variability of 0.89 and 1.29 for the expression of sfGFP in CRISPRi+ from plasmid and linear DNA, respectively (**Supplementary Table S14**).

Fast Prototyping *de-novo* Designed Sensors

We prepared CRISPRi-optimized extracts using the BL21-Gold-dLac strain and tested the RNA sensing capability of ZIKV toehold sensors expressed directly from PCR linear products. Using these CRISPRi extracts, both ZIKV sensors, 8 and 27, were able to detect RNA in a sequence-specific manner and behaved similarly to the plasmid derived sensors (**Supplementary Figure S13**).

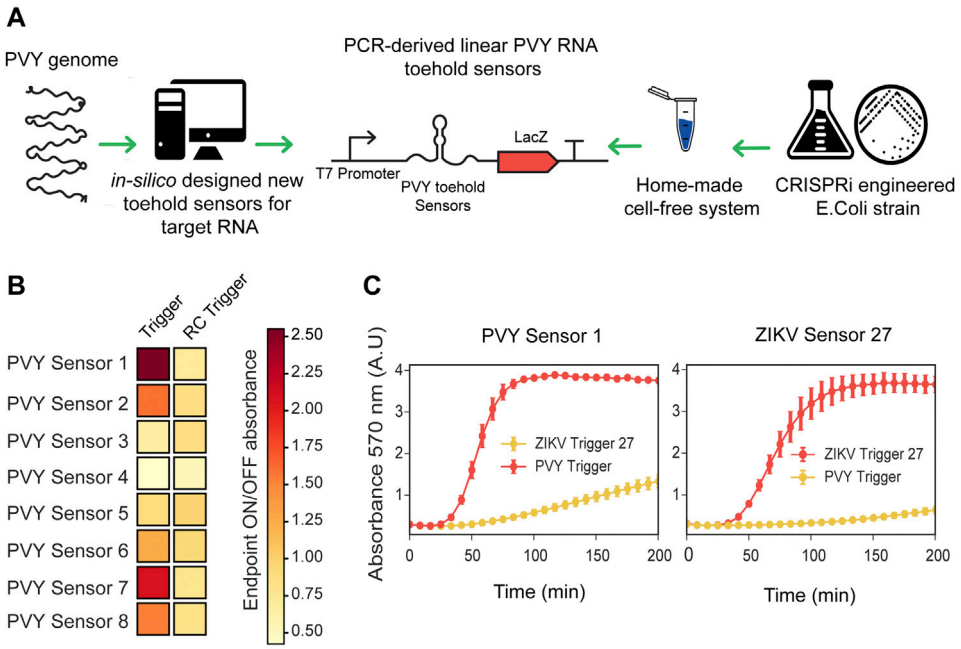


FIGURE 3 | Fast prototyping *de-novo* designed sensors in low-cost optimized cell-free systems. **(A)** Scheme representation of fast prototyping *de-novo* designed RNA toeholds sensors against synthetic fragments of PVY virus. **(B)** PCR-purified transcriptional units (at 10 nM final concentration) encoding for PVY RNA toehold sensors were incubated with synthetic RNA direct or RNA reverse complementary (RC) (at 300 nM final concentration) and absorbance at 570 nm was measured in plate reader. Endpoint ON/OFF absorbance was calculated at 200 min with respect to an untriggered control. Heatmap values correspond to the average of six experimental replicates from two PCR amplifications performed on different days. **(C)** PVY Sensor 1 and ZIKV Sensor 27 encoded in plasmids (at 3.5 nM final concentration) were incubated with ZIKV Trigger 27 RNA or with PVY Trigger RNA (at 480 nM final concentration). Plotted values correspond to arithmetic mean and standard deviation of two independent experiments.

Subsequently, we used these CRISPRi-optimized crude extracts to prototype novel RNA sensors for Potato Virus Y (PVY), a virus affecting local agriculture in Chile. First, we selected a conserved sequence from a wide range of PVY strains (**Supplementary Figure S14**), which was subsequently generated as a synthetic fragment, transcribed *in vitro* and used for screening potential toehold sensors (**Figure 3**; **Supplementary Table S5, 6**).

The majority of published toehold sensors have been designed with NUPACK, a software tool for modeling the thermodynamics properties of DNA and RNA molecules (Green et al., 2014; Pardee et al., 2016; Ma et al., 2018). In this work, we have implemented an algorithm that leverages NUPACK and can be generally applied to the analysis and design of novel toehold sensors according to ideal designs (**Supplementary Figure S1**). Using NupackSensors (see Methods), we analyzed the consensus sequence identified previously and selected eight different toehold RNA sensors for experimental screening. Transcriptional units encoding for each PVY sensor were prepared by PCR amplification and incubated with synthetic RNA direct or reverse complementary (RC) to test their ability to detect RNA in a sequence-specific manner (**Figure 3A**; **Supplementary Figure S15**). A control without RNA was used as a baseline in the ON/OFF endpoint calculations. Among the eight sensors tested, two PVY sensors had ON/OFF endpoint measurements greater than two and background level lower than 1.2 (**Figure 3B**).

In order to validate the fast characterization of PVY toehold sensors using PCR-derived linear DNA, the best scored sensors (PVY Sensor one and PVY Sensor 7) were cloned in plasmids and compared side by side to the plasmid version of ZIKV Sensor 27 in a gradient of input DNA concentrations ranging from 1 to 4 nM (**Supplementary Figure S16**). Consistent with previous observations (**Supplementary Figure S7**), input DNA concentrations affect the performance of the toehold sensors. Some sensors (e.g. PVY Sensor 7) showed significant background activation (i.e., being active without a trigger) (**Supplementary Figure S16**). This background activation is proportional to input DNA concentration. The best performance of PVY Sensor seven was found at a lower concentration (1–2 nM) than for PVY Sensor 1, whose optimal concentration was between 2 and 4 nM (**Supplementary Figure S16**). Once a good working concentration was identified, PVY Sensor one and ZIKV Sensor 27 were incubated in home-made cell-free reactions (at 3.5 nM final concentration) and supplemented with ZIKV Trigger 27 or PVY RNA (at 480 nM final concentration). PVY Sensor one displayed sequence-specific activation comparable with ZIKV Sensor 27 (**Figure 3C**). These results showed that the fast-prototyping pipeline presented here, followed by an optimization of input DNA concentration, is a suitable strategy for identifying novel RNA toehold sensors of good performance.

DISCUSSION

The increasing frequency of infectious disease outbreaks and the long-standing impact of viral pathogens on agriculture and livestock highlight the need for more efficient, cost-effective and ubiquitous surveillance of viruses. Cell-free toehold sensors have emerged as a promising technology to address this challenge (Pardee et al., 2014; Pardee et al., 2016; Ma et al., 2018). However, in Latin America the high cost and shipping constraints of the commercial *in vitro* transcription and translation systems has limited the local production and distributed development of new sensors. These constraints could lead to centralized production models and fragile technological dependencies prone to failure, as evidenced recently by the disrupted supply chain of diagnostic reagents for SARS-CoV-2 testing (The COVID-19 testing debacle, 2020; Nkengasong, 2020). These shortages highlight the relevance of biotechnological tools that can be shared, produced and deployed locally at a lower cost and without the restrictions of intellectual property (Kellner et al., 2020; Mascuch et al., 2020; Thomas G.W. et al., 2020; JOGL (Just One Gigant Lab)-Open COVID19 Initiative, 2020; Reclone, 2020). In-house prepared cell-free toehold sensors, implemented in cell lysates or home-made PURE systems (Shimizu et al., 2001; Villarreal et al., 2018; Lavickova and Maerkl, 2019), could contribute to the decentralization of this technology and its adaptation to local needs.

Here, we demonstrate the implementation and rapid prototyping of novel sensors using low-cost, locally produced cell-free lysates. As our cost analysis showed, we estimate that this approach could enable the manufacture of bespoke diagnostic tests to meet local needs (for instance in Chile the cost was $0.069 USD/test, considering a 5 μL reaction). While the RNA amplification step still requires commercial inputs, such reactions can also be made in-house (Compton, 1991; Aufdembrink et al., 2020). Future work will focus on the in-house preparation of NASBA from locally purified, and IP-free RNAseH, T7 RNAP, pyrophosphatase, and reverse transcriptase enzymes (Aufdembrink et al., 2020; Reclone, 2020; Matute et al., 2021).

Future lines of research will also address the cost efficiency of crude extract preparation. 45–60% of the maltodextrin-based reaction cost relates to growing, inducing, and processing cells, where IPTG (used for induction) and the micro-chromatography columns (used for separation of the cell extract from the glass beads) add up to more than one third of the cost per reaction (**Supplementary Figure S4**). Therefore, sonication (Shrestha et al., 2012), or other column-free methods for disrupting the cells should be evaluated, along with the use of alternatives to the relatively expensive IPTG such as use of the autolysis strain (Didovyk et al., 2017) or the autoinduction media formulation (Levine et al., 2019).

A key aspect of cell-free biology is the possibility of freeze-drying genetic systems into pellets or paper-based reactions that can be transported at room temperature and deployed in the field. Here, we have tested the performance of toehold sensors in lyophilized cell extracts stored at room temperature for up to 7 days. The decay of the activity after lyophilization highlights the need for further optimization of the storage conditions. Modifications of the storage atmosphere conditions either by replacing air with an inert gas, such as nitrogen or argon, in addition to the use of vacuum sealing, oxygen absorbers, and desiccants (Jung et al., 2020), remains to be explored.

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. All the data collected in this work is publicly available at: https://zenodo.org/record/4988741#.YM0pzWhKhPY.

AUTHOR CONTRIBUTIONS

FF conceived the project, supervised the research, assisted in preparing the figures and writing the manuscript. AA designed and performed the experiments, analyzed the results, prepared the figures and wrote the first draft of the manuscript. ND assisted the coding writing, data analysis and statistics. JP, FC, CG, and TM conceived and performed experiments and analyzed the results. KP, JM, JH, and ND supervised the research and assisted in writing the manuscript.

FUNDING

AA was founded by CONICYT scholarship No. 21140714. FF, AA and TM were funded by ANID—Millennium Science Initiative Program—ICN17_022, Fondo de Desarrollo de Áreas Prioritarias—Center for Genome Regulation ANID/FONDAP/15090007 and ICGEB CRP/CHL19-01. FF, AA, and JP were funded by Proyecto Investigación Interdisciplinaria 2018 VRI. FC was supported by EPSRC EP/R014000/1 under the project LCVD: Low-cost Cell-extract Viral Diagnostics.

ACKNOWLEDGMENTS

We thank Jonathan Dhalin (Technical University of Denmark) and Vincent Noireaux (University of Minnesota) for providing positive controls and examples for cell-free protein synthesis reactions at early stages of the project. We thank the Pardee lab (University of Toronto) for welcoming Anibal during his internship. We thank members of the lab de tecnología libre (Pontificia Universidad Católica de Chile) for comments and feedback on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2021.727584/full#supplementary-material

REFERENCES

Anonymous (2020). The COVID-19 testing debacle (2020). *Nat. Biotechnol.* 38, 653. doi:10.1038/s41587-020-0575-3

Aufdembrink, Lauren. M., Khan, Pavana, J Gaut, Nathaniel, Adamala, Katarzyna. P., and Aaron, E. (2020). Highly specific, multiplexed isothermal pathogen detection with fluorescent aptamer readout. *RNA* 26, 1283–1290. doi:10.1261/rna.075192.120

Bassett, C. L., and Rawson, J. R. (1983). *In vitro* coupled transcription-translation of linear DNA fragments in a lysate derived from a recB rna pnp strain of Escherichia coli. *J. Bacteriol.* 156, 1359–1362. doi:10.1128/JB.156.3.1359-1362.1983

Bundy, B. C., Franciszkowicz, M. J., and Swartz, J. R. (2008). Escherichia coli-based cell-free synthesis of virus-like particles. *Biotechnol. Bioeng.* 100, 28–37. doi:10.1002/bit.21716

Caschera, F., and Noireaux, V. (2015a). A cost-effective polyphosphate-based metabolism fuels an all E. coli cell-free expression system. *Metab. Eng.* 27, 29–37. doi:10.1016/j.ymben.2014.10.007

Caschera, F., and Noireaux, V. (2015b). Preparation of amino acid mixtures for cell-free expression systems. *BioTechniques* 58 , 40–43. doi:10.2144/000114249

Compton, J. (1991). Nucleic acid sequence-based amplification. *Nature* 350, 91–92. doi:10.1038/350091a0

Didovyk, A., Tonooka, T., Tsimring, L., and Hasty, J. (2017). Rapid and Scalable Preparation of Bacterial Lysates for Cell-free Gene Expression. *ACS Synth. Biol.* 6, 2198–2208. doi:10.1021/acssynbio.7b00253

Dillingham, M. S., and Kowalczykowski, S. C. (2008). RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks. *Microbiol. Mol. Biol. Rev.* 72, 642–671. doi:10.1128/MMBR.00020-08

Dudley, Q. M., Anderson, K. C., and Jewett, M. C. (2016). Cell-Free Mixing of *Escherichia coli* Crude Extracts to Prototype and Rationally Engineer High-Titer Mevalonate Synthesis. *ACS Synth. Biol.* 5, 1578–1588. doi:10.1021/acssynbio.6b00154

Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016). The All *E. coli* TX-TL Toolbox 2.0: A Platform for Cell-free Synthetic Biology. *ACS Synth. Biol.* 5, 344–355. doi:10.1021/acssynbio.5b00296

Garenne, D., and Noireaux, V. (2019). Cell-free transcription–translation: engineering biology from the nanometer to the millimeter scale. *Curr. Opin. Biotechnol.* 58, 19–27. doi:10.1016/j.copbio.2018.10.007

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi:10.1038/nmeth.1318

Green, A. A., Silver, P. A., Collins, J. J., and Yin, P. (2014). Toehold Switches: De-novo-designed Regulators of Gene Expression. *Cell* 159, 925–939. doi:10.1016/j.cell.2014.10.002

Grubbe, W. S., Rasor, B. J., Krüger, A., Jewett, M. C., and Karim, A. S. (2020). Cell-free styrene biosynthesis at high titers. *Metab. Eng.* 61, 89–95. doi:10.1016/j.ymben.2020.05.009

Hoffmann, Thomas., Nemetz, Cordula., Schweizer, Regina., Mutter, Wolfgang., and Watzele, Mandred. (2001). "High-Level Cell-free Protein Expression from PCR-Generated DNA Templates," in *Cell-Free Translation Systems*. Editor A. S. Spirin (Berlin, Heidelberg: Springer).

Hong, S. H., Kwon, Y.-C., Martin, R. W., Des Soye, B. J., de Paz, A. M., Swonger, K. N., et al. (2015). Improving Cell-free Protein Synthesis through Genome Engineering of *Escherichia coli* Lacking Release Factor 1. *ChemBioChem* 16, 844–853. doi:10.1002/cbic.201402708

Horiya, S., Bailey, J. K., and Krauss, I. J. (2017). "Directed Evolution of Glycopeptides Using mRNA Display," in *Methods in Enzymology* (Amsterdam, Netherlands: Elsevier), 83–141. doi:10.1016/bs.mie.2017.06.029

JOGL (Just One Gigant Lab)-Open COVID19 Initiative (2020). JOGL-open COVID19 Initiat. Available at: https://app.jogl.io/program/opencovid19.

Jung, J. K., Alam, K. K., Verosloff, M. S., Capdevila, D. A., Desmau, M., Clauer, P. R., et al. (2020). Cell-free biosensors for rapid detection of water contaminants. *Nat. Biotechnol.* 38, 1451–1459. doi:10.1038/s41587-020-0571-7

Karig, D. K., Bessling, S., Thielen, P., Zhang, S., and Wolfe, J. (2017). Preservation of protein expression systems at elevated temperatures for portable therapeutic production. *J. R. Soc. Interf.* 14, 20161039. doi:10.1098/rsif.2016.1039

Kellner, M. J., Ross, J. J., Schnabl, J., Dekens, M. P. S., Heinen, R., Grishkovskaya, I., et al. (2020). A rapid, highly sensitive and open-access SARS-CoV-2 detection assay for laboratory and home testing. *Mol. Biol.* doi:10.1101/2020.06.23.166397

Kim, T.-W., Keum, J.-W., Oh, I.-S., Choi, C.-Y., Park, C.-G., and Kim, D.-M. (2006). Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system. *J. Biotechnol.* 126, 554–561. doi:10.1016/j.jbiotec.2006.05.014

Kim, T.-W., Kim, H.-C., Oh, I.-S., and Kim, D.-M. (2008). A highly efficient and economical cell-free protein synthesis system using the S12 extract of *Escherichia coli*. *Biotechnol. Bioproc. Eng.* 13, 464–469. doi:10.1007/s12257-008-0139-8

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. doi:10.1038/srep08663

Laohakunakorn, N., Grasemann, L., Lavickova, B., Michielin, G., Shahein, A., Swank, Z., et al. (2020). Bottom-Up Construction of Complex Biomolecular Systems With Cell-free Synthetic Biology. *Front. Bioeng. Biotechnol.* 8, 213. doi:10.3389/fbioe.2020.00213

Larson, M. H., Gilbert, L. A., Wang, X., Lim, W. A., Weissman, J. S., and Qi, L. S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* 8, 2180–2196. doi:10.1038/nprot.2013.132

Lavickova, B., and Maerkl, S. J. (2019). A Simple, Robust, and Low-Cost Method to Produce the PURE Cell-free System. *ACS Synth. Biol.* 8, 455–462. doi:10.1021/acssynbio.8b00427

Lehman, I. R., Roussos, G. G., and Pratt, E. A. (1962). The Deoxyribonucleases of *Escherichia coli*. *Int. J. Biol. Chem.* 233.

Lehr, F.-X., Hanst, M., Vogel, M., Kremer, J., Göringer, H. U., Suess, B., et al. (2019). Cell-Free Prototyping of AND-Logic Gates Based on Heterogeneous RNA Activators. *ACS Synth. Biol.* 8, 2163–2173. doi:10.1021/acssynbio.9b00238

Levine, M. Z., So, B., Mullin, A. C., Watts, K. R., and Oza, J. P. (2019). Redesigned upstream processing enables a 24-hour workflow from *E. coli* cells to cell-free protein synthesis. *Synth. Biol.*. doi:10.1101/729699

Liu, X., Silverman, A. D., Alam, K. K., Iverson, E., Lucks, J. B., Jewett, M. C., et al. (2020). Design of a Transcriptional Biosensor for the Portable, On-Demand Detection of Cyanuric Acid. *ACS Synth. Biol.* 9, 84–94. doi:10.1021/acssynbio.9b00348

Ma, D., Shen, L., Wu, K., Diehnelt, C. W., and Green, A. A. (2018). Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment. *Synth. Biol.* 3 , ysy018. doi:10.1093/synbio/ysy018

Marshall, R., Maxwell, C. S., Collins, S. P., Beisel, C. L., and Noireaux, V. (2017). Short DNA containing χ sites enhances DNA stability and gene expression in *E. coli* cell-free transcription-translation systems: Enhancing TXTL-Based Expression With χ-Site DNA. *Biotechnol. Bioeng.* 114, 2137–2141. doi:10.1002/bit.26333

Mascuch, S. J., Fakhretaha-Aval, S., Bowman, J. C., Ma, M. T. H., Thomas, G., Bommarius, B., et al. (2020). A blueprint for academic labs to produce SARS-CoV-2 RT-qPCR test kits. *J. Biol. Chem.*, 015434. doi:10.1074/jbc.RA120.015434

Matute, T., Nuñez, I., Rivera, M., Reyes, J., Blázquez-Sánchez, P., Arce, A., et al. (2021). Homebrew reagents for low cost RT-LAMP. *Public Glob. Health*. doi:10.1101/2021.05.08.21256891

McNerney, M. P., Zhang, Y., Steppe, P., Silverman, A. D., Jewett, M. C., and Styczynski, M. P. (2019). Point-of-care biomarker Quantification enabled by sample-specific calibration. *Sci. Adv.* 5, eaax4473. doi:10.1126/sciadv.aax4473

Michel-Reydellet, N., Woodrow, K., and Swartz, J. (2005). Increasing PCR Fragment Stability and Protein Yields in a Cell-free System with Genetically Modified *Escherichia coli* Extracts. *J. Mol. Microbiol. Biotechnol.* 9, 26–34. doi:10.1159/000088143

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. doi:10.1042/BST20170011

Niederholtmeyer, H., Sun, Z. Z., Hori, Y., Yeung, E., Verpoorte, A., Murray, R. M., et al. (2015). Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife* 4, e09771. doi:10.7554/eLife.09771

Nirenberg, M. W., and Matthaei, J. H. (1961). The Dependence of Cell- Free Protein Synthesis in E. coli upon Naturally Occurring or Synthetic Polyribonucleotides. *Proc. Natl. Acad. Sci. U. S. A.* 47, 1588–1602.

Nkengasong, John. (2020). Let Africa into the market for COVID-19 diagnostics. *Nature* 580, 556. doi:10.1038/d41586-020-01265-0

Nuñez, I. N., Matute, T. F., Del Valle, I. D., Kan, A., Choksi, A., Endy, D., et al. (2017). Artificial Symmetry-Breaking for Morphogenetic Engineering Bacterial Colonies. *ACS Synth. Biol.* 6, 256–265. doi:10.1021/acssynbio.6b00149

Pardee, K., Green, A. A., Ferrante, T., Cameron, D. E., DaleyKeyser, A., Yin, P., et al. (2014). Paper-Based Synthetic Gene Networks. *Cell* 159, 940–954. doi:10.1016/j.cell.2014.10.004

Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., et al. (2016). Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* 165, 1255–1266. doi:10.1016/j.cell.2016.04.059

Pieters, P. A., Nathalia, B. L., van der Linden, A. J., Yin, P., Kim, J., Huck, W. T. S., et al. (2021). Cell-Free Characterization of Coherent Feed-Forward Loop-Based Synthetic Genetic Circuits. *ACS Synth. Biol.* 10, 1406–1416. doi:10.1021/acssynbio.1c00024

Pollak, B., Matute, T., Nuñez, I., Cerda, A., Lopez, C., Vargas, V., et al. (2020). Universal Loop assembly (uLoop): open, efficient, and species-agnostic DNA fabrication. *Synth. Biol.* 5, ysaa001. doi:10.1101/744854

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

Reclone (2020). Recloneorg Glob. Collab. Equitable Access Biotechnol. Available at: https://reclone.org/.

Rustad, M., Eastlund, A., Jardine, P., and Noireaux, V. (2018). Cell-free TXTL synthesis of infectious bacteriophage T4 in a single test tube reaction. *Synth. Biol.* 3, ysy002. doi:10.1093/synbio/ysy002

Salehi, A. S. M., Yang, S. O., Earl, C. C., Shakalli Tang, M. J., Porter Hunt, J., Smith, M. T., et al. (2018). Biosensing estrogenic endocrine disruptors in human blood and urine: A RAPID cell-free protein synthesis approach. *Toxicol. Appl. Pharmacol.* 345, 19–25. doi:10.1016/j.taap.2018.02.016

Salgado, H., Peralta-Gil, M., Gama-Castro, S., Santos-Zavaleta, A., Muñiz-Rascado, L., García-Sotelo, J. S., et al. (2013). RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res.* 41, D203–D213. doi:10.1093/nar/gks1201

Seki, E., Matsuda, N., Yokoyama, S., and Kigawa, T. (2008). Cell-free protein synthesis system from *Escherichia coli* cells cultured at decreased temperatures improves productivity by decreasing DNA template degradation. *Anal. Biochem.* 377, 156–161. doi:10.1016/j.ab.2008.03.001

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., et al. (2001). Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755. doi:10.1038/90802

Shin, J., and Noireaux, V. (2012). An *E. coli* Cell-free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells. *ACS Synth. Biol.* 1, 29–41. doi:10.1021/sb200016s

Shrestha, P., Holland, T. M., and Bundy, B. C. (2012). Streamlined extract preparation for *Escherichia coli* -based cell-free protein synthesis by sonication or bead vortex mixing. *BioTechniques* 53, 163–174. doi:10.2144/0000113924

Siegal-Gaskins, D., Tuza, Z. A., Kim, J., Noireaux, V., and Murray, R. M. (2014). Gene Circuit Performance Characterization and Resource Usage in a Cell-free "Breadboard. *ACS Synth. Biol.* 3, 416–425. doi:10.1021/sb400203p

Silverman, A. D., Akova, U., Alam, K. K., Jewett, M. C., and Lucks, J. B. (2020). Design and Optimization of a Cell-free Atrazine Biosensor. *ACS Synth. Biol.* 9, 671–677. doi:10.1021/acssynbio.9b00388

Silverman, A. D., Kelley-Loughnane, N., Lucks, J. B., and Jewett, M. C. (2019). Deconstructing Cell-free Extract Preparation for *in Vitro* Activation of Transcriptional Genetic Circuitry. *ACS Synth. Biol.* 8, 403–414. doi:10.1021/acssynbio.8b00430

Sitaraman, K., Esposito, D., Klarmann, G., Le Grice, S. F., Hartley, J. L., and Chatterjee, D. K. (2004). A novel cell-free protein synthesis system. *J. Biotechnol.* 110, 257–263. doi:10.1016/j.jbiotec.2004.02.014

Sun, Z. Z., Hayes, C. A., Shin, J., Caschera, F., Murray, R. M., and Noireaux, V. (2013). Protocols for Implementing an *Escherichia coli* Based TX-TL Cell-free Expression System for Synthetic Biology. *J. Vis. Exp.*, 50762. doi:10.3791/50762

Sun, Z. Z., Yeung, E., Hayes, C. A., Noireaux, V., and Murray, R. M. (2014). Linear DNA for Rapid Prototyping of Synthetic Biological Circuits in an *Escherichia coli* Based TX-TL Cell-free System. *ACS Synth. Biol.* 3, 387–397. doi:10.1021/sb400131a

Swartz, J. (2020). Opportunities toward hydrogen production biotechnologies. *Curr. Opin. Biotechnol.* 62, 248–255. doi:10.1016/j.copbio.2020.03.002

Takahashi, M. K., Chappell, J., Hayes, C. A., Sun, Z. Z., Kim, J., Singhal, V., et al. (2015). Rapidly Characterizing the Fast Dynamics of RNA Genetic Circuitry with Cell-free Transcription–Translation (TX-TL) Systems. *ACS Synth. Biol.* 4, 503–515. doi:10.1021/sb400206c

Takahashi, M. K., Tan, X., Dy, A. J., Braff, D., Akana, R. T., Furuta, Y., et al. (2018). A low-cost paper-based synthetic biology platform for analyzing gut microbiota and host biomarkers. *Nat. Commun.* 9, 3347. doi:10.1038/s41467-018-05864-4

Tayar, A. M., Karzbrun, E., Noireaux, V., and Bar-Ziv, R. H. (2017). Synchrony and pattern formation of coupled genetic oscillators on a chip of artificial cells. *Proc. Natl. Acad. Sci.* 114, 11609–11614. doi:10.1073/pnas.1710620114

Thavarajah, W., Silverman, A. D., Verosloff, M. S., Kelley-Loughnane, N., Jewett, M. C., and Lucks, J. B. (2020). Point-of-Use Detection of Environmental Fluoride *via* a Cell-free Riboswitch-Based Biosensor. *ACS Synth. Biol.* 9, 10–18. doi:10.1021/acssynbio.9b00347

Thomas, G. W., Dailey, Gina. M., and Claire Dugast-Darzacq, M. N. E. (2020). BEARmix: Basic Economical Amplification Reaction one-step RT-qPCR master mix. *BEARmix*. Available at: https://gitlab.com/tjian-darzacq-lab/bearmix.

Verosloff, M., Chappell, J., Perry, K. L., Thompson, J. R., and Lucks, J. B. (2019). PLANT-dx: A Molecular Diagnostic for Point-of-Use Detection of Plant Pathogens. *ACS Synth. Biol.* 8, 902–905. doi:10.1021/acssynbio.8b00526

Villarreal, F., Contreras-Llano, L. E., Chavez, M., Ding, Y., Fan, J., Pan, T., et al. (2018). Synthetic microbial consortia enable rapid assembly of pure translation machinery. *Nat. Chem. Biol.* 14, 29–35. doi:10.1038/nchembio.2514

Yang, H. L., Ivashkiv, L., Chen, H. Z., Zubay, G., and Cashel, M. (1980). Cell-free coupled transcription-translation system for investigation of linear DNA segments. *Proc. Natl. Acad. Sci.* 77, 7029–7033. doi:10.1073/pnas.77.12.7029

Yelleswarapu, M., van der Linden, A. J., van Sluijs, B., Pieters, P. A., Dubuc, E., de Greef, T. F. A., et al. (2018). Sigma Factor-Mediated Tuning of Bacterial Cell-free Synthetic Genetic Oscillators. *ACS Synth. Biol.* 7, 2879–2887. doi:10.1021/acssynbio.8b00300

Zhu, B., Gan, R., Cabezas, M. D., Kojima, T., Nicol, R., Jewett, M. C., et al. (2020). Increasing cell-free gene expression Yields from linear templates in *Escherichia coli* and *Vibrio natriegens* extracts by using DNA-binding proteins. *Biotechnol. Bioeng. Bit*, 27538. doi:10.1002/bit.27538

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership