NORTHWESTERN UNIVERSITY

Cell-free Technologies for On-demand Glycoprotein Biomanufacturing and Hands-on Biology Education

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Chemical Engineering

By

Jessica Carol Stark

EVANSTON, ILLINOIS

September 2019

Since their introduction nearly a century ago, protein vaccines and therapeutics have revolutionized our ability to prevent and treat human disease. However, existing production processes for biopharmaceuticals are technically complex and rely on living cells, which necessitates highly centralized manufacturing in large-scale production facilities, specialized equipment, and cold-chain distribution. With increasing demands for medicines tailored to individuals or relatively small patient populations, there is growing interest in scaled down bioprocesses that can accommodate production of many different biologic molecules. In addition, the need for cold-chain refrigeration limits our ability to supply life-saving biologics to underdeveloped regions as well as in emergency situations, prompting efforts to develop on-demand protein production technologies that can be distributed without refrigeration. Overall, technologies for small-scale, decentralized biomanufacturing represent an emerging paradigm that promises to enable portable and personalized protein medicines. Still, existing technologies have been limited in their ability to produce glycosylated protein products, which represent over 70% of protein therapeutics and vaccines approved or in clinical development.

This work seeks to address this limitation through the development of new technologies for on-demand biomanufacturing of glycosylated protein therapeutics and vaccines. We first developed cell-free glycoprotein synthesis (CFGpS) technology with the ability to produce glycosylated protein therapeutics at the point-of-care. The CFGpS platform uses crude *Escherichia coli* cell lysates containing the biological machinery for both protein synthesis and glycosylation to produce glycoproteins in simple, one-pot reactions. We show that CFGpS can produce glycoprotein medicines such as erythropoietin, as well as proteins bearing a range of bacterial and eukaryotic glycans. In parallel, we developed the *in vitro* bioconjugate <u>va</u>ccine e<u>x</u>pression (iVAX) platform that enables on-demand and portable biosynthesis of antibacterial vaccines via coordinated cell-free protein synthesis and glycosylation. iVAX reactions can synthesize single doses of vaccines against diverse bacterial pathogens in one hour, including the highly virulent *Franciscella tularensis* subsp. *tularensis* (type A) strain Schu S4 and pathogenic *E. coli* strains O78 and O7. In particular, we showed that anti-*F. tularensis* vaccines can be produced in iVAX for ~$6 per human dose and elicited pathogen-specific immune responses in mice. Together, the CFGpS and iVAX platforms represent key first steps toward modular, on-demand production of glycosylated protein

therapeutics and vaccines, joining an emerging set of decentralized biomanufacturing platforms that promise to increase global access to costly drugs.

Finally, we demonstrate that portable cell-free protein production platforms can be adapted to enable educational kits for teaching molecular and synthetic biology, which we call BioBits^{TM} kits. The BioBits^{TM} kits alleviate many of the economic and logistical challenges associated with implementing hands-on molecular and synthetic biology activities in classrooms and other non-laboratory settings. As such, these kits have the potential to increase scientific literacy through the integration of cutting-edge molecular and synthetic biology topics into K-12 STEM education.

Acknowledgements

Every PhD journey is unique, and I have many people to thank for contributing to mine. First of all, thank you to my advisor, Dr. Michael Jewett, for his guidance and for the freedom to create a PhD project that couldn't have been more perfect for me. You've helped me learn how to be myself in science, which is perhaps the most important lesson an advisor can convey. Thank you to my committee for your support: Dr. Matt DeLisa, Dr. Bill Miller, and Dr. Josh Leonard. To Matt, who was also my undergraduate research advisor, the fact that you've contributed to my growth as an independent scientist for almost 10 years is truly special and I am so grateful. Thank you for your unwavering support as a colleague, mentor, and advocate. Thapakorn Jaroentomeechai, I couldn't have asked for a better partner for the CFGpS project. Thank you for being so generous with your time, knowledge, and gifts from Thailand. Thanks to Ally Huang, co-founder of the BioBits™ project, I am lucky to have had a wonderful collaborator for this unconventional and important work and I am excited to see how it continues to grow under your leadership. Thanks also to collaborators Ty Moeller, Dr. Peter Nguyen, Aravind Natarajan, Thomas Ferrante, Nina Donghia, Dr. Melissa Takahashi, Dr. Keith Pardee, and Dr. James Collins. Thank you to Dr. Linda Broadbelt, my teaching apprentice program mentor and academic role model. On multiple occasions, you have gone out of your way to offer your support and it means the world to me.

If the PhD is a battlefield, then your labmates are the ones with you on the front lines. Thank you to Anne d'Aquino, who has literally sat next to me through almost all of it – my PhD wouldn't have been filled with nearly so many bad puns without you. Thank you for single-handedly keeping me sane. To Dr. Ashty Karim, our friendship has been one of the greatest gifts of my professional career. Whether collaborating or competing, I know I can always count on your support. To Dr. Erik Carlson and Dr. Jessica Perez, my Jewett lab role models, thank you for being such wonderful examples of scientists, educators, and friends. To Jazzy Hershewe, I am so lucky to be able to work with such a close friend. Thank you for taking a chance on team glyco. To Weston Kightlinger, you have always listened to and valued my opinion and for that I am grateful. I had the privilege of working with two extremely talented

undergraduates: Karen Hsu and Rachel Dubner, who made mentoring a true joy. Thank you to Lauren for everything you do really, but especially for being my first bench mate and for being you. Thank you to Tom Martinez, an honorary Jewett lab member, for being open to new educational adventures and for being the first and most enthusiastic advocate of BioBits™. Final thanks to all other members of team glyco past and present: Dr. Jennifer Schoborg, Dr. James Kath, and Katie Warfel.

Contrary to popular belief, I did try to maintain a life outside of lab. Thank you to the puppy pals, Danica, Paul, and Rye Carlson, Dr. Jennifer Greene and DJ Jellybeans, and Dr. Louisa Saveride, for providing much-needed comic relief and delicious snacks on the regular. Thank you to Jill Mockaitis and Heather Martello, my weekly trips to yoga were my sanctuary. Thank you to Diana Hamann, Dr. Meghan Fritz, Morgan, and the rest of the Wine Goddess crew for the sacred work that you do. Thank you to Audrey Denvir and Dr. Amit Kumar, who helped us eat our way through Chicago winters. Thank you to Dr. Jacob Olshansky, one of my oldest friends, and Youjin Lee, one of my newest, for moving to Chicago so we could hang out. And, of course, thank you to Molly Jerrard, Meg Taylor, Lauren Regan, Rebecca Varnhagen, Caitlyn Dewitt, Caitie Waldron, Tina Rho, Will Barclay, Dan Bonomo, and Candace Lucke. I am lucky to know so many girls and guys with so much personality.

Finally, I wouldn't be where I am today without the support of my family. Mom and Dad, thank you for everything. It has been so wonderful having you just a short drive away. Christie, I'm learning more and more from you every day. Greg, I'm so glad you stayed in Chicago – thanks for bringing the laughs up to Evanston. Millie, you're a good dog. And Jordan, you always know what to say. Thank you for being a constant source of support during this crazy journey.

Dedication

To my Grandpa, Dr. Robert Martin Stark, who would have loved his copy.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 On-demand biomanufacturing technologies promise to enable portable and personalized medicine

Since the introduction of the tetanus and diphtheria toxoid vaccines in the 1920s[1] and the approval of recombinant insulin in 1982[2], protein vaccines and therapeutics have revolutionized our ability to prevent and treat human disease. However, current biomanufacturing strategies are key contributors to the escalating cost to develop novel biologics (estimated at $2.5B per new molecule in 2014)[3], due to the high costs ($300-500M) and long time scales (4-5 years) associated with building large-scale commercial facilities[4]. Additionally, with advancements in cell line engineering enabling order of magnitude increases in recombinant protein titers[5], as well as increasing demands for medicines tailored to biologically-stratified patient populations[6], there is growing interest in scaled down bioprocesses that can accommodate production of multiple biologic molecules. Finally, the current centralized biomanufacturing paradigm necessitates refrigerated supply chains for distribution of many protein vaccines and therapeutics. The need for cold-chain refrigeration presents significant economic and logistical challenges for supplying life-saving biologics to remote or underdeveloped regions with limited infrastructure as well as in battlefield or emergency situations[7-9].

As a result, novel biomanufacturing paradigms are emerging that can enable decentralized and potentially portable production of protein therapeutics and vaccines at small scales. A handful of technologies have been described including automated sub-liter scale bioreactors[10,11] and downstream processing modules[11] for cell-based biomanufacturing in yeast, as well as lyophilized bacterial[12,13] and mammalian[14] cell-free protein synthesis systems that can be activated with water and combined with portable, automated purification platforms[14] to enable cold chain-independent biomanufacturing. To date, recombinant interferon alpha-2b, human growth hormone, erythropoietin, G-CSF, onconase, diphtheria toxoid, and a panel of ten antimicrobial peptides have been made using point-of-care production technologies, with some achieving clinical-grade purity, safety, and potency[11,14]. However, so far these

technologies have been limited in their ability to produce complex protein biologics, including glycosylated protein therapeutics that represent the majority of FDA-approved protein products.

This work develops new technologies for on-demand biomanufacturing of glycosylated protein therapeutics and vaccines, adding an exciting new capability to the emerging area of portable, on demand biomanufacturing. In addition, we demonstrate that portable, cell-free protein production platforms can be adapted to create educational kits that alleviate many of the challenges associated with implementing hands-on molecular and synthetic biology activities in classrooms.

1.2 Glycosylation plays key roles in the efficacy of protein therapeutics and vaccines

Glycosylation is critically important for the production of recombinant protein therapeutics. Approximately 70% of the >100 protein products approved by U.S. and European regulatory agencies and the ~500 candidates in clinical trials are glycosylated. The majority of glycoprotein therapeutics contain glycans attached primarily to asparagine residues (*N*-linked) or serine or threonine residues (*O*-linked), which are known to impact many therapeutically relevant protein properties including pharmacokinetics, immunogenicity, and biological activity[15-17]. Because of the important roles glycosylation plays in therapeutic efficacy, the majority of glycoprotein therapeutics are produced in mammalian cells to ensure proper modification[18]. However, mammalian cell culture is expensive, laborious, and challenging to scale, resulting in long development and production timelines. These challenges have prompted recent efforts to enable glycoprotein production in *Escherichia coli* through the functional expression of orthogonal glycosylation machinery[19-22]. Bacteria like *E. coli* provide a blank canvas for which to engineer synthetic glycosylation pathways, as they lack native glycosylation machinery. To date, the initial steps of human *N*-linked glycosylation have been recapitulated in *E. coli,* demonstrating production of glycoproteins bearing the eukaryotic trimannosyl core glycan (**Figure 1.1**)[23]. This advance opens the door to production of glycosylated protein therapeutics in bacterial systems.
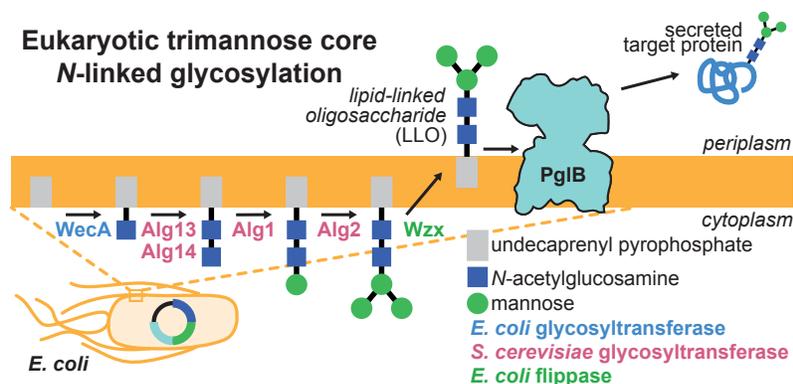
Figure 1.1 Production of proteins bearing the eukaryotic trimannose core glycan in *E. coli*. Expression of a synthetic glycosylation pathway in *E. coli* results in the assembly of the Man$_3$GlcNAc$_2$ (Man: mannose; GlcNAc: *N*-acetylglucosamine) core eukaryotic *N*-glycan on the bacterial lipid undecaprenol pyrophosphate. This glycan structure is then transferred from the lipid-linked oligosaccharide precursor to proteins via an *N*-linked glycosylation reaction catalyzed by the *C. jejuni* oligosaccharyltransferase enzyme PglB.

Glycosylation also plays a critical role in the efficacy of conjugate vaccines, which are among the safest and most effective methods for preventing life-threatening bacterial infections[24]. Conjugate vaccines are composed of an immunogenic (CD4+ T-cell dependent antigen) protein carrier chemically conjugated to a polysaccharide antigen from the surface of the target bacterium, typically capsular polysaccharide (CPS)- or lipopolysaccharide (LPS)-derived antigens. These protein-polysaccharide conjugates induce immune responses directed against the bacterial polysaccharide antigen that are characterized by polysaccharide-specific IgM-to-IgG class switching, memory B cell development, and long-lived T-cell memory[25-29]. As a result, conjugate vaccines have proven to be a highly efficacious and safe strategy for protecting against virulent bacterial pathogens, including *Haemophilus influenzae*, *Neisseria meningitidis*, and *Streptococcus pneumoniae*[29-31], resulting in several marketed products and many others in clinical development[25,30].

Despite their effectiveness, conjugate vaccines are particularly challenging to develop and distribute due to their structural complexity, associated biosafety concerns, and refrigeration requirements. The process to produce conjugate vaccines involves biosynthesis, purification, and chemical conjugation of the polysaccharide and protein components, which is complex, costly, and low yielding (**Figure 1.2, left**)[32]. Additionally, biosynthesis of the polysaccharide component typically requires

large-scale cultivation of pathogenic bacteria, which is accompanied by biosafety regulations, high costs, and limits the development of new conjugate vaccines to bacterial targets that are amenable to large-scale fermentation under normal laboratory conditions. To address these challenges, it was recently demonstrated that polysaccharide-protein conjugates could be made in *E. coli* via protein-glycan coupling technology (PGCT)[22]. In this approach, engineered *E. coli* cells covalently attach heterologously expressed CPS or O-PS antigens to specific residues on carrier proteins via an *N*-linked glycosylation reaction (**Figure 1.2, right**). To date, this technology has yielded a handful of vaccine candidates, termed "bioconjugates" to highlight the *in vivo* production process, including those directed against important human pathogens including *Burkholderia pseudomallei*, *E. coli* O121, *E. coli* O157:H7, *Francisella tularensis*, and *Staphylococcus aureus*[33-37]. In all cases tested, the vaccines either stimulated serum bactericidal antibodies[36] or provided protection against pathogen challenge[33-35].
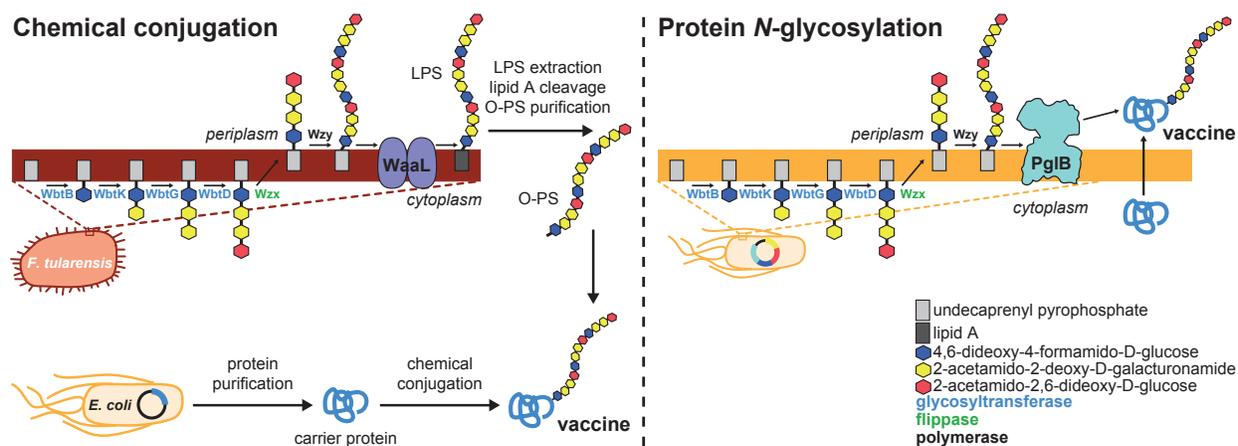
Figure 1.2 Strategies for production of conjugate and bioconjugate vaccines. Conjugate vaccine production (**left**) involves large-scale fermentations of bacterial pathogens and complex bioprocessing steps to express, isolate, and conjugate the polysaccharide and protein components. In contrast, production of bioconjugate vaccines in *E. coli* via *N*-linked glycosylation (**right**) greatly simplifies biomanufacturing. Schematic illustrates examples of chemical conjugation and *N*-linked glycosylation approaches to produce vaccines directed against the bacterial pathogen *F. tularensis*.

While advances in enabling eukaryotic or bacterial glycosylation in *E. coli* could reduce the time and cost associated with glycoprotein therapeutic and vaccine production, these approaches are still limited in their modularity and portability. Biomanufacturing in living *E. coli* is associated with biosafety

and product purity regulations, subject to lengthy *in vivo* process development timelines, and dependent upon skilled operators and specialized equipment, which altogether necessitate centralized production facilities and cold-chain distribution strategies. Refrigeration is especially critical for conjugate vaccines, as they are prone to precipitation and significant loss of the pathogen-specific carbohydrate component upon both heating and freezing due to the delicate nature of polysaccharide antigen structures[8,32,38]. A portable platform for point-of-care glycoprotein biomanufacturing could significantly enhance our ability to distribute sophisticated protein therapeutics and vaccines to both the developed and the developing world.

1.3 Cell-free systems enable rapid, portable protein production

Cell-free protein synthesis (CFPS) is an emerging field that allows for on-demand production of proteins in crude cell lysates[39,40] (**Figure 1.3**). CFPS technology was first used over 50 years ago by Nirenberg and Matthaei to decipher the genetic code[41]. In the late 1960s and early 1970s, CFPS was employed to help elucidate the regulatory mechanisms of the *E. coli* lactose[42] and tryptophan[43] operons. In the last two decades, CFPS platforms have experienced a surge in development to meet the increasing demand for recombinant protein expression technologies[44].

CFPS offers several advantages for recombinant protein expression. In particular, the open reaction environment allows for addition or removal of substrates for protein synthesis, as well as precise, on-line reaction monitoring. Additionally, the CFPS reaction environment can be wholly directed toward and optimized for production of the protein product of interest, as the use of cell lysates rather than living cells circumvents many cell viability constraints. CFPS also offers shortened protein production timelines compared to *in vivo* approaches, as proteins can be synthesized at relevant yields (~1 mg/mL) in less than a day. The *E. coli* CFPS system in particular has been widely adopted because of (i) its high batch yields, with up to 2.3 g/L of green fluorescent protein (GFP) reported[45], (ii) inexpensive required substrates[46-48], and (iii) the ability to linearly scale reaction volumes over 10^6 L[49]. Importantly, recent work demonstrated that the *E. coli* CFPS platform could be lyophilized for cold chain-independent distribution and storage. These freeze-dried cell-free (FD-CF) reactions can be activated by just adding water to produce protein therapeutics, vaccines, and diagnostics[12,13,50,51]. These factors make the *E. coli* CFPS

system an attractive platform for portable, on-demand biomanufacturing. However, production of glycoprotein therapeutics is not possible in existing systems, as *E. coli* lacks endogenous glycosylation machinery.
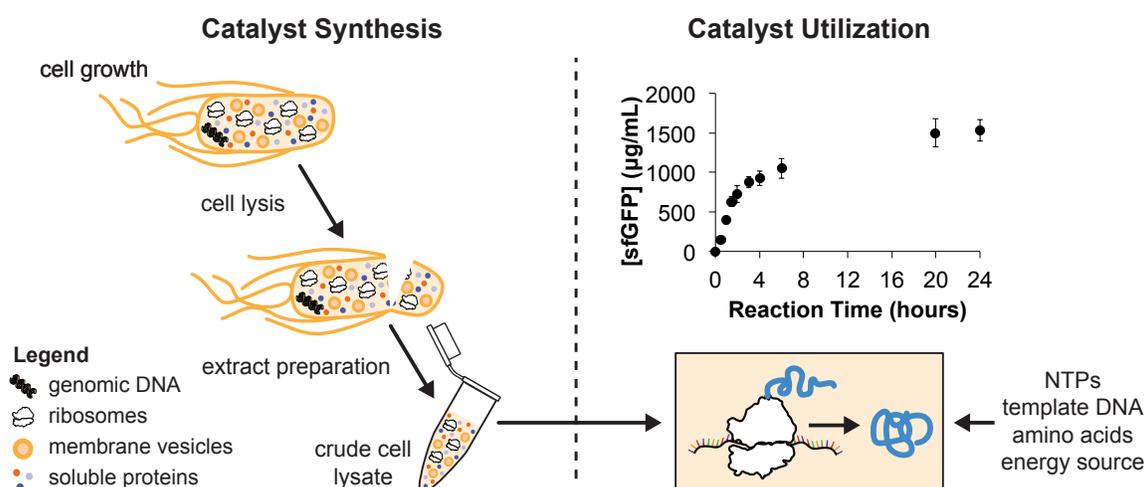
Figure 1.3 CFPS systems enable on-demand protein production. The CFPS approach uses cell lysates, rather than living cells, to synthesize proteins *in vitro*. Relevant yields (1 mg/mL or higher) of protein are produced in just a few hours, making CFPS systems attractive platforms for on-demand protein production.

In fact, CFPS systems have been limited in general by their inability to co-activate efficient protein synthesis and glycosylation. Glycosylation is possible in some eukaryotic CFPS systems, including those prepared from insect cells[52], trypanosomes[53], hybridomas[54], or mammalian cells[14,55,56]. However, these platforms are limited to endogenous machinery for performing glycosylation, meaning that (i) the possible glycan structures are restricted to those naturally synthesized by the host cells and (ii) the glycosylation process is carried out in a black box and thus difficult to engineer or control. Additionally, eukaryotic CFPS systems are technically difficult to prepare, often requiring supplementation with microsomes[57-59], and suffer from inefficient protein synthesis and glycosylation yields due to inefficient trafficking of nascent polypeptide chains to microsomes[53,59]. A glycoengineered bacterial cell-free system capable of efficient protein synthesis and glycosylation could

enable facile, portable glycoprotein production and enable glycosylation reaction optimization and engineering not possible in eukaryotic platforms.

In Chapters 2 and 3, I describe the development of the cell-free glycoprotein synthesis (CFGpS) and *in vitro* bioconjugate vaccine expression (iVAX) platforms, two new technologies that enable on-demand and modular biomanufacturing of glycosylated protein therapeutics and bioconjugate vaccines. Additionally, in Chapter 4, I describe the ability to reduce cell-free reaction cost by activating native metabolism to power protein synthesis in a yeast cell-free protein synthesis system. Combined with engineered yeast strains that can produce proteins bearing human-like glycans[60-62], this platform could provide an alternative strategy for portable production of therapeutic glycoproteins.

1.4   Portable cell-free protein synthesis systems enable hands-on biology education

Another area where portable protein synthesis technologies have great potential for impact is in facilitating hands-on biology education. Hands-on science activities have been shown to improve student engagement and performance in science classes[63], prompting the development of hands-on kits for teaching STEM subjects such as chemistry, physics, engineering, and computer science. However, compared to other STEM disciplines, implementation of hands-on molecular and synthetic biology curriculum is challenging due to the expensive equipment and expertise required to grow cells.

Synthetic biology-based educational efforts such as BioBuilder Educational Foundation, the International Genetically Engineered Machines competition, Amino Labs, and The ODIN have made great strides toward integration of cutting-edge, hands-on biology research into classrooms. However, these resources rely on cell-based experimentation, which requires (i) expensive equipment and specialized expertise to grow and engineer cells[64,65], (ii) extended instructor prep time and in-class time due to the time scales associated with cell growth[66], and (iii) compliance with biosafety regulations that can limit the ability to work with cells outside of a laboratory setting[67,68].

In contrast, FD-CF reactions circumvent many of the biosafety and biocontainment regulations that exist for living cells because they use cell lysates, rather than intact cells, to synthesize proteins. Further, FD-CF reactions eliminate the need for specialized equipment or experimental expertise as they

are shelf-stable[50] and can be activated simply by adding water and other desired inputs (*e.g.*, DNA, small molecules, enzymes) to a freeze-dried pellet of reagents. The simplified nature of reaction setup and analysis would minimize both the amount of in-class time and out-of-class instructor preparation time required to incorporate hands-on lab activities, which have been cited as limiting factors for high school biology teachers[69,70].

In Chapter 5, I describe the adaptation of freeze-dried cell-free (FD-CF) technology to enable educational kits for teaching molecular and synthetic biology, which we call BioBits[TM] kits. This work promises to facilitate the implementation of hands-on biology activities in classrooms and other non-laboratory settings.

1.5   Publishing information

Section 1.3 was adapted from the following publication:

2. Single-pot glycoprotein biosynthesis in lysates enriched with glycosylation machinery

2.1 Abstract

The emerging discipline of bacterial glycoengineering has made it possible to produce designer glycans and glycoconjugates for use as vaccines and therapeutics. Unfortunately, cell-based production of homogeneous glycoproteins remains a significant challenge due to cell viability constraints and the inability to control glycosylation components at precise ratios in vivo. To address these challenges, we describe a novel cell-free glycoprotein synthesis (CFGpS) technology that seamlessly integrates protein biosynthesis with asparagine-linked protein glycosylation. This technology leverages a glyco-optimized *Escherichia coli* strain to source cell extracts that are selectively enriched with glycosylation components, including oligosaccharyltransferases (OSTs) and lipid-linked oligosaccharides (LLOs). The resulting extracts enable a one-pot reaction scheme for efficient and site-specific glycosylation of target proteins. The CFGpS platform is highly modular, allowing the use of multiple distinct OSTs and structurally diverse LLOs. As such, we anticipate CFGpS will facilitate fundamental understanding in glycoscience and make possible applications in on demand biomanufacturing of glycoproteins.

2.2 Introduction

Asparagine-linked (*N*-linked) protein glycosylation is one of the most common post-translational modifications in eukaryotes, and profoundly affects protein properties such as folding, stability, immunogenicity, and pharmacokinetics[71-73]. The attached *N*-glycans can participate in a wide spectrum of biological processes such as immune recognition/response[74,75] and stem cell fate[76]. Moreover, the intentional engineering of protein-associated glycans can be used to manipulate protein therapeutic properties such as enhancing in vivo activity and half-life[77].

At present, however, the inherent structural complexity of glycans and the corresponding difficulties producing homogeneously glycosylated proteins have slowed advances in our understanding of glycoprotein functions and limited opportunities for biotechnological applications. Moreover, because glycan biosynthesis is neither template-driven nor genetically encoded, glycans cannot be produced from recombinant DNA technology. Instead, *N*-glycans are naturally made by coordinated expression of multiple glycosyltransferases (GTs) across several subcellular compartments. This mode of biosynthesis combined with the lack of a strict proofreading system results in inherent glycan heterogeneity and accounts for the large diversity of structures in the expressed glycan repertoire of a cell or organism[78,79]. Further complicating matters is the paucity of structure–function relationships for GTs, which hinders a priori prediction of glycan structure. Altogether, these factors have frustrated production of homogeneous glycans and glycoconjugates in biological systems and restricted our capacity to elucidate the biochemical and biophysical effects of glycans on the proteins to which they are attached. Thus, there is an unmet need for a technology capable of rapidly producing useful quantities of proteins featuring user-specified glycosylation for biochemical and structural biology studies.

Recent pioneering efforts in glycoengineering of cellular systems including mammalian[80], yeast[60], and bacterial cells[23] have expanded our ability to reliably synthesize chemically defined glycans and glycoproteins. Despite the promise of these systems, protein expression yields often remain low and design-build-test (DBT) cycles—iterations of re-engineering organisms to test new sets of enzymes—can be slow. One promising alternative to cell-based systems is cell-free protein synthesis (CFPS) in which protein synthesis occurs *in vitro* without using intact, living cells. Recently, a technical renaissance has revitalized CFPS systems to help meet increasing demands for simple and efficient protein synthesis, with *Escherichia coli*-based CFPS systems now exceeding grams of protein per liter reaction volume[39], with the ability to support co- or post-translational modifications[81-85]. As a complement to in vivo expression systems, cell-free systems offer several potential advantages. First, the open nature of the reaction allows the user to directly influence biochemical systems of interest. As a result, new components can be added or synthesized, and maintained at precise

concentrations[86,87]. Second, cell-free systems bypass viability constraints making possible the production of proteins at titers that would otherwise be toxic in living cells[88]. Third, processes that take days or weeks to design, prepare, and execute in vivo can be done more rapidly in a cell-free system[89,90], leading to high-throughput production campaigns on a whole-proteome scale[91] with the ability to automate[92].

Unfortunately, CFPS systems have been limited by their inability to co-activate efficient protein synthesis and glycosylation. The best characterized and most widely adopted CFPS systems use *E. coli* lysates to activate *in vitro* protein synthesis, but these systems are incapable of making glycoproteins because *E. coli* lacks endogenous glycosylation machinery. Glycosylation is possible in some eukaryotic CFPS systems, including those prepared from insect cells[52], trypanosomes[53], hybridomas[54], or mammalian cells[55-57]. However, these platforms are limited to endogenous machinery for performing glycosylation, meaning that (i) the possible glycan structures are restricted to those naturally synthesized by the host cells and (ii) the glycosylation process is carried out in a black box and thus difficult to engineer or control. Additionally, eukaryotic CFPS systems are technically difficult to prepare, often requiring supplementation with microsomes[57-59], and suffer from inefficient protein synthesis and glycosylation yields due to inefficient trafficking of nascent polypeptide chains to microsomes[53,59].

Despite progress in eukaryotic cell-free systems, cell-free extracts from bacteria like *E. coli* offer a blank canvas for studying glycosylation pathways, provided they can be activated *in vitro*. A recent work from our group highlights the ability of CFPS to enable glycoprotein synthesis in bacterial cell-free systems by augmenting commercial *E. coli*-based cell-free translation systems with purified components from a bacterial *N*-linked glycosylation pathway[93]. While these results established the possibility of *E. coli* lysate-based glycoprotein production, there are several drawbacks of using purified glycosylation components that limit system utility. First, preparation of the glycosylation components required time-consuming and cost-prohibitive steps, namely purification of a multipass transmembrane oligosaccharyltransferase (OST) enzyme and organic solvent-based extraction of lipid-linked oligosaccharide (LLO) donors from bacterial membranes. These steps significantly

lengthen the process development timeline, requiring 3–5 days each for preparation of the LLO and OST components, necessitate skilled operators and specialized equipment, and result in products that must be refrigerated and are stable for only a few months to a year. Second, glycoproteins were produced using a sequential translation/glycosylation strategy, which required 20 h for cell-free synthesis of the glycoprotein target and an additional 12 h for post-translational protein glycosylation.

Here, we addressed these drawbacks by developing an integrated cell-free glycoprotein synthesis (CFGpS) technology that bypasses the need for purification of OSTs and organic solvent-based extraction of LLOs. The creation of this streamlined CFGpS system was made possible by two important discoveries: (i) crude extract prepared from the glyco-optimized *E. coli* strain, CLM24, is able to support cell-free protein expression and *N*-linked glycosylation; and (ii) OST- and LLO-enriched extracts derived from CLM24 are able to reproducibly co-activate protein synthesis and *N*-glycosylation in a reaction mixture that minimally requires priming with DNA encoding the target glycoprotein of interest. Importantly, the CFGpS system decouples production of glycoprotein synthesis components (i.e., OSTs, LLOs, translation machinery) and the glycoprotein target of interest, providing significantly reduced cell viability constraints compared to in vivo systems. The net result is a one-pot bacterial glycoprotein biosynthesis platform whereby different acceptor proteins, OSTs, and/or oligosaccharide structures can be functionally interchanged and prototyped for customizable glycosylation.

2.3   Methods

2.3.1 Bacterial strains and plasmids

The following *E. coli* strains were used in this study: DH5α, BL21(DE3) (Novagen), CLM24, and Origami2(DE3) *gmd*::kan Δ*waaL*. DH5α was used for plasmid cloning and purification. BL21(DE3) was used for expression and purification of the scFv13-R4[DQNAT] acceptor protein that was used in all *in vitro* glycosylation reactions. CLM24 is a glyco-optimized derivative of W3110 that carries a deletion in the gene encoding the WaaL ligase, thus facilitating the accumulation of preassembled glycans on Und-PP[22]. CLM24 was used for purification of the *Cj*OST enzyme, organic solvent-based extraction of all LLOs bearing bacterial glycans, and the source strain for preparing extracts with and without

selectively enriched glycosylation components. Origami2(DE3) *gmd*::kan Δ*waaL* was used for producing Man$_3$GlcNAc$_2$-bearing LLOs and was generated by sequential mutation with P1*vir* phage transduction using the respective strains from the Keio collection[94] as donors, which were obtained from the Coli Genetic Stock Center (CGSC). In brief, donor lysate was generated from strain JW3597-1 (Δ*rfaL734*::kan) and the resulting phage was used to infect Origami2(DE3) target cells. After plating transformants on LB plates containing kanamycin (Kan), successful transductants were selected and their Kan resistance cassettes were removed by transforming with temperature-sensitive plasmid pCP20[95]. The resulting strain, Origami2(DE3) Δ*waaL*, was then used for subsequent deletion of the *gmd* gene according to an identical strategy but using donor strain JW2038-1 (Δ*gmd751*::kan).

All plasmids used in the study are listed in **Table 2.1**. Plasmids constructed in this study were made using standard cloning protocols and confirmed by DNA sequencing. These included the following. Plasmid pJL1-scFv13-R4[DQNAT] was generated by first PCR amplifying the gene encoding scFv13-R4[DQNAT] from pET28a-scFv13-R4(N34L, N77L)[DQNAT], where the N34L and N77L mutations were introduced to eliminate putative internal glycosylation sites in scFv13-R4[96]. The resulting PCR product was then ligated between *Nco*I and *Sal*I restriction sites in plasmid pJL1, a pET-based vector used for CFPS[97]. Plasmid pJL1-sfGFP[217-DQNAT] was generated by ligating a commercially-synthesized DNA fragment encoding sfGFP[217-DQNAT] (Integrated DNA Technologies) into pJL1. This version of sfGFP contains an additional GT insertion after K214, which extends this flexible loop before the final beta sheet[98]. Into this flexible loop, immediately after T216, we grafted a 21-amino acid sequence containing the *C. jejuni* AcrA N123 glycosylation site[93], but with an optimal DQNAT sequon in place of the native AcrA sequon. Similar procedures were used to generate plasmids pJL1-sfGFP[217-AQNAT], pJL1-hEPO[22-DQNAT-26], pJL1-hEPO[36-DQNAT-40], and pJL1-hEPO[81-DQNAT-85]. In the case of pJL1-hEPO[22-DQNAT-26], the gene for mature human EPO was designed such that the native sequon at N24 was changed from 22-AENIT-26 to an optimal bacterial sequon, DQNAT. Identical cloning strategies were carried out to separately introduce optimal DQNAT motifs in place of the native hEPO sequons 36-NENIT-40 and 81-LVNSS-85. Recombinant expression of the *E. coli* O9 primer-adaptor glycan (Man$_3$GlcNAc) on Und-PP was achieved by cloning the genes encoding the WbdB and WbdC

mannosyltransferase enzymes derived from *E. coli* ATCC31616 for assembling the glycan, and RfbK and RfbM, also derived from *E. coli* ATCC31616 for increasing the pool of available GDP-mannose, in *E. coli* MG1655. Plasmid pConYCGmCB was constructed by isothermal Gibson assembly and encodes an artificial operon comprised of: (i) the yeast glycosyltransferases Alg13, Alg14, Alg1, and Alg2 for $Man_3GlcNAc_2$ glycan biosynthesis[23] and (ii) the *E. coli* enzymes phosphomannomutase (ManB) and mannose-1-phosphate guanylyltransferase (ManC), which together increase availability of GDP-mannose substrates for the Alg1 and Alg2 enzymes.

2.3.2 Protein expression and purification

Purification of *Cj*PglB was performed according to a previously described protocol[93]. Briefly, a single colony of *E. coli* CLM24 carrying plasmid pSN18[99] was grown overnight at 37 °C in 50 mL of Luria-Bertani (LB; 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl, pH 7.2) supplemented with ampicillin (Amp) and 0.2% (w/v%) D-glucose. Overnight cells were subcultured into 1 L of fresh terrific broth (TB; 12 g L^{-1} tryptone, 24 g L^{-1} yeast extract, 0.4% (v/v%) glycerol, 10% (v/v%) 0.17 M KH_2PO_4/0.72 M K_2HPO_4 phosphate buffer), supplemented with Amp and grown until the absorbance at 600 nm (Abs_{600}) reached a value of ~0.7. The incubation temperature was adjusted to 16 °C, after which protein expression was induced by the addition of L-arabinose to a final concentration of 0.02% (w/v%). Protein expression was allowed to proceed for 20 h at 16 °C. Cells were harvested by centrifugation and then disrupted using a homogenizer (Avestin C5 EmulsiFlex). The lysate was centrifuged to remove cell debris and the supernatant was ultracentrifuged (100,000×*g*) for 2 h at 4 °C. The resulting pellet containing the membrane fraction was fully resuspended with a Potter-Elvehjem tissue homogenizer in buffer containing 50 mM HEPES, 250 mM NaCl, 10% (v/v%) glycerol, and 1% (w/v%) *n*-dodecyl-β-D-maltoside (DDM) at pH 7.5. The suspension was incubated at room temperature for 1 h to facilitate detergent solubilization of *Cj*PglB from native *E. coli* lipids, which were removed by subsequent ultracentrifugation (100,000×*g*) for 1 h at 4 °C. The supernatant containing DDM-solubilized *Cj*PglB was purified using Ni-NTA resin (Thermo) according to manufacturer's specification with the exception that all buffers were supplemented with 1% (w/v%) DDM. The elution fraction from Ni-NTA purification was then subjected to size exclusion chromatography (SEC) using an

ÄKTA Explorer FPLC system (GE Healthcare) with Superdex 200 10/300 GL column. Purified protein was stored at a final concentration of 1–2 mg mL^{-1} in OST storage buffer (50 mM HEPES, 100 mM NaCl, 5% (v/v%) glycerol, 0.01% (w/v%) DDM, pH 7.5) at 4 °C. Glycerol concentration in the sample was adjusted to 20% (v/v%) for long-term storage at −80 °C.

Purification of acceptor protein scFv13-R4^{DQNAT} was carried out as described previously[96]. Briefly, *E. coli* strain BL21(DE3) carrying plasmid pET28a-scFv13-R4(N34L, N77L)^{DQNAT} was grown in 1 L of TB supplied with kanamycin. The culture was incubated at 37°C until Abs$_{600}$ reached ~0.7, at which point protein expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. Protein expression was allowed to proceed for 20 h at 25°C. Cells were harvested and disrupted identically as described above. The scFv13-R4^{DQNAT} protein was purified using Ni-NTA resin followed by SEC according to manufacturer's protocols. Protein was stored at a final concentration of 1–2 mg mL^{-1} in storage buffer (50 mM HEPES, 250 mM NaCl, 1 mM EDTA, pH 7.5) at 4 °C.

2.3.3 Extraction of LLOs

The protocol for organic solvent extraction of LLOs from *E. coli* membranes was adapted from a previously described protocol[93,100]. In most cases, a single colony of strain CLM24 carrying a plasmid for target glycan biosynthesis (**Table 2.1**) was grown overnight in LB media. The notable exceptions were LLOs bearing the *W. succinogenes N*-glycan (*Ws*LLOs), which were produced using DH5α cells carrying the pEpiFOS-5*pgl5* fosmid (kindly provided by Dr. Markus Aebi), and LLO sbearing Man$_3$GlcNAc$_2$, which were produced using Origami2(DE3) *gmd*::kan Δ*waaL* cells carrying plasmid pConYCGmCB. Overnight cells were subcultured into 1 L of TB supplemented with an appropriate antibiotic and grown until the Abs$_{600}$ reached ~0.7. The incubation temperature was adjusted to 30 °C for biosynthesis of all glycans except for Man$_3$GlcNAc$_2$, which was adjusted to 16 °C. For plasmid pMW07-pglΔB, protein expression was induced with L-arabinose at a final concentration of 0.2% (w/v%) while for fosmid pEpiFOS-5*pgl5* induction was with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 1 mM. All other plasmids involved constitutive promoters and thus did not require chemical inducers. After 16 h, cells were harvested by

centrifugation and cell pellets were lyophilized to complete dryness at −70 °C. For extraction of CjLLOs, native and engineered ClLLOs, $E.$ $coli$ O9 primer-adaptor LLOs, and WsLLOs, the lyophilisates were suspended in 10:20:3 volumetric ratio of $CHCl_3:CH_3OH:H_2O$ solution and incubated at room temperature for 15 min to facilitate extraction of LLOs. For extraction of LLOs bearing $Man_3GlcNAc_2$ glycan, lyophilisate was successively suspended in 10:20 (v/v%) $CHCl_3:CH_3OH$ solution, water, and 10:20:3 $CHCl_3:CH_3OH:H_2O$ solution with 15 min of incubation at room temperature between each step. In each case, the final suspension was centrifuged (4000×g) for 15 min, after which the organic layer (bottom layer) was collected and dried with a vacuum concentrator followed by lyophilization. Lyophilisates containing active LLOs were resuspended in cell-free glycosylation buffer (10 mM HEPES, pH 7.5, 10 mM $MnCl_2$, and 0.1% (w/v%) DDM) and stored at 4 °C.

2.3.4 Preparation of crude S30 extracts

CLM24 source strains were grown in 2×YTPG (10 g L^{-1} yeast extract, 16 g L^{-1} tryptone, 5 g L^{-1} NaCl, 7 g L^{-1} K_2HPO_4, 3 g L^{-1} KH_2PO_4, 18 g L^{-1}glucose, pH 7.2) until the Abs_{600} reached ~3. To generate OST-enriched extract, CLM24 carrying plasmid pSF-CjPglB, pSF-CcPglB, pSF-DdPglB, pSF-DgPglB, or pSF-DvPglB[96] was used as the source strain. To generate LLO-enriched extract, CLM24 carrying plasmid pMW07-pglΔB was used as the source strain. To generate one-pot extract containing both OST and LLOs, CLM24 carrying pMW07-pglΔB and pSF-CjOST was used as the source strain. As needed, the expression of glycosylation components was induced with L-arabinose at final concentration of 0.02% (w/v%). After induction, protein expression was allowed to proceed at 30 °C to a density of OD_{600} ~3, at which point cells were harvested by centrifugation (5000×g) at 4 °C for 15 min. All subsequent steps were carried out at 4 °C unless otherwise stated. Pelleted cells were washed three times in S30 buffer (10 mM tris acetate, 14 mM magnesium acetate, 60 mM potassium acetate, pH 8.2). After the last wash, cells were pelleted at 7000×g for 10 min and flash-frozen on liquid nitrogen. To make lysate, cells were thawed and resuspended to homogeneity in 1 mL of S30 buffer per 1 g of wet cell mass. Cells were disrupted using an Avestin EmulsiFlex-B15 high-pressure homogenizer at 20,000–25,000 psi with a single passage. The lysate was then centrifuged twice at 30,000×g for 30 min to remove cell debris. Supernatant was transferred to a new vessel and incubated

with 250 rpm shaking at 37 °C for 60 min to degrade endogenous mRNA transcripts and disrupt existing polysome complexes in the lysate. Following centrifugation (15,000×g) for 15 min at 4 °C, supernatant was collected, aliquoted, flash-frozen in liquid nitrogen, and stored at −80 °C. S30 extract was active for about three freeze-thaw cycles and contained ~40 g L^{-1} total protein as measured by Bradford assay.

2.3.5 Cell-free glycoprotein synthesis

For *in vitro* glycosylation of purified acceptor protein, reactions were carried out in a 50 μL volume containing 3 μg of scFv13-R4^{DQNAT}, 2 μg of purified CjPglB, and 5 μg extracted LLOs (in the case of Man$_3$GlcNAc$_2$LLOs, 20 μg was used) in *in vitro* glycosylation buffer (10 mM HEPES, pH 7.5, 10 mM MnCl$_2$, and 0.1% (w/v%) DDM). The reaction mixture was incubated at 30 °C for 16 h. For crude extract-based expression of glycoproteins, a two-phase scheme was implemented. In the first phase, protein synthesis was carried out with a modified PANOx-SP system[101]. Specifically, 1.5 mL microcentrifuge tubes were charged with 15-μL reactions containing 200 ng plasmid DNA, 30% (v/v%) S30 extract and the following: 12 mM magnesium glutamate, 10 mM ammonium glutamate, 130 mM potassium glutamate, 1.2 mM adenosine triphosphate (ATP), 0.85 mM guanosine triphosphate (GTP), 0.85 mM uridine triphosphate (UTP), 0.85 mM cytidine triphosphate (CTP), 0.034 mg mL^{-1} folinic acid, 0.171 mg mL^{-1} *E. coli* tRNA (Roche), 2 mM each of 20 amino acids, 30 mM phosphoenolpyruvate (PEP, Roche), 0.4 mM nicotinamide adenine dinucleotide (NAD), 0.27 mM coenzyme-A (CoA), 4 mM oxalic acid, 1 mM putrescine, 1.5 mM spermidine, and 57 mM HEPES. For scFv13-R4^{DQNAT}, hEPO$^{22-DQNAT-26}$, hEPO$^{36-DQNAT-40}$, and hEPO$^{81-DQNAT85}$, this phase was carried out at 30 °C for 4 h under oxidizing conditions while for sfGFP$^{217-DQNAT}$ and sfGFP$^{217-AQNAT}$ this phase was carried out at 30 °C for 5 min under reducing conditions. For oxidizing conditions, extract was pre-conditioned with 750 μM iodoacetamide in the dark at room temperature for 30 min and the reaction mix was supplied with 200 mM glutathione at a 3:1 ratio between oxidized and reduced forms. The active sfGFP yields from cell-free reactions were quantified by measuring fluorescence in-lysate and converting into concentration using a standard curve as previously described[102]. In the second phase, protein glycosylation was initiated by the addition of MnCl$_2$ and DDM at a final concentration of 10 mM and

0.1% (w/v%), respectively, and allowed to proceed at 30 °C for 16 h. As needed, reactions were supplemented with 2 µg of purified CjPglB (i.e., for CFGpS with LLO-enriched extracts) or 5 µg solvent-extracted CjLLOs (i.e., for CFGpS with OST-enriched extracts). All reactions were stopped by adding Laemmli sample buffer containing 5% βME, after which samples were boiled at 100 °C for 15 min and analyzed by SDS-PAGE and Western blotting.

2.3.6 Western blot analysis

Samples containing 0.5 µg of acceptor protein were loaded into SDS-PAGE gels. Following electrophoretic separation, proteins were transferred from gels onto Immobilon-P polyvinylidene difluoride (PVDF) membranes (0.45 µm) according to manufacturer's protocol. Membranes were washed twice with TBS buffer (80 g L^{-1} NaCl, 20 g L^{-1}KCl, and 30 g L^{-1} Tris-base) followed by incubation for 1 h in blocking solution (50 g L^{-1} non-fat milk in TBST (TBS with 0.05% (v/v%) Tween-20). After blocking, membranes were washed 4 times with TBST with 10 min incubation between each wash. A first membrane was probed with 6xHis-polyclonal antibody (Abcam, ab137839, 1:7500) that specifically recognizes hexahistidine epitope tags while a second replicate membrane was probed with one of the following: hR6 (1:10,000) serum from rabbit that recognizes the native *C. jejuni* and *C. lari* glycan as well as engineered *C. lari* glycan or ConA-HRP (Sigma, L6397, 1:2500) that recognizes Man$_3$GlcNac and Man$_3$GlcNAc$_2$. Probing of membranes was performed for at least 1 h with shaking at room temperature, after which membranes were washed with TBST in the same manner as described above. For development, membranes were incubated briefly at room temperature with Western ECL substrate (BioRad) and imaged using a ChemiDocTM XRS+System. OST enzymes enriched in extracts were detected by an identical SDS-PAGE procedure followed by Western blot analysis with a polyclonal antibody specific to the FLAG epitope tag (Abcam, ab49763, 1:7500). The glycan component of LLOs enriched in extracts was detected by directly spotting 10 µL of extracts onto nitrocellulose membranes followed by detection with hR6 serum.

2.3.7 Mass spectrometry analysis

Approximately 2 µg of scFv13-R4^{DQNAT} protein in solution was denatured with 6 M urea, reduced with 10 mM DTT, incubated at 34 °C for 1 h, then alkylated with 58 mM iodoacetamide for 45 min in the dark at room temperature and quenched by final 36 mM DTT. The solution was then diluted with 50 mM ammonium bicarbonate (pH 8.0) to a final buffer concentration of 1 M urea prior to trypsin digestion. Sample was digested with 0.2 µg of trypsin for 18 h at 37 °C. The digestion was stopped by addition of TFA to a final pH 2.2–2.5. The samples were then desalted with SOLA HRP SPE Cartridge (ThermoFisher Scientific). The cartridges were conditioned with 1 × 0.5 mL 90% methanol, 0.1% trifluoroacetic acid (TFA) and equilibrated with 2 × 0.5 mL 0.1% (v/v%) TFA. The samples were diluted 1:1 with 0.2% (v/v%) TFA and run slowly through the cartridges. After washing with 2 × 0.5 mL of equilibration solution, peptides were eluted by 1 × 0.5 mL of 50% (v/v%) acetonitrile (ACN), 0.1% (v/v%) TFA and dried in a speed vacuum centrifuge.

The nanoLC–MS/MS analysis was carried out using UltiMate3000 RSLCnano (Dionex) coupled to an Orbitrap Fusion (ThermoFisher Scientific) mass spectrometer equipped with a nanospray Flex Ion Source. Each sample was reconstituted in 22 µL of 0.5% (w/v%) FA and 10 µL was loaded onto an Acclaim PepMap 100 C18 trap column (5 µm, 100 µm × 20 mm, 100 Å, ThermoFisher Scientific) with nanoViper Fittings at 20 µL min^{-1} of 0.5% FA for on-line desalting. After 2 min, the valve switched to allow peptides to be separated on an Acclaim PepMap C18 nano column (3 µm, 75 µm × 25 cm, ThermoFisher Scientific), in a 90 min gradient of 5 to 23% to 35% B at 300 nL min^{-1} (3 to 73 to 93 min, respectively), followed by a 9-min ramping to 90% B, a 9-min hold at 90% B and quick switch to 5% B in 1 min. The column was re-equilibrated with 5% B for 20 min prior to the next run. The Orbitrap Fusion was operating in positive ion mode with nanospray voltage set at 1.7 kV and source temperature at 275 °C. External calibration for FT, IT and quadrupole mass analyzers was performed prior to the analysis. The Orbitrap full MS survey scan (*m/z* 400–1800) was followed by Top 3 s data-dependent Higher Collision dissociation product ion triggered ETD (HCD-pd-ETD) MS/MS scans for precursor peptides with 2–7 charges above a threshold ion count of 50,000 with normalized collision energy of 32%. MS survey scans were acquired at a resolving power of 120,000 (FWHM

at m/z 200), with Automatic Gin Control (AGC) = 2e5 and maximum injection time (Max IT) = 50 ms, and HCD MS/MS scans at a resolution of 30,000 with AGC = 5e4, Max IT = 60 ms and with Q isolation window (m/z) at 3 for the mass range m/z 105–2000. Dynamic exclusion parameters were set at 1 within 60 s exclusion duration with ±10 ppm exclusion mass width. Product Ion trigger list consisted of peaks at 204.0867 Da (HexNAc oxonium ion), 138.0545 Da (HexNAc fragment), and 366.1396 Da (HexHexNAc oxonium ions). If one of the HCD product ions in the list was detected, two charge-dependent ETD MS/MS scans (EThcD) with HCD supplemental activation (SA) on the same precursor ion were triggered and collected in a linear ion trap. For doubly charged precursors, the ETD reaction time as set 150 ms and the SA energy was set at 30%, while the same parameters at 125 ms and 20%, respectively, were used for higher charged precursors. For both ion triggered scans, fluoranthene ETD reagent target was set at 2e5, AGC target at 1e4, Max IT at 105 ms and isolation window at 3. All data were acquired using Xcalibur 3.0 operation software and Orbitrap Fusion Tune Application v. 2.1 (ThermoFisher Scientific).

All MS and MS/MS raw spectra from each sample were searched using Byonics v. 2.8.2 (Protein Metrics) using the *E coli* protein database with added scFv13-R4[DQNAT] protein target sequence. The peptide search parameters were as follows: two missed cleavage for full trypsin digestion with fixed carbamidomethyl modification of cysteine, variable modifications of methionine oxidation, and deamidation on asparagine/glutamine residues. The peptide mass tolerance was 10 ppm and fragment mass tolerance values for HCD and EThcD spectra were 0.05 and 0.6 Da, respectively. Both the maximum number of common and rare modifications were set at two. The glycan search was performed against a list of 309 mammalian *N*-linked glycans in Byonic software. Identified peptides were filtered for maximum 2% FDR. The software exported the results of the search to a spreadsheet.

2.3.8 GFP fluorescence activity

The activity of cell-free-derived sfGFP was determined using an in-lysate fluorescence analysis as described previously[102]. Briefly, 2 µL of cell-free synthesized glycosylated sfGFP reaction was diluted into 48 µL of nanopure water. The solution was then placed in a Costar 96-well black

assay plate (Corning). Excitation and emission wavelengths for sfGFP fluorescence were 485 and 528 nm, respectively.

2.3.9 Enzyme-linked immunosorbent assay (ELISA)

Costar 96-well ELISA plates (Corning) were coated overnight at 4 °C with 50 µl of 1 mg mL^{-1} *E. coli* β-gal (Sigma-Aldrich) in 0.05 M sodium carbonate buffer (pH 9.6). After blocking with 5% (w/v%) bovine serum albumin (BSA) in PBS for 3 h at room temperature, the plates were washed four times with PBST buffer (PBS, 0.05% (v/v%) Tween-20, 0.3% (w/v%) BSA) and incubated with serially diluted purified scFv13-R4 samples or soluble fractions of CFGpS lysates for 1 h at room temperature. Samples were quantified by the Bradford assay and an equivalent amount of total protein was applied to the plate. After washing four times with the same buffer, anti-6×-His-HRP conjugated rabbit polyclonal antibody (Abcam) in 3% PBST was added to each well for 1 h. Plates were washed and developed using standard protocols.

2.3.10   *In vitro* cell proliferation assay

Human erythroleukemia TF-1 cells (Sigma) that require granulocyte–macrophage colony-stimulating factor (GM-CSF), interleukin 3 (IL-3), or hEPO for growth and survival were used. Cells were maintained in RPMI-1640 media supplemented with 10% FBS, 50 U mL^{-1} penicillin, 50 mg mL^{-1} streptomycin, 2 mM glutamine, and 2 ng mL^{-1} GM-CSF at 37 °C in a humidified atmosphere containing 5% CO_2. After 16 h incubation in RPMI-1640 media without GM-CSF, cells were counted, harvested, and resuspended in fresh media. 5×10^3 TF-1 cells per well were seeded in a 96-well assay plate, and EPO standards or samples were added to final desired concentrations to each well. Cells were incubated with for 6 h in humid incubator before adding alamarBlue®. After 12 h, fluorescence signal was measured at 560 nm/590 nm excitation/emission wavelength.

2.4 Results

2.4.1 Efficient glycoprotein synthesis in extracts from a glyco-optimized chassis strain

To develop a one-pot glycoprotein synthesis system, the bacterial protein glycosylation locus (*pgl*) present in the genome of the Gram-negative bacterium *Campylobacter jejuni* was chosen as a

model glycosylation system (**Figure 2.1**). This gene cluster encodes an asparagine-linked (*N*-linked) glycosylation pathway that is functionally similar to that of eukaryotes and archaea[103], involving a single-subunit OST, PglB, that catalyzes the en bloc transfer of a preassembled 1.4 kDa GlcGalNAc$_5$Bac heptasaccharide (where Bac is bacillosamine) from the lipid carrier undecaprenyl pyrophosphate (Und-PP) onto asparagine residues in a conserved motif (D/E-X$_{-1}$-N-X$_{+1}$-S/T, where X$_{-1}$ and X$_{+1}$ are any residues except proline) within acceptor proteins. PglB was selected because we previously showed that *N*-glycosylated acceptor proteins were reliably produced when cell-free translation kits were supplemented with (i) *C. jejuni* PglB (*Cj*PglB) purified from *E. coli* cells and (ii) LLOs extracted from glycoengineered *E. coli* cells expressing the enzymes for producing the *C. jejuni* *N*-glycan on Und-PP (*Cj*LLOs)[93]. Additionally, PglB has been used in engineered *E. coli* for transferring eukaryotic trimannosyl chitobiose glycans (mannose$_3$-*N*-acetylglucosamine$_2$, Man$_3$GlcNAc$_2$) to specific asparagine residues in target proteins[23].
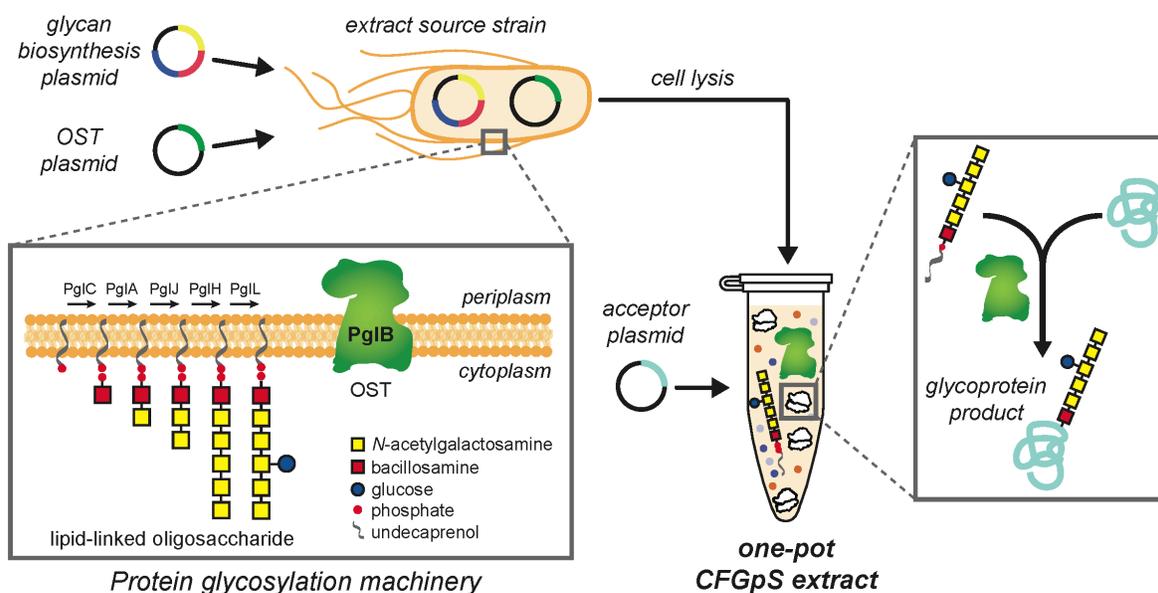
Figure 2.1. Schematic of single-pot CFGpS technology. Glycoengineered *E. coli* that are modified with (i) genomic mutations that benefit glycosylation reactions and (ii) plasmid DNA for producing essential glycosylation components (i.e., OSTs, LLOs) serve as the source strain for producing crude S30 extracts. Candidate glycosylation components can be derived from all kingdoms of life and include single-subunit OSTs like *C. jejuni* PglB and LLOs bearing *N*-glycans from *C. jejuni* that are assembled on Und-PP by the Pgl pathway enzymes. Following extract preparation by lysis of the source strain, one-pot biosynthesis of *N*-glycoproteins is initiated by priming the extract with DNA encoding the acceptor protein of interest.

Establishing a CFGpS system first required crude cell extracts suitable for glycoprotein synthesis; hence, we selected *E. coli* strain CLM24 that was previously optimized for in vivo protein glycosylation[22]. CLM24 has two attributes that we hypothesized would positively affect cell-free protein glycosylation. First, CLM24 does not synthesize *O*-polysaccharide antigen due to an inactivating insertion in w*bbL*, which encodes a rhamnosyl transferase that transfers the second sugar of the O16 subunit to Und-PP[104]. Thus, absence of WbbL should allow uninterrupted assembly of engineered glycans, such as the *C. jejuni* heptasaccharide, on Und-PP. Second, CLM24 cells lack the *waaL* gene, which encodes the ligase that transfers *O*-polysaccharide antigens from Und-PP to lipid A-core. Because WaaL can also promiscuously transfer engineered glycans that are assembled on Und-PP[23,105], the absence of this enzyme should favor accumulation of target glycans on Und-PP.

To determine whether CLM24 could be used as a chassis strain to support integrated cell-free transcription, translation, and glycosylation, we first prepared crude S30 extract from these cells using a rapid and robust procedure for extract preparation based on sonication[102]. Then, 15-µL batch-mode, sequential CFGpS reactions were performed using CLM24 crude extract that was supplemented with the following: (i) an OST catalyst in the form of purified *Cj*PglB[93]; (ii) oligosaccharide donor in the form of *Cj*LLOs that were isolated by organic solvent extraction from the membrane fraction of glycoengineered *E. coli* cells[93]; and (iii) plasmid DNA encoding the model acceptor protein scFv13-R4^{DQNAT}, an anti-β-galactosidase (β-gal) single-chain variable fragment (scFv) antibody modified C-terminally with a single DQNAT motif[23]. The glycosylation status of scFv13-R4^{DQNAT}was analyzed by SDS-PAGE and immunoblotting with an anti-polyhistidine (anti-His) antibody or hR6 serum that is specific for the *C. jejuni* heptasaccharide glycan[106]. Following an overnight reaction at 30 °C, highly efficient glycosylation was achieved as evidenced by the mobility shift of scFv13-R4^{DQNAT} entirely to the mono-glycosylated (g1) form in anti-His immunoblots and the detection of the *C. jejuni* glycan attached to scFv13-R4^{DQNAT} by hR6 serum (**Figure 2.2a**). For synthesis of scFv13-R4^{DQNAT}, the reaction mixture was modified to be oxidizing, through the addition of iodoacetamide and a 3:1 ratio of oxidized and reduced glutathione, demonstrating the flexibility of CFGpS reaction conditions for producing eukaryotic glycoprotein targets. The efficiency achieved in this CFGpS system rivaled that

of an *in vitro* glycosylation reaction in which the scFv13-R4^{DQNAT} acceptor protein was expressed and purified from *E. coli*, and then incubated overnight with purified *Cj*PglB and extracted *Cj*LLOs (**Figure 2.2a**). As expected, when *Cj*PglB was omitted from the reaction, the scFv13-R4^{DQNAT} acceptor protein was produced only in the aglycosylated (g0) form. The results generated here with CLM24 extract are consistent with our earlier studies using an *E. coli* S30 extract-based CFPS system or purified translation machinery[93], and establish that the *C. jejuni N*-linked protein glycosylation mechanism can be functionally reconstituted outside the cell.
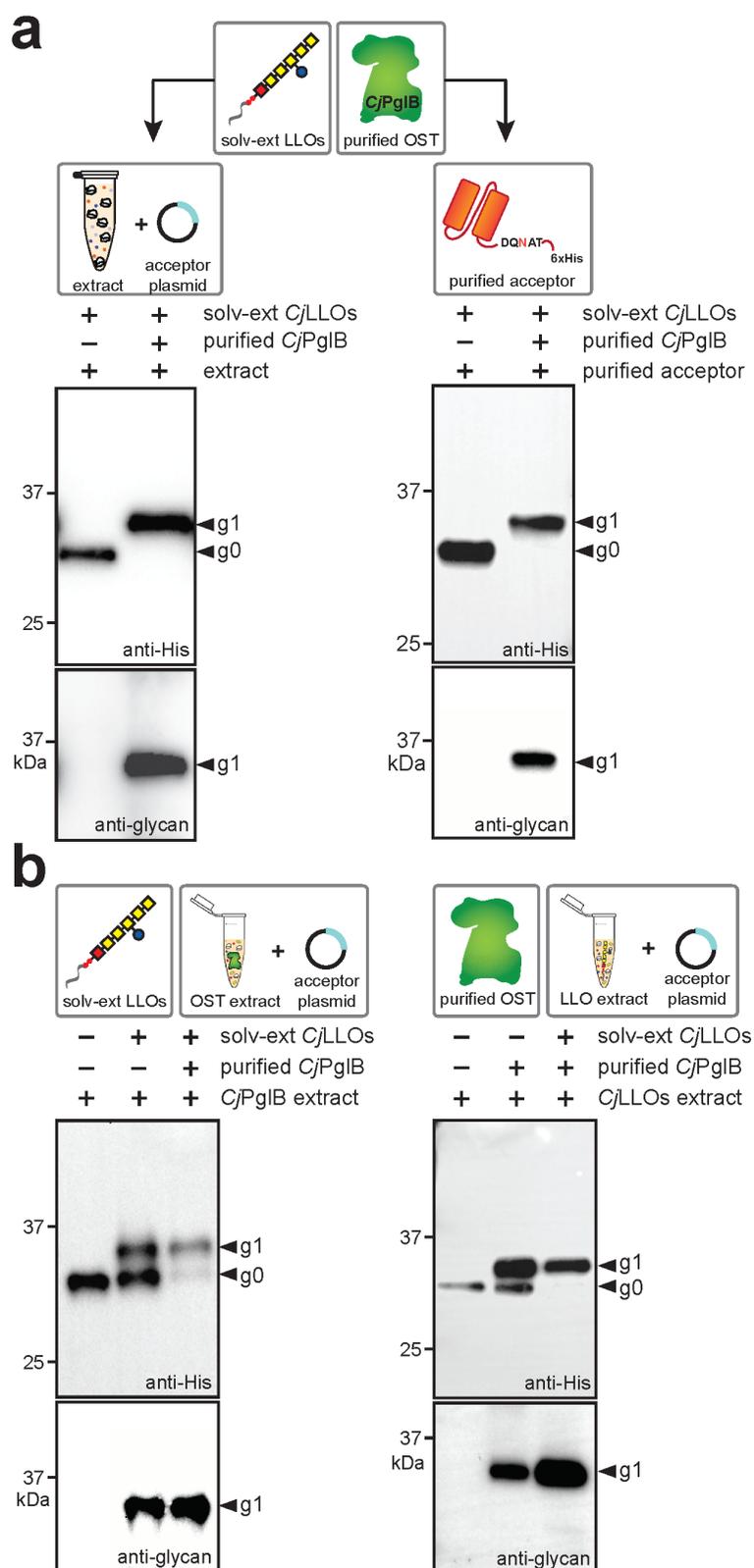
Figure 2.2 Extract from glyco-optimized chassis strain supports CFGpS. (**a**) (**left**) Western blot analysis of scFv13-R4^{DQNAT} produced by crude CLM24 extract supplemented with purified CjPglB and

organic solvent-extracted (solv-ext) *Cj*LLOs, and primed with plasmid pJL1-scFv13-R4[DQNAT]. (**right**) Western blot analysis of *in vitro* glycosylation reaction using purified scFv13-R4[DQNAT]acceptor protein that was incubated with purified *Cj*PglB and organic solvent-extracted (solv-ext) *Cj*LLOs. Control reactions (lane 1 in each panel) were performed by omitting purified *Cj*PglB. (**b**) (**left**) Western blot analysis of scFv13-R4[DQNAT]produced by crude CLM24 extract selectively enriched with *Cj*PglB from heterologous overexpression from pSF-*Cj*PglB. (**right**) Western blot analysis of scFv13-R4[DQNAT] produced by crude CLM24 extract selectively enriched with *Cj*LLOs from heterologous overexpression from pMW07-pglΔB. Reactions were primed with plasmid pJL1-scFv13-R4[DQNAT] and supplemented with purified *Cj*PglB and organic solvent-extracted (solv-ext) *Cj*LLOs as indicated. Control reactions (lane 1 in each panel) were performed by omitting solv-ext *Cj*LLOs in (**left**) or purified *Cj*PglB (**right**) in (**b**). Blots were probed with anti-hexa-histidine antibody (anti-His) to detect the acceptor protein or hR6 serum (anti-glycan) to detect the *N*-glycan. Arrows denote aglycosylated (g0) and singly glycosylated (g1) forms of scFv13-R4[DQNAT]. Molecular weight (MW) markers are indicated at left. Results are representative of at least three biological replicates.

2.4.2 Expanding the glycan repertoire of cell-free glycosylation

To date, only the *C. jejuni* glycosylation pathway has been reconstituted *in vitro*[93], and it remains an open question whether our system can be reconfigured with different LLOs and OSTs. Therefore, to extend the range of glycan structures beyond the *C. jejuni* heptasaccharide, we performed glycosylation reactions in which the solvent-extracted *Cj*LLOs used above were replaced with oligosaccharide donors extracted from *E. coli* cells carrying alternative glycan biosynthesis pathways. These included LLOs bearing the following glycan structures: (i) native *C. lari* hexasaccharide *N*-glycan[106]; (ii) engineered GalNAc$_5$GlcNAc based on the *Campylobacter lari* hexasaccharide *N*-glycan[107]; (iii) native *Wolinella succinogenes* hexasaccharide *N*-glycan containing three 216-Da monosaccharides and an unusual 232-Da residue at the non-reducing end[108]; (iv) engineered *E. coli* O9 primer-adaptor glycan, Man$_3$GlcNAc, that links the O-chain and core oligosaccharide in the lipopolysaccharide of several *E. coli* and *Klebsiella pneumoniae* serotypes[109]; and (v) eukaryotic trimannosyl core *N*-glycan, Man$_3$GlcNAc$_2$[23]. Glycosylation of scFv13-R4[DQNAT] with each of these different glycans was observed to occur only in the presence of *Cj*PglB (**Figure 2.3**). It should be noted that 100% glycosylation conversion was observed for each of these glycans except for the Man$_3$GlcNAc$_2$ *N*-glycan, which had a conversion of ~40% as determined by densitometry analysis. While the reasons for this lower efficiency remain unclear, conjugation efficiency of the same Man$_3$GlcNAc$_2$ glycan to acceptor proteins in vivo was reported to be even lower (<5%)[23,110]. Hence, transfer of Man$_3$GlcNAc$_2$ to acceptor proteins *in vitro* appears to overcome some of the yet-to-be-

identified bottlenecks of in vivo glycosylation. This result is likely due to the opportunity with CFGpS

to control the concentration of reaction components, for example, providing a higher local

concentration of LLO donors. Importantly, scFv13-R4^{DQNAT} was uniformly decorated with a

Man$_3$GlcNAc$_2$ glycan as evidenced by liquid chromatography-mass spectrometry (LC–MS).

Specifically, the only major glycopeptide product to be detected was a triply-charged ion containing

an *N*-linked pentasaccharide with m/z = 1032.4583, consistent with the Man$_3$GlcNAc$_2$ glycoform

(**Figure 2.4**). The tandem MS spectra for this triply-charged glycopeptide yielded an excellent y-ion

series and a good b-ion series enabling conclusive determination of the tryptic glycopeptide sequence

and attachment of the Man$_3$GlcNAc$_2$ glycoform at residue N273 of the scFv13-R4^{DQNAT} protein (**Figure

2.5**). Taken together, these results demonstrate that structurally diverse glycans, including those that

resemble eukaryotic structures, can be modularly interchanged in cell-free glycosylation reactions.
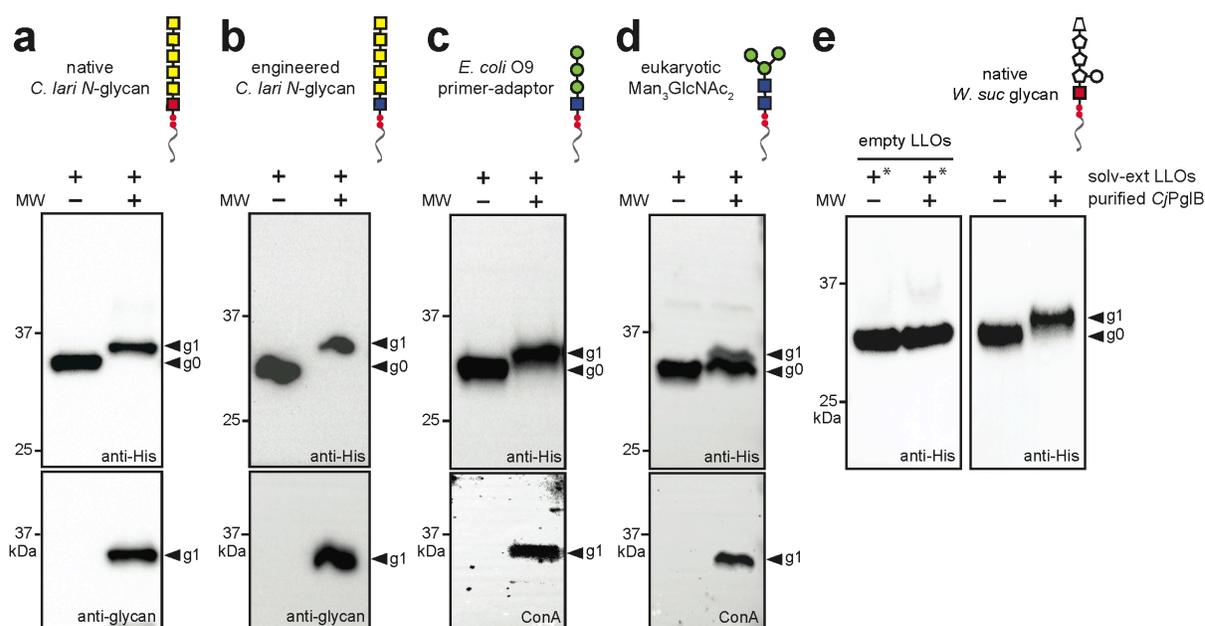
Figure 2.3 Expanding cell-free glycosylation with different oligosaccharide structures. Western blot analysis of *in vitro* glycosylation reaction products generated with purified scFv13-R4^{DQNAT} acceptor protein, purified *Cj*PglB, and organic solvent-extracted (solv-ext) LLOs from cells carrying: (**a**) plasmid pACYC*pgl4* for making the native *C. lari* hexasaccharide *N*-glycan; (**b**) plasmid pACYC*pgl2* for making the engineered *C. lari* hexasaccharide *N*-glycan; (**c**) plasmid pO9-PA for making the *E. coli* O9 'primer-adaptor' Man$_3$GlcNAc structure; (**d**) plasmid pConYCGmCB for making the eukaryotic Man$_3$GlcNAc$_2$ *N*-glycan structure; and (**e**) fosmid pEpiFOS-5*pgl5* for making the native *W. succinogenes* hexasaccharide *N*-glycan. Reactions were run at 30 °C for 16 h. Blots were

probed with anti-His antibody to detect the acceptor protein and one of the following: hR6 serum that cross-reacts with the native and engineered *C. lari* glycans or ConA lectin that binds internal and non-reducing terminal α-mannosyl groups in the $Man_3GlcNAc$ and $Man_3GlcNAc_2$ glycans. Because structural determination of the *W. succinogenes* *N*-glycan is currently incomplete, and because there are no available antibodies, the protein product bearing this *N*-glycan was only probed with the anti-His antibody. As an additional control for this glycan, we included empty LLOs prepared from the same host strain but lacking the pEpiFOS-5*pgl5* fosmid (left hand panel, "+" signs marked with an asterisk). Arrows denote aglycosylated (g0) and singly glycosylated (g1) forms of the scFv13-R4^{DQNAT} protein. Molecular weight (MW) markers are indicated at left. Results are representative of at least three biological replicates.
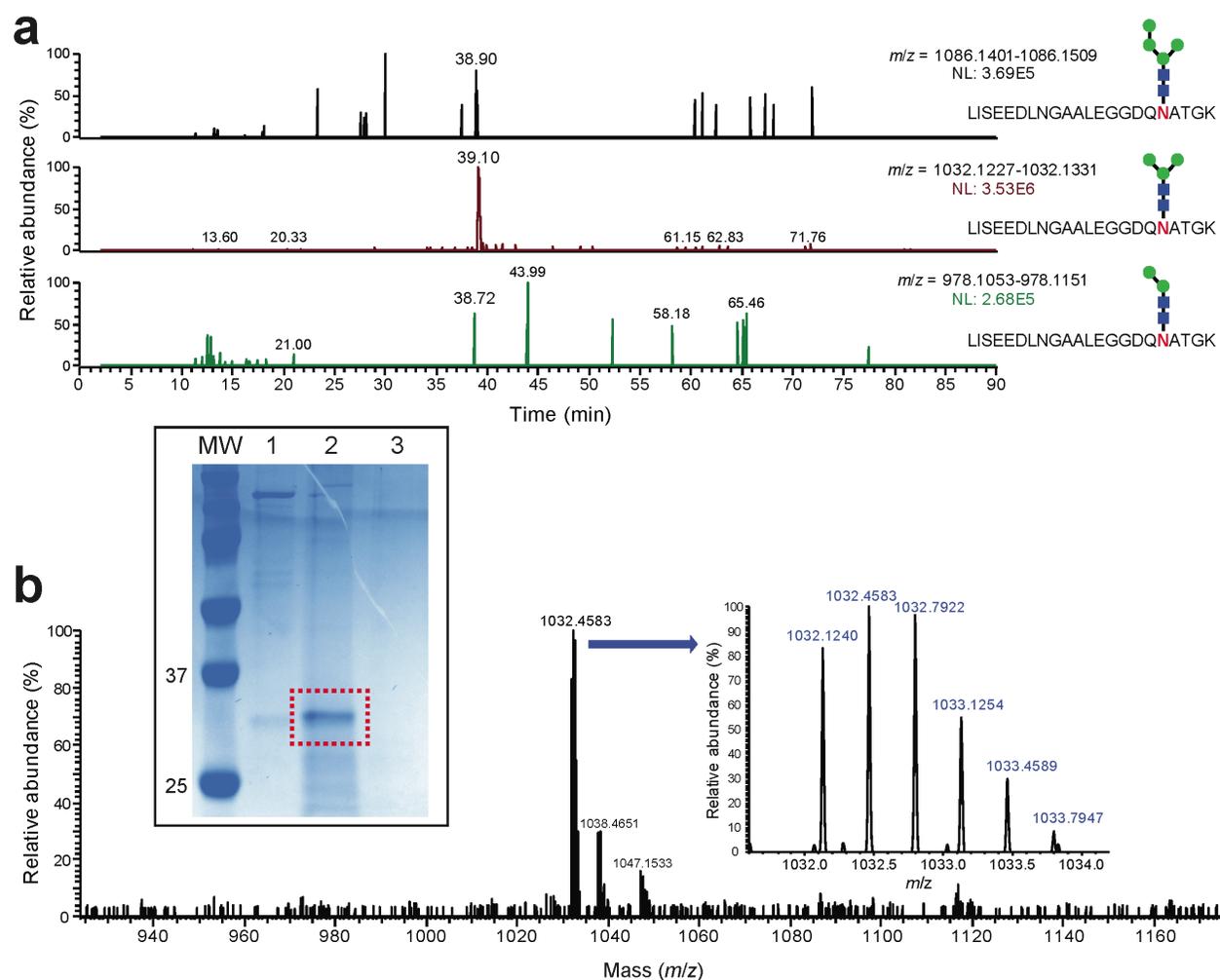
Figure 2.4 MS analysis of scFv13-R4^{DQNAT} glycosylated with $Man_3GlcNAc_2$. Ni-NTA-purified scFv13-R4^{DQNAT} was subjected to *in vitro* glycosylation in the presence of purified *Cj*PglB and organic solvent-extracted $Man_3GlcNAc_2$ LLOs, and then directly loaded into an SDS-PAGE gel. Following staining of gel with Coomassie Brilliant Blue G-250 (**inset**), the glycosylated band (lane 2, indicated by red box) was excised and submitted for MS analysis. Controls included *in vitro* glycosylation reaction performed with solvent-extracted empty LLOs (lane 1) and complete *in vitro* glycosylation reaction mixture lacking purified scFv13-R4^{DQNAT} acceptor protein (lane 3). Molecular weight (MW) ladder loaded on the left. (**a**) Three extracted ion chromatograms (XIC) corresponding to mass ranges for three possible glycopeptide products having masses consistent with the expected $Man_3GlcNAc_2$ (**middle**), as well as $Man_4GlcNAc_2$

(**top**) and Man2GlcNAc2 (**bottom**) attached to N273 site of scFv13-R4^{DQNAT} (mass tolerance at 5 ppm). The individually normalized level (NL) for each glycoform indicates that only a Hex3HexNAc2 glycoform, which eluted at 39.10 min with NL of 3.53E6, was decently detected in the sample (**middle**). A trace amount of a Hex4HexNAc2 glycoform form eluted at 38.9 min with NL of 2.96E5 (**top**), but no Hex2HexNAc2 glycoform was detected. (**b**) MS spectrum of the detected glycopeptide containing an *N*-linked pentasaccharide consistent with Man3GlcNAc2 at m/z = 1032.4583. The MS inset shows an expanded view of the glycopeptide ion with triple charge.
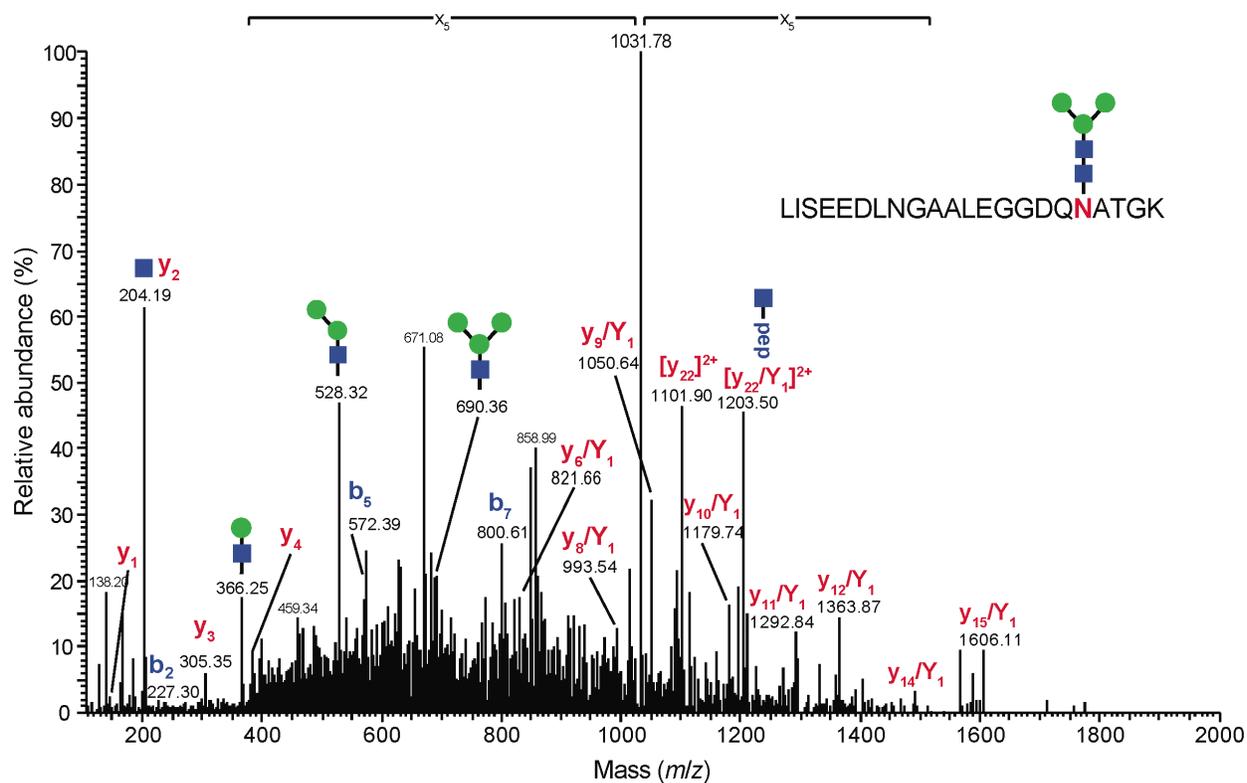
Figure 2.5 Tandem mass spectrometry of scFv13-R4^{DQNAT} glycosylated with Man3GlcNAc2. MS/MS spectrum of the triply-charged precursor (m/z 1032.12), identifying the glycopeptide with core pentasaccharide (Hex3HexNAc2) attached to residue N273 (shown in red) in scFv13-R4^{DQNAT}. A series of y-ions covering from y1 to y4 and a second series of y-ions with the added mass of 203.08 Da at N273 site were found covering from y6/Y1 to y15/Y1, leading to the confident identification of tryptic peptide 256-LISEEDLNGAALEGGDQNATGK-277 and providing direct evidence for HexNAc as the innermost monosaccharide (Y1) attached to the N273 site. This result is also consistent with the previous observation that a relatively tight bond exists for the Y1-peptide compared to the fragile internal glycan bonds.

2.4.3 Extracts enriched with OSTs or LLOs co-activate glycosylation

To circumvent the need for exogenous addition of purified glycosylation components, we hypothesized that heterologous overexpression of OST or GT enzymes directly in the chassis strain

would yield extracts that are selectively enriched with the requisite glycosylation components. This strategy was motivated by a recent metabolic engineering approach whereby multiple cell-free lysates were each selectively enriched with an overexpressed metabolic enzyme and then combinatorially mixed to construct an intact pathway[87,89]. However, a fundamental difference in our system is the fact that the OST and LLOs are not soluble components but instead reside natively in the inner cytoplasmic membrane. This is potentially problematic because of the significant breakup of the cell membrane during S30 extract preparation. However, it has been established that fragments of the *E. coli* inner membrane reform into membrane vesicles, some of which are inverted but others that are orientated properly[111], and thus could supply the OST and LLOs in a functionally accessible conformation within the extract.

To test this hypothesis, we used a high-pressure homogenization method to prepare crude S30 extract from CLM24 cells carrying a plasmid-encoded copy of *Cj*PglB such that the resulting cell-free lysates were selectively enriched with detectable quantities of full-length OST enzyme as confirmed by Western blot analysis (**Figure 2.6a**). Similarly, crude S30 extract from CLM24 cells overexpressing the *C. jejuni* glycan biosynthesis enzymes produced lysate that was selectively enriched with *Cj*LLOs as confirmed by dot blot analysis with hR6 serum (**Figure 2.6b**). It should be noted that the amount of *Cj*LLOs enriched in the crude extract rivaled that produced by the significantly more tedious organic solvent extraction method. Importantly, when 15-μL batch-mode sequential CFGpS reactions were performed using the OST-enriched crude extract that was supplemented with solvent-extracted *Cj*LLOs and plasmid DNA encoding scFv13-R4^{DQNAT}, clearly detectable glycosylation of the acceptor protein was observed (**Figure 2.2b**). The conversion of acceptor protein to glycosylated product was ~50%; however, further supplementation with purified *Cj*PglB increased the conversion to nearly 100%, indicating that the amount of OST in the crude extract might have been limiting under the conditions tested. When similar CFGpS reactions were performed using the *Cj*LLO-enriched crude extract supplemented with purified *Cj*PglB and plasmid DNA encoding scFv13-R4^{DQNAT}, >80% glycosylation of the acceptor protein was observed, which reached 100% when additional donor glycans were supplemented (**Figure 2.2b**).
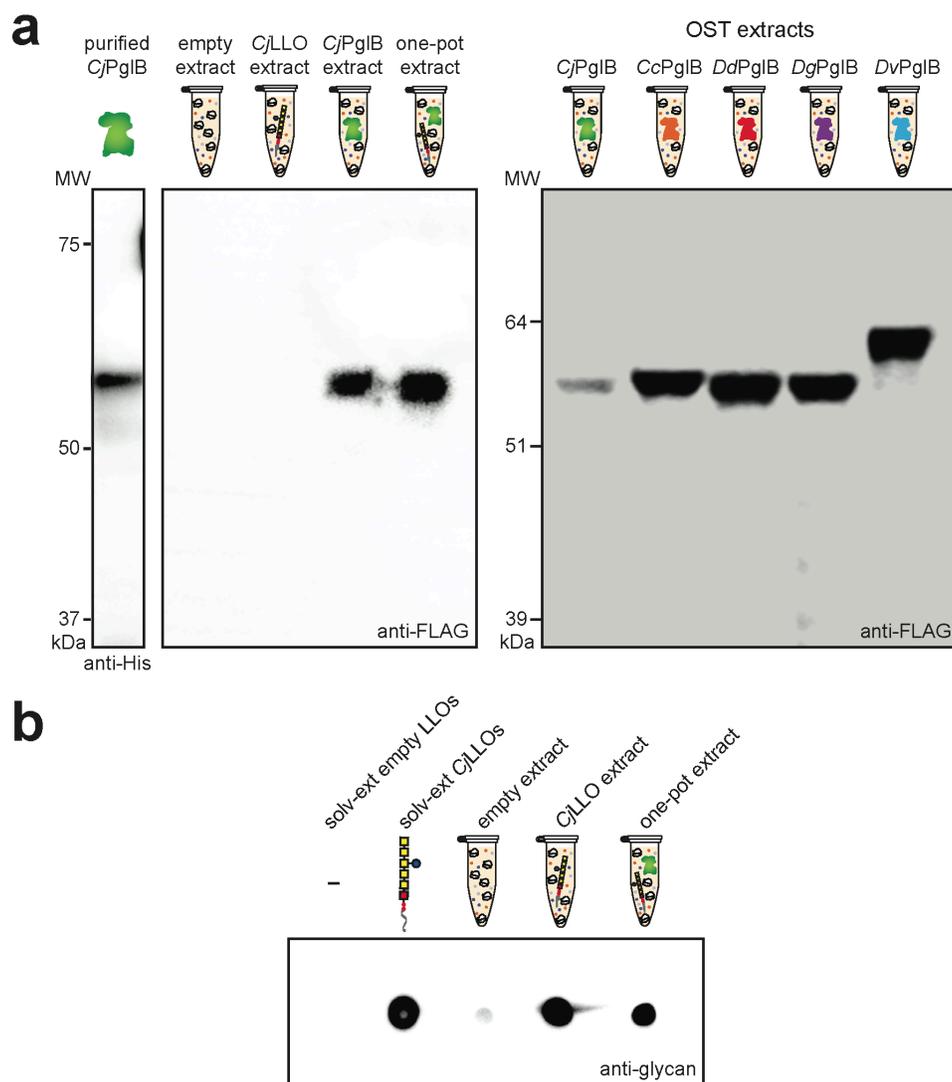
Figure 2.6 Crude cell extracts are enriched with glycosylation machinery. (**a**) Western blot analysis of *Cj*PglB in the following samples: (**left-hand panel**) 1 µg of purified *Cj*PglB; (**center panel**) crude cell extracts derived from CLM24 cells with no plasmid (empty extract), CLM24 cells carrying pMW07-pgl∆B (CjLLO extract), CLM24 cells carrying pSF-*Cj*PglB (*Cj*PglB extract) or CLM24 cells carrying pMW07-pgl∆B and pSF-*Cj*PglB (one-pot extract); and (**right-hand panel**) crude cell extracts derived from CLM24 cells carrying pSF-based plasmids encoding different PglB homologs as indicated. Blots were probed with anti-His antibody and anti-FLAG antibody as indicated. Molecular weight (MW) markers are indicated at left. Results are representative of at least three biological replicates. (**b**) Dot blot analysis of LLOs in the following samples: organic solvent extract from membrane fractions of CLM24 cells with no plasmid (solv-ext empty LLOs) or from CLM24 cells carrying plasmid pMW07-pgl∆B (solv-ext *Cj*LLOs); crude cell extracts derived from CLM24 cells with no plasmid (empty extract), CLM24 cells carrying pMW07-pglB (CjLLO extract) or CLM24 cells carrying pMW07-pgl∆B and pSF-*Cj*PglB (one-pot extract). 10 µl of extracted LLOs or crude cell extract was spotted onto nitrocellulose membrane and probed with hR6 serum (anti-glycan).

2.4.4 Modularity enables glycosylation components to be interchanged

Given the open nature of cell-free biosynthesis, we postulated that it should be possible to functionally interchange and prototype alternative biochemical reaction components. One straightforward way that this can be accomplished is by combining separately prepared extracts, each of which is selectively enriched with a given enzyme, such that the resulting reaction mixture comprises a functional biological pathway[87,89]. As proof of this concept, separately prepared *Cj*LLO and *Cj*PglB extracts were mixed and subsequently primed with DNA encoding the scFv13-R4^{DQNAT} acceptor. The resulting mixture promoted efficient glycosylation of scFv13-R4^{DQNAT} as observed in Western blots probed with anti-His antibody and hR6 serum (**Figure 2.7a**). In addition to scFv13-R4^{DQNAT}, we also expressed a different model acceptor protein that was created by grafting a 21-amino acid sequence from the *C. jejuni* glycoprotein AcrA[93], which was further modified with an optimized DQNAT glycosylation site, into a flexible loop of superfolder GFP (sfGFP$^{217-DQNAT}$). The mixed lysate reaction scheme was able to glycosylate the sfGFP$^{217-DQNAT}$ acceptor protein with 100% conversion (**Figure 2.7a**). It is noteworthy that the high conversion observed for both acceptor proteins was achieved in mixed lysates without the need to supplement the reactions with purified OST or organic solvent-extracted *Cj*LLOs.
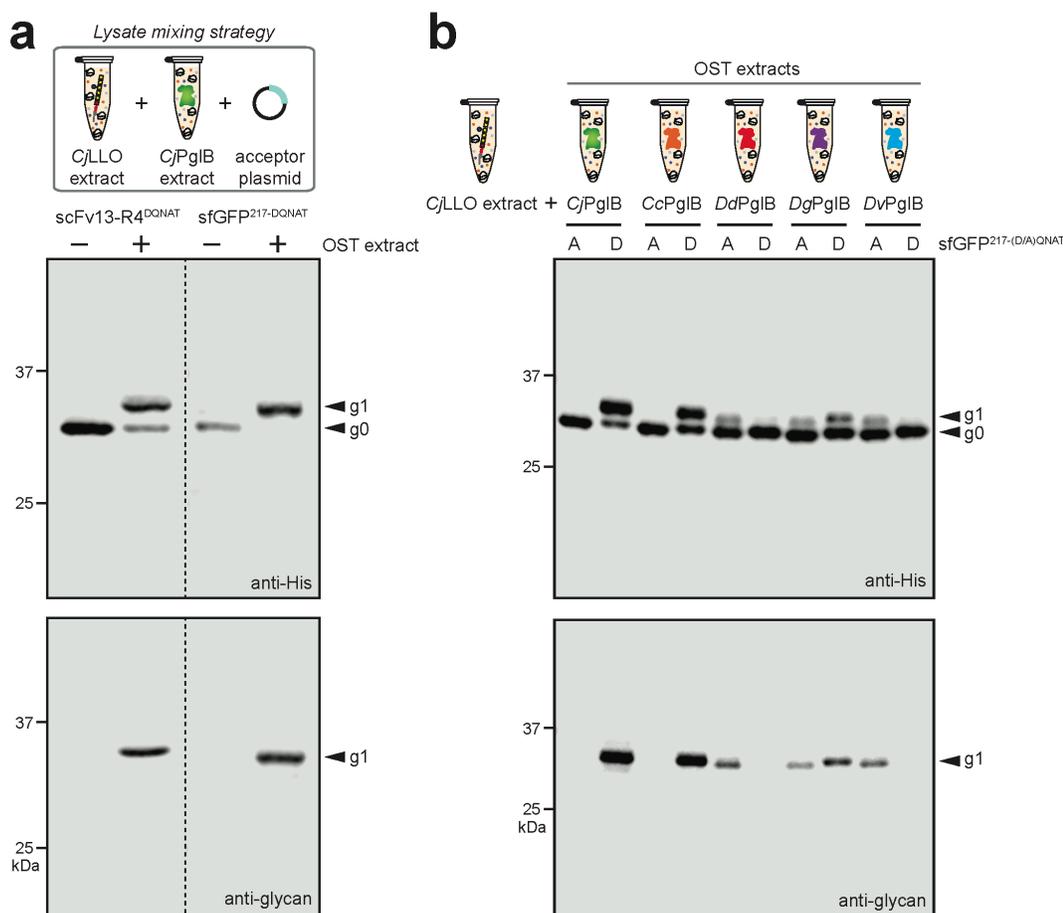
Figure 2.7 Mixing of CFGpS extracts enables rapid prototyping of different OST enzymes. (a) Western blot analysis of CFGpS reactions performed using lysate mixing strategy whereby *Cj*LLO lysate derived from CLM24 cells carrying pMW07-pglΔB was mixed with *Cj*PglB lysate derived from CLM24 cells carrying pSF-*Cj*PglB, and the resulting CFGpS mixture was primed with plasmid DNA encoding either scFv13-R4^{DQNAT} or sfGFP$^{217\text{-}DQNAT}$. **(b)** Western blot analysis of CFGpS reactions performed using *Cj*LLO lysate mixed with extract derived from CLM24 cells carrying a pSF plasmid encoding one of the following OSTs: *Cj*PglB, *Cc*PglB, *Dd*PglB, *Dg*PglB, or *Dv*PglB. Mixed lysates were primed with plasmid DNA encoding either sfGFP$^{217\text{-}DQNAT}$ (D) or sfGFP$^{217\text{-}AQNAT}$ (A). Blots were probed with anti-His antibody to detect the acceptor proteins (**top panels**) and hR6 serum against the *C. jejuni* glycan (**bottom panels**). Arrows denote aglycosylated (g0) and singly glycosylated (g1) forms of the acceptor proteins. Molecular weight (MW) markers are indicated at left. Results are representative of at least three biological replicates.

Next, we sought to demonstrate that the mixed lysate approach could be used to rapidly prototype the activity of four additional bacterial OSTs. Crude extracts were separately prepared from CLM24 source strains heterologously overexpressing one of the following bacterial OSTs: *Campylobacter coli* PglB (*Cc*PglB), *Desulfovibrio desulfuricans* PglB (*Dd*PglB), *Desulfovibrio gigas* PglB (*Dg*PglB), or *Desulfovibrio vulgaris* PglB (*Dv*PglB). The resulting extracts were selectively

enriched with full-length OST proteins at levels that were comparable to CjPglB (**Figure 2.6a**). Each OST extract was mixed with the CjLLO-enriched extract and then supplemented with plasmid DNA encoding sfGFP[217-DQNAT] or a modified version of this target protein where the residue in the −2 position of the acceptor sequon was mutated to alanine. Upon completion of CFGpS reactions, the expression and glycosylation status of sfGFP[217-DQNAT] and sfGFP[217-AQNAT] was followed by western blot analysis, which revealed information about the sequon preferences for these homologous enzymes. For example, the mixed lysate containing CcPglB was observed to efficiently glycosylate sfGFP[217-DQNAT] but not sfGFP[217-AQNAT] (**Figure 2.7b**). This activity profile for CcPglB was identical to that observed for CjPglB, which was not surprising based on its high sequence similarity (~81%) to CjPglB. In contrast, lysate mixtures containing OSTs from *Desulfovibrio* sp., which have low sequence identity (~15-20%) to CjPglB, showed more relaxed sequon preferences (**Figure 2.7b**). Specifically, DgPglB-enriched extract mixtures modified both (D/A)QNAT motifs with nearly equal efficiency while mixed lysates containing DdOST and DvOST preferentially glycosylated the AQNAT sequon.

2.4.5 One-pot extract promotes biosynthesis of diverse glycoproteins

To create a fully integrated CFGpS platform that permits one-pot synthesis of *N*-glycoproteins without the need for supplementation of either purified OSTs or solvent-extracted LLOs (**Figure 2.1**), we produced crude S30 extract from CLM24 cells heterologously overexpressing CjPglB and the *C. jejuni* glycan biosynthesis enzymes. The resulting extract was selectively enriched with both CjPglB and CjLLOs donor to an extent that was indistinguishable from the separately prepared extracts (**Figure 2.6**). Using this extract, CFGpS reactions were performed by addition of plasmid DNA encoding either scFv13-R4[DQNAT] or sfGFP[217-DQNAT]. In both cases, 100% protein glycosylation was achieved without the need for exogenous supplementation of separately prepared glycosylation components (**Figure 2.8a**). Independent extract preparations yielded identical results for both protein substrates, confirming the reproducibility of the CFGpS system (**Figure 2.9**). Importantly, the *in vitro* synthesized scFv13-R4[DQNAT] and sfGFP[217-DQNAT] proteins retained biological activity that was unaffected by *N*-glycan addition (**Figure 2.10**; **Figure 2.11**). From the activity data, the yield of

glycosylated scFv13-R4^{DQNAT} and sfGFP$^{217-DQNAT}$ proteins produced by the one-pot CFGpS system was determined to be ~20 and ~10 mg L^{-1}, respectively.
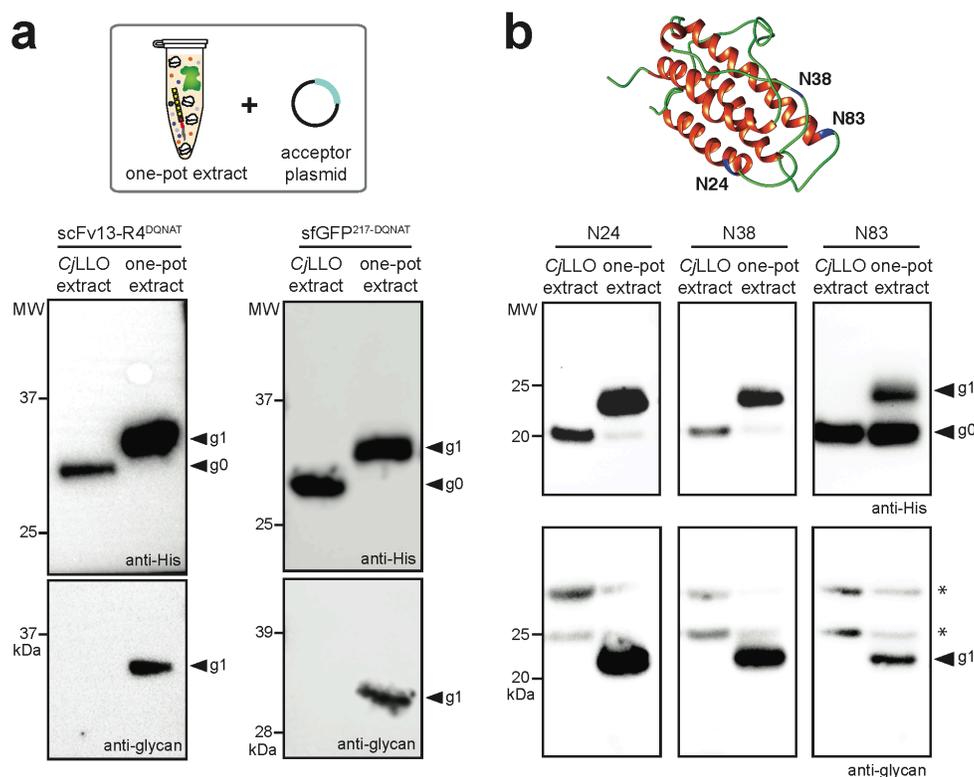
Figure 2.8 One-pot CFGpS using extracts selectively enriched with OSTs and LLOs. (**a**) Western blot analysis of scFv13-R4^{DQNAT} or sfGFP$^{217-DQNAT}$ produced by crude CLM24 extract selectively enriched with (i) *Cj*PglB from heterologous overexpression from pSF-*Cj*PglB and (ii) *Cj*LLOs from heterologous overexpression from pMW07-pglΔB. Reactions were primed with plasmid pJL1-scFv13-R4^{DQNAT} or pJL1-sfGFP$^{217-DQNAT}$. (**b**) Ribbon representation of human erythropoietin (PDB code 1BUY) with α-helices and flexible loops colored in red and green, respectively. Glycosylation sites modeled by mutating the native sequons at N24 (22-AENIT-26), N38 (36-NENIT-40), or N83 (81-LVNSS-85) to DQNAT, with asparagine residues in each sequon colored in blue. Image prepared using UCSF Chimera package.[112] Glycoengineered hEPO variants in which the native sequons at N24 (22-AENIT-26), N38 (36-NENIT-40), or N83 (81-LVNSS-85) were individually mutated to an optimal bacterial sequon, DQNAT (shown in blue). Western blot analysis of hEPO glycovariants produced by crude CLM24 extract selectively enriched with (i) *Cj*PglB from heterologous overexpression from pSF-*Cj*PglB and (ii) *Cj*LLOs from heterologous overexpression from pMW07-pglΔB. Reactions were primed with plasmid pJL1-hEPO$^{22-DQNAT-26}$ (N24), pJL1-hEPO$^{36-DQNAT-40}$ (N38), or pJL1-hEPO$^{81-DQNAT-85}$ (N83) as indicated. All control reactions (lane 1 in each panel) were performed using *Cj*LLO-enriched extracts that lacked *Cj*PglB. Blots were probed with anti-hexa-histidine antibody (anti-His) to detect the acceptor proteins or hR6 serum (anti-glycan) to detect the *N*-glycan. Arrows denote aglycosylated (g0) and singly glycosylated (g1) forms of the protein targets. Asterisks denote bands corresponding to non-specific serum antibody binding. Molecular weight (MW) markers are indicated at left. Results are representative of at least three biological replicates (see **Figure 2.9** for replicate data).
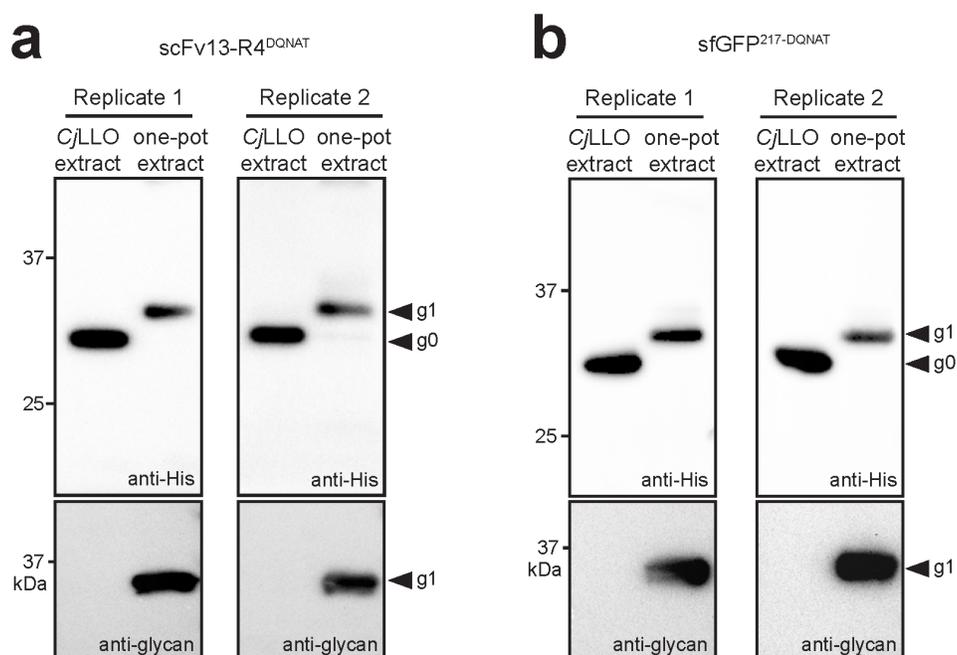
Figure 2.9 Independent biological replicates for one-pot CFGpS reactions. Western blot analysis replicated twice for both the (**a**) scFv13-R4^{DQNAT} and (**b**) sfGFP$^{217\text{-}DQNAT}$ acceptor proteins produced using crude CLM24 extract selectively enriched with (i) *Cj*PglB from heterologous overexpression from pSF-CjPglB and (ii) *Cj*LLOs from heterologous overexpression from pMW07-pglΔB. Each replicate experiment involved charging freshly prepared cell-free extracts with freshly purified pJL1-scFv13-R4^{DQNAT} or pJL1-sfGFP$^{217\text{-}DQNAT}$ plasmid DNA. Control reactions (lane 1 in each panel) were performed using *Cj*LLO-enriched extracts that lacked *Cj*PglB. Blots were probed with anti-hexa-histidine antibody (anti-His) to detect acceptor proteins or hR6 serum (anti-glycan) to detect the N-glycan. Arrows denote aglycosylated (g0) and singly glycosylated (g1) forms of the protein targets. Molecular weight (MW) markers are indicated at left.

To determine whether human glycoproteins could be similarly produced in our one-pot system, we constructed plasmids for cell-free expression of human erythropoietin (hEPO) glycovariants in which the native sequons at residue N24 (22-AENIT-26), N38 (36-NENIT-40) or N83 (81-LVNSS-85) were individually mutated to the optimal bacterial sequon, DQNAT (**Figure 2.8b**). CFGpS reactions were then initiated by priming the all-in-one extract with plasmid DNA encoding hEPO$^{22\text{-}DQNAT\text{-}26}$, hEPO$^{36\text{-}DQNAT\text{-}40}$, or hEPO$^{81\text{-}DQNAT\text{-}85}$. Western blot analysis revealed clearly detectable glycosylation of each hEPO glycovariant with 100% glycosylated product for the N24 and N38 sites and ~30–40% for the N83 site (**Figure 2.8b**). As with the model glycoproteins scFv13-R4^{DQNAT} and sfGFP$^{217\text{-}DQNAT}$ above, all three glycosylated hEPO variants retained biological activity that was indistinguishable from the activity measured for the corresponding aglycosylated counterparts, with

yields in the ~10 mg L^{-1} range (**Figure 2.12**). Collectively, these findings establish that one-pot

CFGpS extracts are capable of co-activating protein synthesis and *N*-glycosylation in a manner that

yields efficiently glycosylated proteins including those of human origin.

Figure 2.10 CFGpS expression of active sfGFP. In-lysate fluorescence activity for glycosylated (one-pot CFGpS) and aglycosylated (*Cj*LLOs extract) sfGFP[217-DQNAT] produced in cell-free reactions charged with plasmid pJL1-sfGFP[217- DQNAT] (blue) or with no plasmid DNA (red). Following 2-h reactions, cell-free reactions containing glycosylated and aglycosylated sfGFP[217-DQNAT] were diluted 10 times with water and then subjected to fluorescence measurement. Excitation and emission wavelengths for sfGFP were 485 and 528 nm, respectively. Calibration curve was prepared by measuring fluorescence intensity of aglycosylated sfGFP[217-DQNAT] expressed and purified from *E. coli* cells and mixed with empty extract. Linear regression analysis (**inset**) was used to calculate the concentration of glycosylated sfGFP[217-DQNAT] in the samples, which was determined to be ~10 mg L^{-1}. Data are the average of three biological replicates and error bars represent the standard deviation of the mean.
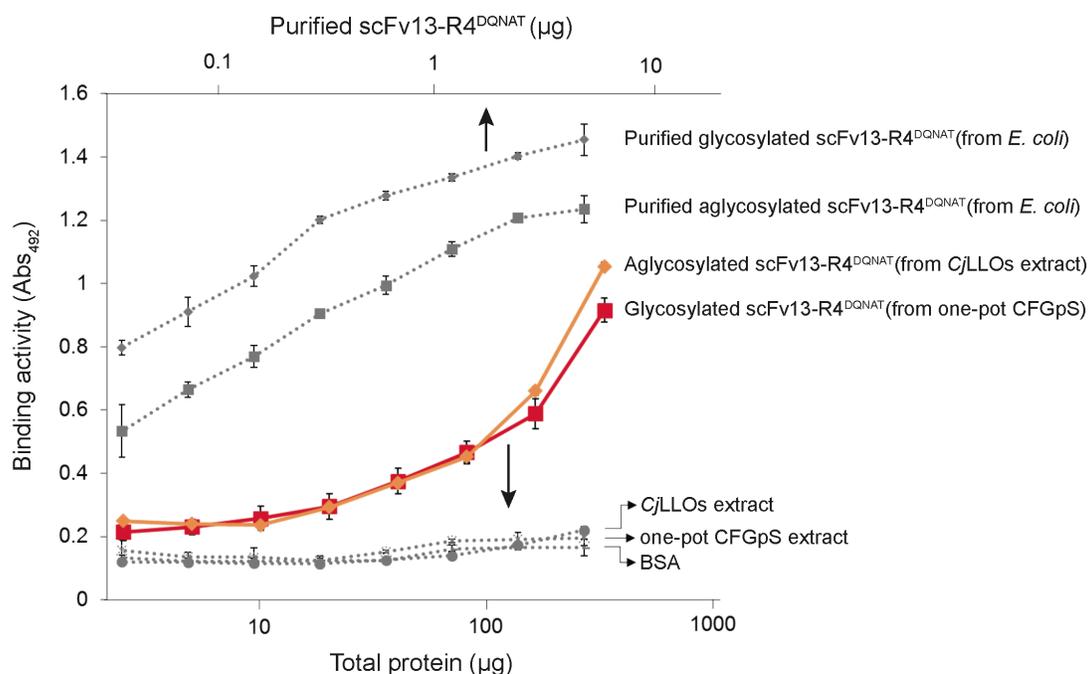
Figure 2.11 CFGpS expression of active scFv antibody fragment. Antigen-binding activity for β-gal-specific scFv13-R4^{DQNAT} measured by ELISA with E. coli β-gal as immobilized antigen. The scFv13-R4^{DQNAT} acceptor was produced as a glycosylated protein in one-pot CFGpS (red) or an aglycosylated protein in control extracts containing *Cj*LLOs but not *Cj*PglB (orange). Extracts were primed with plasmid pJL1-scFv13-R4^{DQNAT}. Positive controls included the same scFv13-R4^{DQNAT} protein produced *in vivo* by recombinant expression in *E. coli* in the presence (glycosylated) or absence (aglycosylated) of glycosylation machinery. Negative controls included extracts without plasmid and BSA. Comparing to signals from purified protein, the concentration of glycosylated scFv13-R4^{DQNAT} was determined to be ~20 mg L^{-1}. Data are the average of three biological replicates and error bars represent the standard deviation of the mean.

To determine whether human glycoproteins could be similarly produced in our one-pot system, we constructed plasmids for cell-free expression of human erythropoietin (hEPO) glycovariants in which the native sequons at residue N24 (22-AENIT-26), N38 (36-NENIT-40) or N83 (81-LVNSS-85) were individually mutated to the optimal bacterial sequon, DQNAT (**Figure 2.8b**). CFGpS reactions were then initiated by priming the all-in-one extract with plasmid DNA encoding hEPO[22-DQNAT-26], hEPO[36-DQNAT-40], or hEPO[81-DQNAT-85]. Western blot analysis revealed clearly detectable glycosylation of each hEPO glycovariant with 100% glycosylated product for the N24 and N38 sites and ~30–40% for the N83 site (**Figure 2.8b**). As with the model glycoproteins scFv13-R4^{DQNAT} and sfGFP[217-DQNAT] above, all three glycosylated hEPO variants retained biological activity that was indistinguishable from the activity measured for the corresponding aglycosylated counterparts, with

yields in the ~10 mg L^{-1} range (**Figure 2.12**). Collectively, these findings establish that one-pot CFGpS extracts are capable of co-activating protein synthesis and *N*-glycosylation in a manner that yields efficiently glycosylated proteins including those of human origin.

Figure 2.12 CFGpS-derived hEPO glycovariants stimulate cell proliferation. Stimulation of human erythroleukemia TF-1 cell proliferation following incubation with purified rhEPO standard or hEPO variants produced in cell-free reactions. For CFGpS-derived hEPO glycovariants, TF-1 cells were treated with either glycosylated hEPO variants produced in one-pot CFGpS (blue) or aglycosylated hEPO variants produced in control extracts containing *Cj*LLOs but not *Cj*PglB (red). To produce the hEPO variants, extracts were primed with plasmid pJL1-hEPO[22-DQNAT-26] (N24), pJL1-hEPO[36-DQNAT-40] (N38), or pJL1-hEPO[81-DQNAT-85] (N83). For positive control rhEPO samples, cells were treated with serial dilutions of commercial rhEPO that was purified from CHO cells and thus glycosylated (green). TF-1 cells incubated with empty extracts or PBS (unstimulated) served as negative controls while RPMI media without cells was used as the blank. Regression analysis (**inset**) was performed to determine the concentration of hEPO variants in the samples, which was found to be at ~10 mg L^{-1}. Data are the average of three biological replicates and error bars represent the standard deviation of the mean.

2.5 Discussion

In this work, we successfully created a technology for one-pot biosynthesis of *N*-linked glycoproteins in the absence of living cells. This was accomplished by uniting cell-free transcription and translation with the necessary reaction components for *N*-linked protein glycosylation through a process of crude extract enrichment. By preparing OST- and LLO-enriched crude S30 extracts from a glyco-optimized chassis strain, glycosylation-competent lysates were capable of supplying efficiently glycosylated target proteins, with conversion levels at or near 100% in most instances. The glycoprotein yields obtained for three structurally diverse proteins were in the 10–20 mg L^{-1} range, which compare favorably to some of the yields reported previously for these proteins in different CFPS kits or in-house generated extracts. For example, Jackson et al.[113] produced 3.6 mg L^{-1} of GFP using the PURExpress system, Stech et al.[114] produced ~12 mg L^{-1} of an anti-SMAD2 scFv using a CHO cell-derived lysate, Ahn et al.[115] produced 55 mg L^{-1} of hEPO using an *E. coli*-derived S30 lysate, and Gurramkonda et al.[57] produced ~120 mg L^{-1} of hEPO using a CHO cell-derived lysate supplemented with CHO microsomes.

Furthermore, this work represents the first demonstration of extract enrichment with catalytically active multipass transmembrane enzymes (and their corresponding lipid-linked substrates) without the need for domain truncation or supplementation of extra scaffold molecules[97], and provides a blueprint for other CFPS-based applications beyond glycosylation that involve this important class of proteins. Moreover, the ability of OST- or LLO-enriched crude extracts to co-activate glycosylation partially bypassed the need for costly, labor-intensive preparation of glycosylation components and paved the way for a modular single-pot CFGpS platform in which protein synthesis and *N*-linked glycosylation were integrated.

A major advantage of the CFGpS system developed here is the level of control it affords over each of the glycosylation components (i.e., catalysts, substrates, and cofactors) in terms of important process variables such as relative concentration, timing of addition, overall reaction time, etc. Likewise, genome engineering of the chassis strain used to supply the extract, such as our recent report enhancing cell-free synthesis containing multiple, identical non-canonical amino acids[83], makes

it possible to eliminate inhibitory substances such as glycosidases that catalyze the undesired hydrolysis of glycosidic linkages. This user-level control provides an opportunity to overcome system bottlenecks that effectively limit glycosylation efficiency as we showed with both the *C. jejuni* heptasaccharide and the eukaryotic $Man_3GlcNAc_2$ glycan. Moreover, the open nature of the CFGpS system could be further exploited in the future to introduce components that may otherwise be incompatible with chassis strain expression such as unusual and/or non-natural LLOs that cannot be assembled or flipped in vivo.

An additional advantage of the CFGpS system is that it does not rely on commercial cell-free kits to support protein synthesis. For comparison, the glycoproteins yields obtained here were ~10–20 mg L^{-1} in reactions costing ~\$0.01–0.03 per µL (**Table 2.2** and also ref. [116]) versus previous kit-based (e.g., Promega L110; NEB® E6800S) glycoprotein yields of ~100 mg L^{-1} [93] in reactions costing ~\$1 µL^{-1} [117]. As a result, our system can synthesize ~1 µg glycoprotein/\$ reagents compared to the previously published approach that can synthesize ~0.1 µg glycoprotein/\$ reagents, representing an order of magnitude improvement in relative protein synthesis yields. It is also worth noting that this cost analysis does not take into account the cost of purifying OSTs or extracting LLOs that were used to supplement the commercial kits in our previous work[93]. We anticipate this reduction in cost will encourage adoption of the CFGpS platform.

Perhaps the most important feature of the CFGpS platform is its modularity, which was evidenced by the interchangeability of: (i) OST enzymes from different bacterial species; (ii) engineered LLOs with glycan moieties derived from bacteria and eukaryotes; and (iii) diverse acceptor protein targets including naturally occurring human *N*-glycoproteins with terminal or internal acceptor sequons. Importantly, enriched extracts could be readily mixed in a manner that enabled screening of an OST panel whose activities in CFGpS were in line with previously reported activities in vivo[96], thereby validating this lysate mixing strategy as a useful tool for rapid characterization of glycosylation enzyme function and for prototyping glycosylation reactions. In light of this modularity, we envision that lysate enrichment could be further expanded beyond the glycosylation components/substrates tested here. For example, extracts could be heterologously enriched with alternative membrane-bound

or soluble OSTs that catalyze *N*-linked or *O*-linked glycosyl transfer reactions. Such biocatalyst swapping is expected to be relatively straightforward in light of the growing number of prokaryotic and eukaryotic OST enzymes that have been recombinantly expressed in functional conformations and used to promote *in vitro* glycosylation reactions[82,96,97,118-121]. Likewise, as newly engineered glycan biosynthesis pathways emerge[20], these could be readily integrated into the CFGpS platform through heterologous expression of GTs in the chassis strain. The ability to modularly reconfigure and quickly interrogate glycosylation systems *in vitro* should make the CFGpS technology a useful new addition to the glycoengineering toolkit for increasing our understanding of glycosylation and, in the future, advancing applications of on demand biomolecular manufacturing[13,50,51,122].

2.6 Additional information

2.6.1 Acknowledgements

2.6.2 Author contributions

T.J. and J.C.S. designed research, performed all research, analyzed all data, and wrote the paper. A.N., C.G., L.E.Y., and K.J.H. performed research. M.M. aided in data interpretation. M.C.J. and M.P.D. directed research, analyzed data, and wrote the paper.

2.6.3 Publishing and patent information

Published manuscript

Jaroentomeechai, T. †, Stark, J. C. †, Natarajan, A., Glasscock, C., Mrksich, M., Jewett, M. C. *, & DeLisa, M. P. * (2018) Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nature Communications*, doi: 10.1038/s41467-018-05110-x.

† contributed equally; * corresponding author

Patent application

Jewett, M. C., Stark, J. C., DeLisa, M. P., & Jaroentomeechai, T. (2015) Cell-free glycoprotein synthesis (CFGpS) in prokaryotic cell lysates enriched with components for glycosylation. PCT International Patent Application Serial No. PCT/US2016/069512.

2.7 Tables

Table 2.1 Plasmids used in this study.

Plasmids	Description	Reference
pSN18	modified pBAD expression plasmid encoding *C. jejuni pglB* with a C-terminal decahistidine affinity tag.	99
pET28a-scFv13-R4 (N34L, N77L)^{DQNAT}	pET28a(+) plasmid encoding scFv13-R4 modified with a C-terminal DQNAT glycosylation tag and two mutations (N34, N77L) to eliminate putative internal glycosylation sites.	96
pMW07-pglΔB	pMW07 plasmid encoding *C. jejuni* protein glycosylation locus (*pgl*) with complete in-frame deletion of *Cj*PglB	123
pACYC*pgl2*	pACYC plasmid encoding modified *C.lari* hexasaccharide glycan biosynthesis gene cluster lacking bacillosamine biosynthesis genes	107
pACYC*pgl4*	pACYC plasmid encoding native *C.lari* hexasaccharide hexasaccharide glycan biosynthesis gene cluster	106
pEpiFOS-5*pgl5*	pEpiFOS-5 encoding the *Wolinella succinogenes* hexasaccharide glycan biosynthesis gene cluster cloned in the Eco72 site	Lab stock
pConYCG-mCB	pMW07 plasmid encoding Man$_3$GlcNAc$_2$ glycan biosynthesis genes and *manCB* genes for GDP-mannose biosynthesis	Lab stock
pJL1-scFv13-R4^{DQNAT}	pJL1 plasmid encoding scFv13-R4 (N34L, N77L)^{DQNAT}	This study
pJL1-sfGFP$^{217-DQNAT}$	pJL1 plasmid encoding superfolder GFP modified after residue T216 with 21 amino acid insertion containing the *C. jejuni* AcrA N123 glycosylation site but with an optimal DQNAT sequon	This study
pJL1-sfGFP$^{217-AQNAT}$	same as pJL1-sfGFP^{DQNAT} but with AQNAT sequon	This study
pJL1-hEPO$^{22-DQNAT-26}$	pJL1 plasmid encoding human erythropoietin with native glycosylation motif surrounding N22 mutated to DQNAT	This study
pJL1-hEPO$^{36-DQNAT-40}$	pJL1 plasmid encoding human erythropoietin with native glycosylation motif surrounding N38 mutated to DQNAT	This study
pJL1-hEPO$^{81-DQNAT-85}$	pJL1 plasmid encoding human erythropoietin with native glycosylation motif surrounding N83 mutated to DQNAT	This study
pSF-*Cj*PglB	pSN18 derivative encoding *C. jejuni* PglB with C-terminal FLAG epitope tag	123
pSF-*Cc*PglB	pSN18-derivative encoding *C. coli* PglB with C-terminal FLAG epitope tag	123
pSF-*Dd*PglB	pSN18-derivative encoding *D. desulfuricans* PglB with C-terminal FLAG epitope tag	123
pSF-*Dg*PglB	pSN18-derivative encoding *D. gigas* PglB with C-terminal FLAG epitope tag	123
pSF-*Dv*PglB	pSN18-derivative encoding *D. vulgaris* PglB with C-terminal FLAG epitope tag	123

Table 2.2 Cost analysis of CFGpS reactions. The total cost to assemble CFGpS reactions is ~$0.01 per µL. In the table, amino acid cost accounts for 2 mM each of the 20 canonical amino acids purchased individually from Sigma. Extract cost is based on a single employee making 50 mL lysate from a 10 L fermentation, assuming 30 extract batches per year and a 5-year equipment lifetime. Component source is included in the table if it is available to purchase directly from a supplier. Homemade or user-supplied components cannot be purchased directly and must be prepared by the end user according to procedures described in the Methods section.

Component	Cost ($/µL rxn)	Supplier	Product No
$Mg(Glu)_2$	negligible	Sigma	49605
NH_4Glu	negligible	MP	02180595
KGlu	negligible	Sigma	G1501
ATP	negligible	Sigma	A2383
GTP	0.000265	Sigma	G8877
UTP	0.000230	Sigma	U6625
CTP	0.000200	Sigma	C1506
Folinic acid	0.0000206	Sigma	47612
tRNA	0.000215	Roche	10109541001
Amino acids	negligible	homemade	
PEP	0.00179	Roche	10108294001
NAD	negligible	Sigma	N8535-15VL
CoA	0.000336	Sigma	C3144
Oxalic acid	negligible	Sigma	P0963
Putrescine	negligible	Sigma	P5780
Spermidine	negligible	Sigma	S2626
HEPES	negligible	Sigma	H3375
$MnCl_2$	negligible	Sigma	63535
DDM	0.000358	Anatrace	D310S
Plasmid	negligible	user-supplied	
Extract	0.00737	homemade	
Total	**0.0108**	**$/µL rxn**	

3. On-demand, cell-free biomanufacturing of conjugate vaccines at the point-of-care

3.1 Abstract

Conjugate vaccines are among the most effective methods for preventing bacterial infections, representing a promising strategy to combat drug-resistant pathogens. However, existing manufacturing approaches limit access to conjugate vaccines due to centralized production and cold chain distribution requirements. To address these limitations, we developed a modular technology for *in vitro* bioconjugate vaccine expression (iVAX) in portable, freeze-dried lysates from detoxified, nonpathogenic *Escherichia coli.* Upon rehydration, iVAX reactions synthesize clinically relevant doses of bioconjugate vaccines against diverse bacterial pathogens in one hour. We show that iVAX synthesized vaccines against the highly virulent pathogen *Franciscella tularensis* subsp. *tularensis* (type A) strain Schu S4 elicited pathogen-specific antibodies in mice at significantly higher levels compared to vaccines produced using engineered bacteria. The iVAX platform promises to accelerate development of new bioconjugate vaccines with increased access through refrigeration-independent distribution and point-of-care production.

3.2 Introduction

Drug-resistant bacteria are predicted to threaten up to 10 million lives per year by 2050[124], necessitating new strategies to develop and distribute antibiotics and vaccines. Conjugate vaccines, typically composed of a pathogen-specific capsular (CPS) or O-antigen polysaccharide (O-PS) linked to an immunostimulatory protein carrier, are among the safest and most effective methods for preventing life-threatening bacterial infections[29,31,125]. In particular, implementation of meningococcal and pneumococcal conjugate vaccines have significantly reduced the occurrence of bacterial meningitis and pneumonia worldwide[126,127], in addition to reducing antibiotic resistance in targeted strains[128]. However, despite their proven safety and efficacy, global childhood vaccination rates for conjugate vaccines remain

as low as ~30%, with lack of access or low immunization coverage accounting for the vast majority of remaining disease burden[129]. In addition, the 2018 WHO prequalification of Typhbar-TCV® to prevent typhoid fever represents the first conjugate vaccine approval in nearly a decade. In order to address emerging drug-resistant pathogens, new technologies to accelerate the development and global distribution of conjugate vaccines are needed.

Contributing to the slow pace of conjugate vaccine development and distribution is the fact that these molecules are particularly challenging and costly to manufacture. The conventional process to produce conjugate vaccines involves chemical conjugation of carrier proteins with polysaccharide antigens purified from large-scale cultures of pathogenic bacteria. Large-scale fermentation of pathogens results in high manufacturing costs due to associated biosafety hazards and process development challenges. In addition, chemical conjugation can alter the structure of the polysaccharide, resulting in loss of the protective epitope[130]. To address these challenges, it was recently demonstrated that polysaccharide-protein "bioconjugates" can be made in *Escherichia coli* using protein-glycan coupling technology (PGCT)[22]. In this approach, engineered *E. coli* cells covalently attach heterologously expressed CPS or O-PS antigens to carrier proteins via an asparagine-linked glycosylation reaction catalyzed by the *Campylobacter jejuni* oligosaccharyltransferase enzyme PglB (*Cj*PglB)[33-37,131,132]. Despite this advance, both chemical conjugation and PGCT approaches rely on living bacterial cells, requiring centralized production facilities from which vaccines are distributed via a refrigerated supply chain. Refrigeration is critical to avoid conjugate vaccine spoilage due to precipitation and significant loss of the pathogen-specific polysaccharide upon both heating and freezing[32,38]. Only one conjugate vaccine, MenAfriVac™, is known to remain active outside of the cold chain for up to 4 days, which enabled increased vaccine coverage and an estimated 50% reduction in costs during vaccination in the meningitis belt of sub-Saharan Africa[133]. However, this required significant investment in the development and validation of a thermostable vaccine. Broadly, the need for cold chain refrigeration creates economic and logistical challenges that limit the reach of vaccination campaigns and present barriers to the eradication of disease, especially in the developing world[7,129].

Cell-free protein synthesis (CFPS) offers opportunities to both accelerate vaccine development and enable decentralized, cold chain-independent biomanufacturing by using cell lysates, rather than living cells, to synthesize proteins *in vitro*[39]. Importantly, CFPS platforms (i) enable point-of-care protein production, as relevant amounts of protein can be synthesized *in vitro* in just a few hours, (ii) can be freeze-dried for distribution at ambient temperature and reconstituted by just adding water[12,14], and (iii) circumvent biosafety concerns associated with the use of living cells outside of a controlled laboratory setting. CFPS has recently been used to enable on-demand and portable production of aglycosylated protein vaccines[12,14]. Moreover, we recently described a cell-free glycoprotein synthesis technology that enables one-pot production of glycosylated proteins, including human glycoproteins and eukaryotic glycans[134]. Despite these advances, cell-free systems and even decentralized manufacturing systems have historically been limited by their inability to synthesize glycosylated proteins at relevant titers and bearing diverse glycan structures, such as the polysaccharide antigens needed for bioconjugate vaccine production[44].

To address these limitations, here we describe the iVAX (*in vitro* bioconjugate va̲ccine ex̲pression) platform that enables rapid development and cold chain-independent biosynthesis of conjugate vaccines in cell-free reactions (**Figure 3.1**). iVAX was designed to have the following features. First, iVAX is fast, with the ability to produce multiple individual doses of bioconjugates in one hour. Second, iVAX is robust, yielding equivalent amounts of bioconjugate over a range of operating temperatures. Third, iVAX is modular, offering the ability to rapidly interchange carrier proteins, including those used in licensed conjugate vaccines, as well as conjugated polysaccharide antigens. We leverage this modularity to create an array of vaccine candidates targeted against diverse bacterial pathogens, including the highly virulent *Franciscella tularensis* subsp. *tularensis* (type A) strain Schu S4, enterotoxigenic (ETEC) *E. coli* O78, and uropathogenic (UPEC) *E. coli* O7. Fourth, iVAX is shelf-stable, derived from freeze-dried cell-free reactions that operate in a just-add-water strategy. Fifth, iVAX is safe, leveraging lipid A remodeling that effectively avoids the high levels of endotoxin present in non-engineered *E. coli* manufacturing platforms. Our results demonstrate that anti-*F. tularensis* bioconjugates derived from freeze-dried, low-endotoxin iVAX reactions elicit pathogen-specific antibody responses in

mice and outperform a bioconjugate produced using the established PGCT approach in living cells. Overall, the iVAX platform offers a new way to deliver the protective benefits of an important class of antibacterial vaccines to both the developed and developing world.
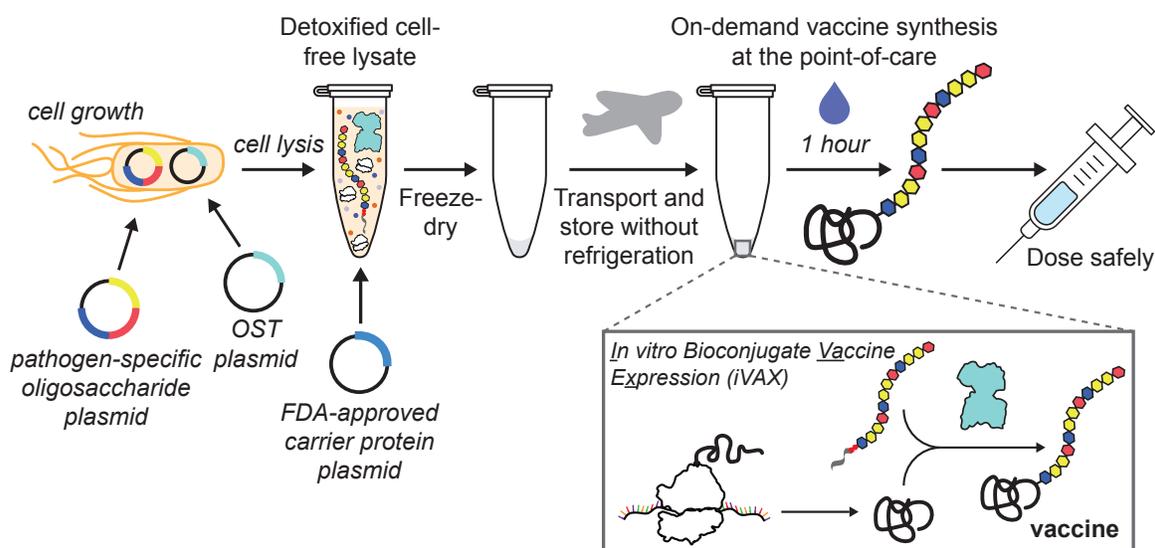
Figure 3.1 iVAX platform enables on-demand and portable production of antibacterial vaccines. The *in vitro* bioconjugate vaccine expression (iVAX) platform provides a rapid means to develop and distribute vaccines against bacterial pathogens. Expression of pathogen-specific polysaccharides (e.g., CPS, O-PS) and a bacterial oligosaccharyltransferase enzyme in engineered nonpathogenic *E. coli* with detoxified lipid A yields low-endotoxin lysates containing all of the machinery required for synthesis of bioconjugate vaccines. Reactions catalyzed by iVAX lysates can be used to produce bioconjugates containing licensed carrier proteins and can be freeze-dried without loss of activity for refrigeration-free transportation and storage. Freeze-dried reactions can be activated at the point-of-care via simple rehydration and used to reproducibly synthesize immunologically active bioconjugates in ~1 h.

3.3   Methods

3.3.1 Bacterial strains and plasmids

E. coli NEB 5-alpha (NEB) was used for plasmid cloning and purification. *E. coli* CLM24 or CLM24 Δ*lpxM* strains were used for preparing cell-free lysates. *E. coli* CLM24 was used as the chassis for expressing bioconjugates *in vivo* using PGCT. CLM24 is a glyco-optimized derivative of W3110 that carries a deletion in the gene encoding the WaaL ligase, facilitating the accumulation of preassembled

glycans on undecaprenyl diphosphate[22]. CLM24 Δ*lpxM* has an endogenous acyltransferase deletion and serves as the chassis strain for production of detoxified cell-free lysates.

E. coli CLM24 Δ*lpxM* was generated using the Datsenko-Wanner gene knockout method[95]. Briefly, CLM24 cells were transformed with the pKD46 plasmid encoding the λ red system. Transformants were grown to an OD_{600} of 0.5-0.7 in 25 mL LB-Lennox media (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract and 5 g L^{-1} NaCl) with 50 µg mL^{-1} carbenicillin at 30°C, harvested and washed three times with 25 mL ice-cold 10% glycerol to make them electrocompetent, and resuspended in a final volume of 100 µL 10% glycerol. In parallel, a *lpxM* knockout cassette was generated by PCR amplifying the kanamycin resistance cassette from pKD4 with forward and reverse primers with homology to *lpxM*. Electrocompetent cells were transformed with 400 ng of the *lpxM* knockout cassette and plated on LB agar with 30 µg mL^{-1} kanamycin for selection of resistant colonies. Plates were grown at 37°C to cure cells of the pKD46 plasmid. Colonies that grew on kanamycin were confirmed to have acquired the knockout cassette via colony PCR and DNA sequencing. These confirmed colonies were then transformed with pCP20 to remove the kanamycin resistance gene via Flp-FRT recombination. Transformants were plated on LB agar with 50 µg mL^{-1} carbenicillin. Following selection, colonies were grown in liquid culture at 42°C to cure cells of the pCP20 plasmid. Colonies were confirmed to have lost both *lpxM* and the knockout cassette via colony PCR and DNA sequencing and confirmed to have lost both kanamycin and carbenicillin resistance via replica plating on LB agar plates with 50 µg mL^{-1} carbenicillin or kanamycin. All primers used for construction and validation of this strain are listed in **Table 3.1**.

All plasmids used in the study are listed in **Table 3.2**. Plasmids pJL1-MBP[4xDQNAT], pJL1-PD[4xDQNAT], pJL1-PorA[4xDQNAT], pJL1-TTc[4xDQNAT], pJL1-TTlight[4xDQNAT], pJL1-CRM197[4xDQNAT], and pJL1-TT[4xDQNAT] were generated via PCR amplification and subsequent Gibson Assembly of a codon optimized gene construct purchased from IDT with a C-terminal 4xDQNAT-6xHis tag[135] between the *NdeI* and *SalI* restriction sites in the pJL1 vector. Plasmid pJL1-EPA[DNNNS-DQNRT] was constructed using the same approach, but without the addition of a C-terminal 4xDQNAT-6xHis tag. Plasmids pTrc99s-ssDsbA-MBP[4xDQNAT], pTrc99s-ssDsbA-PD[4xDQNAT], pTrc99s-ssDsbA-PorA[4xDQNAT], pTrc99s-ssDsbA-TTc[4xDQNAT],

pTrc99s-ssDsbA-TTlight[4xDQNAT], and pTrc99s-ssDsbA-EPA[DNNNS-DQNRT] were created via PCR amplification of each carrier protein gene and insertion into the pTrc99s vector between the *NcoI* and *HindIII* restriction sites via Gibson Assembly. Plasmid pSF-*Cj*PglB-LpxE was constructed using a similar approach, but via insertion of the *lpxE* gene from pE[136] between the *NdeI* and *NsiI* restriction sites in the pSF vector. Inserts were amplified via PCR using Phusion® High-Fidelity DNA polymerase (NEB) with forward and reverse primers designed using the NEBuilder® Assembly Tool (nebuilder.neb.com) and purchased from IDT. The pJL1 vector (Addgene 69496) was digested using restriction enzymes NdeI and SalI-HF® (NEB). The pSF vector was digested using restriction enzymes NdeI and NotI (NEB). PCR products were gel extracted using an EZNA Gel Extraction Kit (Omega Bio-Tek), mixed with Gibson assembly reagents and incubated at 50°C for 1 hour. Plasmid DNA from the Gibson assembly reactions were transformed into *E. coli* NEB 5-alpha cells and circularized constructs were selected using kanamycin at 50 μg ml^{-1} (Sigma). Sequence-verified clones were purified using an EZNA Plasmid Midi Kit (Omega Bio-Tek) for use in CFPS and iVAX reactions.

3.3.2 Cell-free lysate preparation

E. coli CLM24 source strains were grown in 2xYTP media (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, pH 7.2) in shake flasks (1 L scale) or a Sartorius Stedim BIOSTAT Cplus bioreactor (10 L scale) at 37°C. Protein synthesis yields and glycosylation activity were reproducible across different batches of lysate at both small and large scale. To generate *Cj*PglB-enriched lysate, CLM24 cells carrying plasmid pSF-*Cj*PglB[96] was used as the source strain. To generate *Ft*O-PS-enriched lysates, CLM24 carrying plasmid pGAB2[34] was used as the source strain. To generate one-pot lysates containing both *Cj*PglB and *Ft*O-PS, *Ec*O78-PS, or *Ec*O7-PS, CLM24 carrying pSF-*Cj*PglB and one of the following bacterial O-PS biosynthetic pathway plasmids was used as the source strain: pGAB2 (*Ft*O-PS), pMW07-O78 (*Ec*O78-PS), and pJHCV32 (*Ec*O7-PS). *Cj*PglB expression was induced at an OD_{600} of 0.8-1.0 with 0.02% (w/v) L-arabinose and cultures were moved to 30°C. Cells were grown to a final OD_{600} of ~3.0, at which point cells were pelleted by centrifugation at 5,000xg for 15 min at 4°C. Cell pellets were then washed three times with cold S30 buffer (10 mM Tris-acetate pH 8.2, 14 mM magnesium acetate, 60 mM potassium acetate) and pelleted at 5000xg for 10 min at 4°C. After

the final wash, cells were pelleted at 7000xg for 10 min at 4°C, weighed, flash frozen in liquid nitrogen, and stored at -80°C. To make cell lysate, cell pellets were resuspended to homogeneity in 1 mL of S30 buffer per 1 g of wet cell mass. Cells were disrupted via a single passage through an Avestin EmulsiFlex-B15 (1 L scale) or EmulsiFlex-C3 (10 L scale) high-pressure homogenizer at 20,000-25,000 psi. The lysate was then centrifuged twice at 30,000×g for 30 min to remove cell debris. Supernatant was transferred to clean microcentrifuge tubes and incubated at 37°C with shaking at 250 rpm for 60 min. Following centrifugation (15,000xg) for 15 min at 4°C, supernatant was collected, aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C. S30 lysate was active for about 3 freeze-thaw cycles and contained ~40 g/L total protein as measured by Bradford assay.

3.3.3 Cell-free protein synthesis

CFPS reactions were carried out in 1.5 mL microcentrifuge tubes (15 µL scale), 15 mL conical tubes (1 mL scale), or 50 mL conical tubes (5 mL scale) with a modified PANOx-SP system[101]. The CFPS reaction mixture consists of the following components: 1.2 mM ATP; 0.85 mM each of GTP, UTP, and CTP; 34.0 µg mL^{-1} L-5-formyl-5, 6, 7, 8-tetrahydrofolic acid (folinic acid); 170.0 µg mL^{-1} of *E. coli* tRNA mixture; 130 mM potassium glutamate; 10 mM ammonium glutamate; 12 mM magnesium glutamate; 2 mM each of 20 amino acids; 0.4 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); 57 mM HEPES; 13.3 µg mL^{-1} plasmid; and 27% v/v of cell lysate. For reaction volumes ≥1 mL, plasmid was added at 6.67 µg mL^{-1}, as this lower plasmid concentration conserved reagents with no effect on protein synthesis yields or kinetics (**Figure 3.2**). For expression of PorA, reactions were supplemented with nanodiscs at 1 µg mL^{-1}, which were prepared as previously described[97] or purchased (Cube Biotech). For expression of CRM197^{4xDQNAT}, CFPS was carried out at 25°C for 20 hours, unless otherwise noted. For all other carrier proteins, CFPS was run at 30°C for 20 hours, unless otherwise noted.
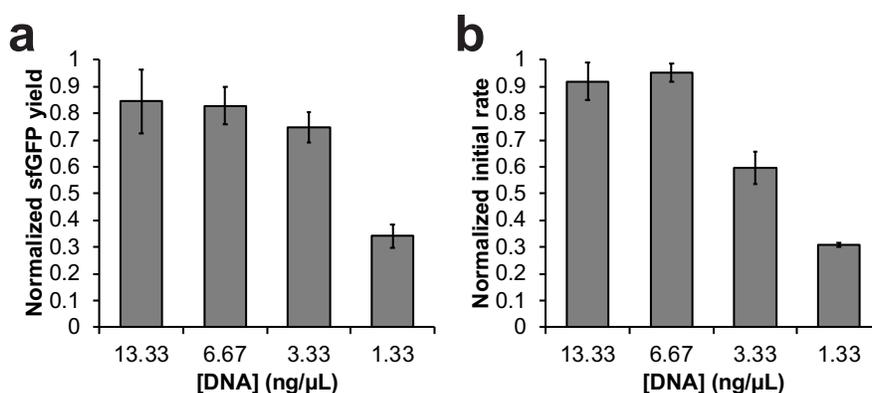
Figure 3.2 DNA concentration in iVAX reactions can be reduced without impacting protein synthesis yields or kinetics. iVAX reactions were prepared containing 13.33, 6.67, 3.33, or 1.33 ng/μL plasmid DNA template encoding sfGFP. We observed that both (**a**) protein synthesis yields after 20 hours and (**b**) initial rates of protein synthesis were conserved with 13.33 or 6.67 ng/μL DNA template. At lower DNA concentrations, DNA template appears to be limiting as lower protein synthesis yields and initial rates are observed.

For expression of TT[4xDQNAT], which contains intermolecular disulfide bonds, CFPS was carried out under oxidizing conditions. For oxidizing conditions, lysate was pre-conditioned with 750 μM iodoacetamide at room temperature for 30 min to covalently bind free sulfhydryls (-SH) and the reaction mix was supplemented with 200 mM glutathione at a 4:1 ratio of oxidized and reduced forms and 10 μM recombinant *E. coli* DsbC[137].

3.3.4 *In vitro* bioconjugate vaccine expression (iVAX)

For *in vitro* expression and glycosylation of carrier proteins in crude lysates, a two-phase scheme was implemented. In the first phase, CFPS was carried out for 15 min at 25-30°C as described above. In the second phase, protein glycosylation was initiated by the addition of $MnCl_2$ and DDM at a final concentration of 25 mM and 0.1% (w/v), respectively, and allowed to proceed at 30°C for a total reaction time of 1 hour. Protein synthesis yields and glycosylation activity were reproducible across biological replicates of iVAX reactions at both small and large scale. Reactions were then centrifuged at 20,000xg for 10 min to remove insoluble or aggregated protein products and the supernatant was analyzed by SDS-PAGE and Western blotting.

Purification of aglycosylated and glycosylated carriers from iVAX reactions was carried out using Ni-NTA agarose (Qiagen) according to manufacturer's protocols. Briefly, 0.5 mL Ni-NTA agarose per 1

mL cell-free reaction mixture was equilibrated in Buffer 1 (300 mM NaCl 50 mM NaH$_2$PO$_4$) with 10 mM imidazole. Soluble fractions from iVAX reactions were loaded on Ni-NTA agarose and incubated at 4°C for 2-4 hours to bind 6xHis-tagged protein. Following incubation, the cell-free reaction/agarose mixture was loaded onto a polypropylene column (BioRad) and washed twice with 6 column volumes of Buffer 1 with 20 mM imidazole. Protein was eluted in 4 fractions, each with 0.3 mL Buffer 1 with 300 mM imidazole per mL of cell-free reaction mixture. All buffers were used and stored at 4°C. Protein was stored at a final concentration of 1-2 mg mL^{-1} in sterile 1xPBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, pH 7.4) at 4°C.

3.3.5 Quantification of cell-free protein synthesis and iVAX yields

To quantify the amount of protein synthesized in iVAX reactions, two approaches were used. Fluorescence units of sfGFP were converted to concentrations using a previously reported standard curve[138]. Yields of all other proteins were assessed via the addition of 10 µM L-^{14}C-leucine (11.1 GBq mmol^{-1}, PerkinElmer) to the CFPS mixture to yield trichloroacetic acid-precipitable radioactivity that was measured using a liquid scintillation counter as described previously[139].

3.3.6 Expression of bioconjugates *in vivo* using protein-glycan coupling technology (PGCT)

Plasmids encoding bioconjugate carrier protein genes preceded by the DsbA leader sequence for translocation to the periplasm were transformed into CLM24 cells carrying pGAB2 and pSF-*Cj*PglB. CLM24 carrying only pGAB2 was used as a negative control. Transformed cells were grown in 5 mL LB media (10 g L^{-1} yeast extract, 5 g L^{-1} tryptone, 5 g L^{-1} NaCl) overnight at 37°C. The next day, cells were subcultured into 100 mL LB and allowed to grow at 37°C for 6 hours after which the culture was supplemented with 0.2% arabinose and 0.5 mM IPTG to induce expression of *Cj*PglB and the bioconjugate carrier protein, respectively. Protein expression was then carried out for 16 hours at 30°C, at which point cells were harvested. Cell pellets were resuspended in 1 mL sterile PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, pH 7.4) and lysed using a Q125 Sonicator (Qsonica, Newtown, CT) at 40% amplitude in cycles of 10 sec on/10 sec off for a total of 5 min. Soluble fractions

were isolated following centrifugation at 15,000 rpm for 30 min at 4°C. Protein was purified from soluble

fractions using Ni-NTA spin columns (Qiagen), following the manufacturer's protocol.

3.3.7 Western blot analysis

Samples were run on 4-12% Bis-Tris SDS-PAGE gels (Invitrogen). Following electrophoretic

separation, proteins were transferred from gels onto Immobilon-P polyvinylidene difluoride (PVDF)

membranes (0.45 μm) according to the manufacturer's protocol. Membranes were washed with PBS (80

g L^{-1} NaCl, 0.2 g L^{-1} KCl, 1.44 g L^{-1} Na$_2$HPO$_4$, 0.24 g L^{-1} KH$_2$PO$_4$, pH 7.4) followed by incubation for 1 hour

in Odyssey® Blocking Buffer (LI-COR). After blocking, membranes were washed 6 times with PBST (80 g

L^{-1} NaCl, 0.2 g L^{-1} KCl, 1.44 g L^{-1} Na$_2$HPO$_4$, 0.24 g L^{-1} KH$_2$PO$_4$, 1 mL L^{-1} Tween-20, pH 7.4) with a 5 min

incubation between each wash. For iVAX samples, membranes were probed with both an anti-6xHis tag

antibody and an anti-O-PS antibody or antisera specific to the O antigen of interest, if commercially

available. Probing of membranes was performed for at least 1 hour with shaking at room temperature,

after which membranes were washed with PBST in the same manner as described above and probed

with fluorescently labeled secondary antibodies. Membranes were imaged using an Odyssey® Fc

imaging system (LI-COR). CRM197 and TT production were compared to commercial DT and TT

standards (Sigma) and orthogonally detected by an identical SDS-PAGE procedure followed by Western

blot analysis with a polyclonal antibody that recognizes diphtheria or tetanus toxin, respectively. All

antibodies and dilutions used are listed in **Table 3.3**.

3.3.8 TLR4 activation assay

HEK-Blue hTLR4 cells (Invivogen) were maintained in DMEM media, high glucose/L-glutamine

supplement with 10% fetal bovine serum, 50 U mL^{-1} penicillin, 50 mg mL^{-1} streptomycin, and 100 μg mL^{-1}

Normacin^{TM} at 37°C in a humidified incubator containing 5% CO$_2$. After reaching ~50-80% confluency,

cells were plated into 96-well plates at a density of 1.4 × 10^5 cells per mL in HEK-Blue detection media

(Invivogen). Antigens were added at the following concentrations: 100 ng μL^{-1} purified protein; and 100 ng

μL^{-1} total protein in lysate. Purified *E. coli* O55:B5 LPS (Sigma-Aldrich) and detoxified *E. coli* O55:B5

(Sigma-Aldrich) were added at 1.0 ng mL^{-1} and served as positive and negative controls, respectively.

Plates were incubated at 37°C, 5% CO_2 for 10–16 h before measuring absorbance at 620 nm. Statistical significance was determined using paired *t*-tests.

3.3.9 Mouse immunization

Six groups of six-week old female BALB/c mice (Harlan Sprague Dawley) were injected subcutaneously with 100 μL PBS (pH 7.4) alone or containing purified aglycosylated MBP, *Ft*O-PS-conjugated MBP, aglycosylated PD, or *Ft*O-PS-conjugated PD, as previously described[140]. Groups were composed of six mice except for the PBS control group, which was composed of five mice. The amount of antigen in each preparation was normalized to 7.5 μg to ensure that an equivalent amount of aglycosylated protein or bioconjugate was administered in each case. Purified protein groups formulated in PBS were mixed with an equal volume of incomplete Freund's Adjuvant (Sigma-Aldrich) before injection. Prior to immunization, material for each group (5 μL) was streaked on LB agar plates and grown overnight at 37°C to confirm sterility and endotoxin activity was measured by TLR4 activation assay. Each group of mice was boosted with an identical dosage of antigen 21 days and 42 days after the initial immunization. Blood was obtained on day -1, 21, 35, 49, and 63 via submandibular collection and at study termination on day 70 via cardiac puncture. Mice were observed 24 and 48 hours after each injection for changes in behavior and physical health and no abnormal responses were reported. This study and all procedures were done in accordance with Protocol 2012-0132 approved by the Cornell University Institutional Animal Care and Use Committee.

3.3.10 Enzyme-linked immunosorbent assay (ELISA)

F. tularensis LPS-specific antibodies elicited by immunized mice were measured via indirect ELISA using a modification of a previously described protocol[140]. Briefly, sera were isolated from the collected blood draws after centrifugation at 5000xg for 10 min and stored at −20 °C; 96-well plates (Maxisorp; Nunc Nalgene) were coated with *F. tularensis* LPS (BEI resources) at a concentration of 5 μg mL^{-1} in PBS and incubated overnight at 4°C. The next day, plates were washed three times with PBST (PBS, 0.05% Tween-20, 0.3% BSA) and blocked overnight at 4°C with 5% nonfat dry milk (Carnation) in PBS. Samples were serially diluted by a factor of two in triplicate between 1:100 and 1:12,800,000 in blocking buffer and added to the plate for 2 hours at 37°C. Plates were washed three times with PBST

and incubated for 1 hour at 37°C in the presence of one of the following HRP-conjugated antibodies (all from Abcam and used at 1:25,000 dilution): goat anti-mouse IgG, anti-mouse IgG1, and anti-mouse IgG2a. After three additional washes with PBST, 3,3′-5,5′-tetramethylbenzidine substrate (1-Step Ultra TMB-ELISA; Thermo-Fisher) was added, and the plate was incubated at room temperature in the dark for 30 min. The reaction was halted with 2 M H_2SO_4, and absorbance was quantified via microplate spectrophotometer (Tecan) at a wavelength of 450 nm. Serum antibody titers were determined by measuring the lowest dilution that resulted in signal 3 SDs above no serum background controls. Statistical significance was determined in RStudio 1.1.463 using one-way ANOVA and the Tukey–Kramer *post hoc* honest significant difference test.

3.3.11   Statistical analysis

Statistical parameters including the definitions and values of *n*, *p*-values, standard deviations, and standard errors are reported in the figures and corresponding figure legends. Analytical techniques are described in the corresponding Method Details section.

3.4   Results

3.4.1 *In vitro* synthesis of licensed vaccine carrier proteins

To demonstrate proof-of-principle for cell-free bioconjugate vaccine production, we first set out to express a set of carrier proteins that are currently used in FDA-approved conjugate vaccines. Producing these carrier proteins in soluble conformations *in vitro* represented an important benchmark because their expression in living *E. coli* has proven challenging, often requiring purification and refolding of insoluble product from inclusion bodies[141,142], fusion of expression partners such as maltose-binding protein (MBP) to increase soluble expression[142,143], or expression of truncated protein variants in favor of the full-length proteins[143]. In contrast, cell-free protein synthesis approaches have recently shown promise for difficult-to-express proteins[44]. The carrier proteins that we focused on here included nonacylated *H. influenzae* protein D (PD), the *N. meningitidis* porin protein (PorA), and genetically detoxified variants of the *Corynebacterium diphtheriae* toxin (CRM197) and the *Clostridium tetani* toxin (TT). We also tested expression of the fragment C (TTc) and light chain (TTlight) domains of TT as well as *E. coli* MBP. While MBP is not a licensed carrier, it has demonstrated immunostimulatory properties[144] and when linked to O-

PS was found to elicit polysaccharide-specific humoral and cellular immune responses in mice[36]. Similarly, the TT domains, TTlight and TTc, have not been used in licensed vaccines, but are immunostimulatory and individually sufficient for protection against *C. tetani* challenge in mice[143]. To enable glycosylation, all carriers were modified at their C-termini with 4 tandem repeats of an optimal bacterial glycosylation motif, DQNAT[145]. A C-terminal 6xHis tag was also included to enable purification and detection via Western blot analysis. A variant of superfolder green fluorescent protein that contained an internal DQNAT glycosylation site (sfGFP$^{217\text{-}DQNAT}$)[134] was used as a model protein to facilitate system development.

All eight carriers were synthesized *in vitro* with soluble yields of ~50-650 µg mL^{-1} as determined by ^{14}C-leucine incorporation (**Figure 3.3a**). In particular, the MBP^{4xDQNAT} and PD^{4xDQNAT} variants were nearly 100% soluble, with yields of 500 µg mL^{-1} and 200 µg mL^{-1}, respectively, and expressed as exclusively full-length products according to Western blot analysis (**Figure 3.3b**). Notably, similar soluble yields were observed for all carriers at 25°C, 30°C, and 37°C, with the exception of CRM197^{4xDQNAT} (**Figure 3.4a**), which is known to be heat sensitive[38]. These results suggest that our method of cell-free carrier biosynthesis is robust over a 13°C range in temperature and could be used in settings where precise temperature control is not feasible.
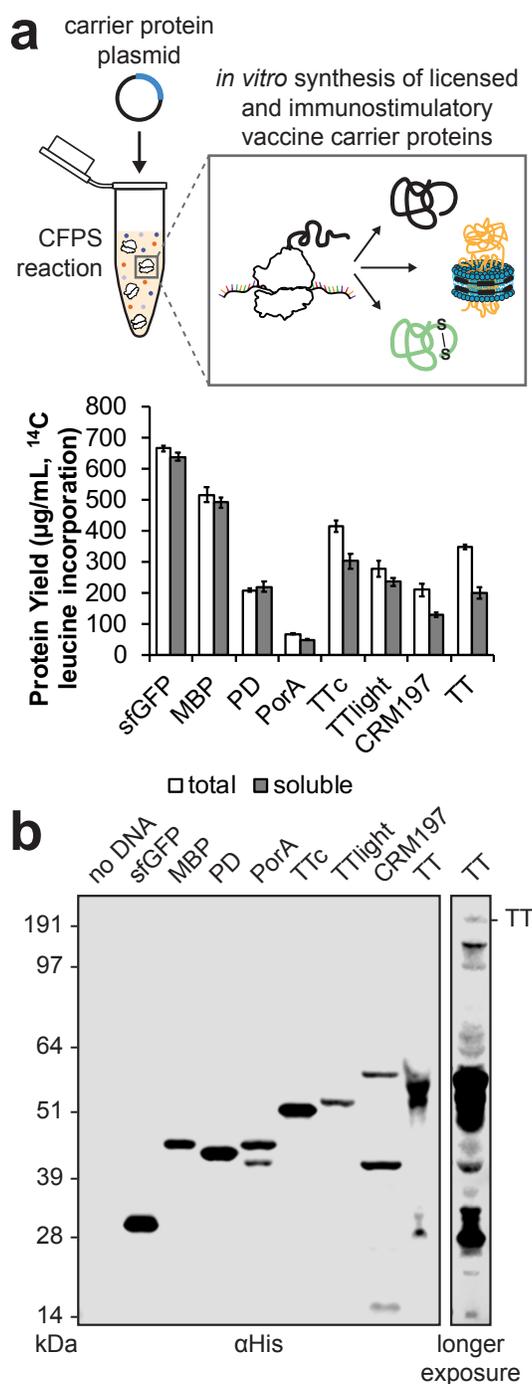
Figure 3.3 *In vitro* synthesis of licensed conjugate vaccine carrier proteins. (**a**) All four carrier proteins used in FDA-approved conjugate vaccines were synthesized solubly *in vitro*, as measured via ^{14}C-leucine incorporation. These include *H. influenzae* protein D (PD), the *N. meningitidis* porin protein (PorA), and genetically detoxified variants of the *C. diphtheriae* toxin (CRM197) and the *C. tetani* toxin (TT). Additional immunostimulatory carriers were also synthesized solubly, including *E. coli* maltose binding protein (MBP) and the fragment C (TTc) and light chain (TTlight) domains of TT. Values represent means and error bars represent standard deviations of biological replicates (*n* = 3). (**b**) Full length product was observed for all proteins tested via Western blot. Different exposures are indicated with solid lines. Molecular weight ladder is shown at left.

The open reaction environment of our cell-free reactions enabled facile manipulation of the chemical and reaction environment to improve production of more complex carriers. For example, in the case of the membrane protein PorA[4xDQNAT], lipid nanodiscs were added to increase soluble expression (**Figure 3.4b**). Nanodiscs provide a cellular membrane mimic to co-translationally stabilize hydrophobic regions of membrane proteins[146]. For expression of TT, which contains an intermolecular disulfide bond, expression was carried out for 2 hours in oxidizing conditions[137], which improved assembly of the heavy and light chains into full-length product and minimized protease degradation of full-length TT (**Figure 3.4c**). *In vitro* synthesized CRM197[4xDQNAT] and TT[4xDQNAT] were comparable in size to commercially available purified diphtheria toxin (DT) and TT protein standards and were reactive with α-DT and α-TT antibodies, respectively (**Figure 3.4d, e**), indicating that both were produced in immunologically relevant conformations. This is notable as CRM197 and TT are FDA-approved vaccine antigens for diphtheria and tetanus, respectively, when they are administered without conjugated polysaccharides. Together, our results highlight the ability of CFPS to express licensed conjugate vaccine carrier proteins in soluble conformations over a range of temperatures.
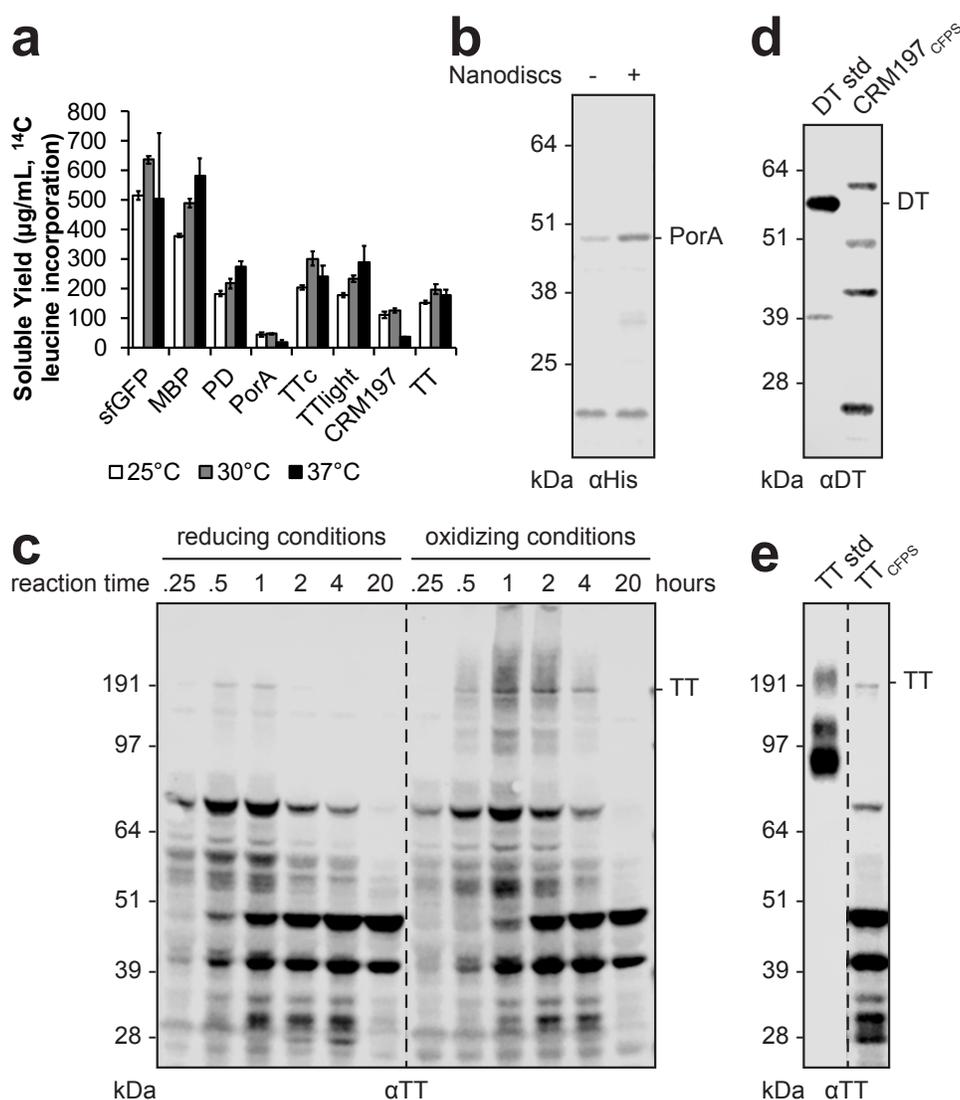
Figure 3.4 *In vitro* synthesis of licensed conjugate vaccine carrier proteins is possible over a range of temperatures and can be readily optimized. (**a**) With the exception of CRM197, all carriers expressed with similar soluble yields at 25°C, 30°C, and 37°C, as measured by ^{14}C-leucine incorporation. Values represent means and error bars represent standard deviations of biological replicates (*n* = 3). (**b**) Soluble expression of PorA was improved through the addition of lipid nanodiscs to the reaction. (**c**) Expression of full-length TT was enhanced by (i) performing *in vitro* protein synthesis in oxidizing conditions to improve assembly of the disulfide-bonded heavy and light chains into full-length TT and (ii) allowing reactions to run for only 2 h to minimize protease degradation. (**d**) CRM197 and (**e**) TT produced in CFPS reactions are detected with α-DT and α-TT antibodies, respectively, and are comparable in size to commercially available purified DT and TT protein standards (50 ng standard loaded). Images are representative of at least three biological replicates. Dashed line indicates samples are from the same blot with the same exposure. Molecular weight ladders are shown at the left of each image.

3.4.2 On-demand synthesis of bioconjugate vaccines

We next sought to synthesize polysaccharide-conjugated versions of these carrier proteins by merging their *in vitro* expression with one-pot, cell-free glycosylation. As a model vaccine target, we focused on the highly virulent *Francisella tularensis* subsp. *tularensis* (type A) strain Schu S4, a gram-negative, facultative coccobacillus and the causative agent of tularemia. This bacterium is categorized as a class A bioterrorism agent due to its high fatality rate, low dose of infection, and ability to be aerosolized[147]. Although there are currently no licensed vaccines against *F. tularensis*, several studies have independently confirmed the important role of antibodies directed against *F. tularensis* LPS, specifically the O-PS repeat unit, in providing protection against the Schu S4 strain[148,149]. More recently, a bioconjugate vaccine comprising the *F. tularensis* Schu S4 O-PS (*Ft*O-PS) conjugated to the *Pseudomonas aeruginosa* exotoxin A (EPA[DNNNS-DQNRT]) carrier protein produced using PGCT[34,131] was shown to be protective against challenge with the Schu S4 strain in a rat inhalation model of tularemia[131]. In light of these earlier findings, we investigated the ability of the iVAX platform to produce anti-*F. tularensis* bioconjugate vaccine candidates on-demand by conjugating the *Ft*O-PS structure to diverse carrier proteins *in vitro*.

The *Ft*O-PS is composed of the 826-Da repeating tetrasaccharide unit Qui4NFm-(GalNAcAN)$_2$-QuiNAc (Qui4NFm: 4,6-dideoxy-4-formamido-D-glucose; GalNAcAN: 2-acetamido-2-deoxy-D-galacturonamide; QuiNAc: 2-acetamido-2,6-dideoxy-D-glucose)[150]. To glycosylate proteins with *Ft*O-PS, we produced an iVAX lysate from glycoengineered *E. coli* cells expressing the *Ft*O-PS biosynthetic pathway and the oligosaccharyltransferase enzyme *Cj*PglB (**Figure 3.5a**). This lysate, which contained lipid-linked *Ft*O-PS and active *Cj*PglB, was used to catalyze iVAX reactions primed with plasmid DNA encoding sfGFP[217-DQNAT]. Control reactions in which attachment of the *Ft*O-PS was not expected were performed with lysates from cells that lacked either the *Ft*O-PS pathway or the *Cj*PglB enzyme. We also tested reactions that lacked plasmid encoding the target protein sfGFP[217-DQNAT] or were primed with plasmid encoding sfGFP[217-AQNAT], which contained a mutated glycosylation site (AQNAT) that is not modified by *Cj*PglB[151]. In reactions containing the iVAX lysate and primed with plasmid encoding sfGFP[217-DQNAT], immunoblotting with anti-His antibody or a commercial monoclonal antibody specific to

*Ft*O-PS revealed a ladder-like banding pattern (**Figure 3.5b**). This ladder is characteristic of *Ft*O-PS attachment, resulting from O-PS chain length variability through the action of the Wzy polymerase[22,34,150]. Glycosylation of sfGFP[217-DQNAT] was observed only in reactions containing a complete glycosylation pathway and the preferred DQNAT glycosylation sequence (**Figure 3.5b**). This glycosylation profile was reproducible across biological replicates from the same lot of lysate (**Figure 3.5c, left**) and using different lots of lysate (**Figure 3.5c, right**). *In vitro* protein synthesis and glycosylation was observed after 1 hour, with the amount of conjugated polysaccharide reaching a maximum between 0.75 and 1.25 hours (**Figure 3.6**). Similar glycosylation reaction kinetics were observed at 37°C, 30°C, 25°C, and room temperature (~21°C), indicating that iVAX reactions are robust over a range of temperatures (**Figure 3.6**).
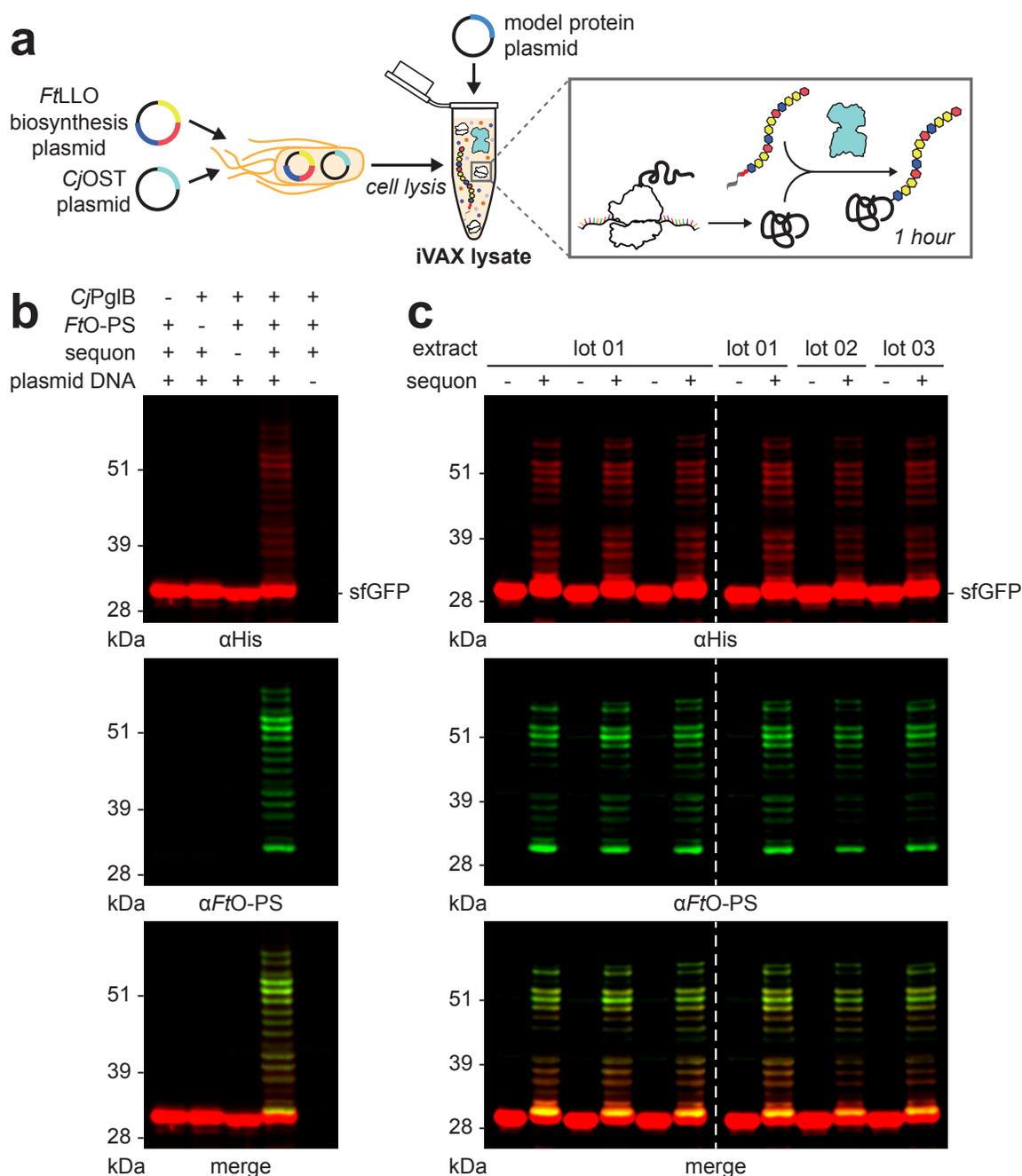
Figure 3.5 Reproducible glycosylation of proteins with *Ft*O-PS in iVAX lysates. (**a**) iVAX lysates were prepared from cells expressing *Cj*PglB and a biosynthetic pathway encoding *Ft*O-PS. (**b**) Glycosylation of sfGFP$^{217\text{-DQNAT}}$ with *Ft*O-PS was only observed when *Cj*PglB, *Ft*O-PS, and the preferred sequon were present in the reaction (lane 3). When plasmid DNA was omitted, sfGFP$^{217\text{-DQNAT}}$ synthesis was not observed. (**c**) Biological replicates of iVAX reactions producing sfGFP$^{217\text{-DQNAT}}$ using the same lot (**left**) or different lots (**right**) of iVAX lysates demonstrated reproducibility of reactions and lysate preparation. Top panels show signal from probing with anti-hexa-histidine antibody (αHis) to detect the carrier protein, middle panels show signal from probing with commercial anti-*Ft*O-PS antibody (α*Ft*O-PS), and bottom panels show αHis and α*Ft*O-PS signals merged. Unless replicates are explicitly shown, images are representative of at least three biological replicates. Dashed lines indicate samples are from the same blot with the same exposure. Molecular weight ladders are shown at the left of each image.
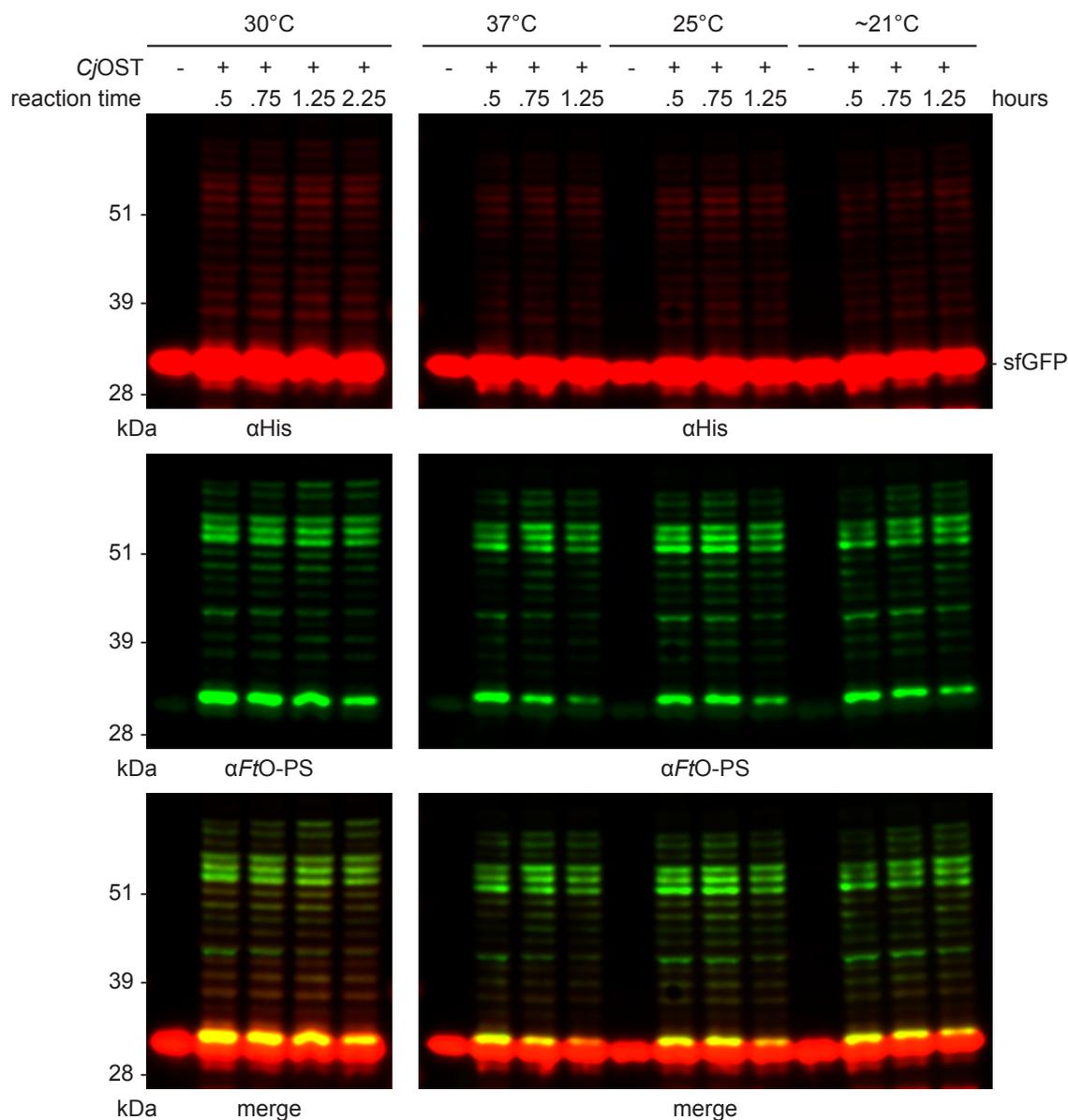
Figure 3.6 Glycosylation in iVAX reactions occurs in 1 h over a range of temperatures. Kinetics of *Ft*O-PS glycosylation at 30°C (**left**), 37°C, 25°C, and room temperature (~21°C) (**right**) are comparable and show that protein synthesis and glycosylation occur in the first hour of the iVAX reaction. These results demonstrate that the iVAX platform can synthesize bioconjugates over a range of permissible temperatures. Top panels show signal from probing with anti-hexa-histidine antibody (αHis) to detect the carrier protein, middle panels show signal from probing with commercial anti-*Ft*O-PS antibody (α*Ft*O-PS), and bottom panels show αHis and α*Ft*O-PS signals merged. Images are representative of at least three biological replicates. Molecular weight ladders are shown at the left of each image.

Next, we investigated whether FDA-approved carriers could be similarly conjugated with *Ft*O-PS in iVAX reactions. Following addition of plasmid DNA encoding MBP[4xDQNAT], PD[4xDQNAT], PorA[4xDQNAT],

TTc[4xDQNAT], TTlight[4xDQNAT], and CRM197[4xDQNAT], glycosylation of each with FtO-PS was observed for iVAX reactions enriched with lipid-linked FtO-PS and CjPglB but not control reactions lacking CjPglB (**Figure 3.7**). We observed conjugation of high molecular weight FtO-PS species (on the order of ~10-20 kDa) to all protein carriers tested, which is important as longer glycan chain length has been shown to increase the efficacy of conjugate vaccines against *F. tularensis*[152]. Notably, our attempts to synthesize the same panel of bioconjugates using the established PGCT approach in living *E. coli* yielded less promising results. Specifically, we observed low levels of FtO-PS glycosylation and lower molecular weight conjugated FtO-PS species for all PGCT-derived bioconjugates compared to their iVAX-derived counterparts (**Figure 3.8**). The same trend was observed for PGCT- versus iVAX-derived bioconjugates involving the most common PGCT carrier protein, EPA[DNNNS-DQNRT33,34,37,131,132] (**Figure 3.7**; **Figure 3.8**). In addition, only limited expression of the PorA membrane protein was achieved *in vivo* (**Figure 3.8**). Collectively, these data indicate that iVAX could provide advantages over PGCT for production of bioconjugate vaccine candidates with high molecular weight O-PS antigens conjugated efficiently to diverse and potentially membrane-bound carrier proteins.

Figure 3.7 On-demand production of bioconjugates against *F. tularensis* using iVAX. (**a**) iVAX reactions were prepared from lysates containing *Cj*PglB and *Ft*O-PS and primed with plasmid encoding immunostimulatory carriers, including those used in licensed vaccines. (**b**) We observed on-demand synthesis of anti-*F. tularensis* bioconjugate vaccines for all carrier proteins tested. Bioconjugates were purified using Ni-NTA agarose from 1 mL iVAX reactions lasting ~1 h. Top panels show signal from probing with anti-hexa-histidine antibody (αHis) to detect the carrier protein, middle panels show signal from probing with commercial anti-*Ft*O-PS antibody (α*Ft*O-PS), and bottom panels show αHis and α*Ft*O-PS signals merged. Images are representative of at least three biological replicates. Dashed lines indicate samples are from the same blot with the same exposure. Molecular weight ladders are shown at the left of each image.

Figure 3.8 Production of bioconjugates against *F. tularensis* using PGCT in living *E. coli*. (**a**) Bioconjugates were produced via PGCT in CLM24 cells expressing *Cj*PglB, the biosynthetic pathway for *Ft*O-PS, and a panel of immunostimulatory carriers including those used in licensed vaccines. (**b**) We observed low expression of PorA, a membrane protein, as well as reduced glycan loading and conjugation of high molecular weight *Ft*O-PS species in all carriers compared to iVAX-derived samples. Top panels show signal from probing with anti-hexa-histidine antibody (αHis) to detect the carrier protein, middle panels show signal from probing with commercial anti-*Ft*O-PS antibody (α*Ft*O-PS), and bottom panels show αHis and α*Ft*O-PS signals merged. Images are representative of at least three biological replicates. Molecular weight ladders are shown at the left of each image.

We next asked whether the yields of bioconjugates produced using iVAX were sufficient to enable production of relevant vaccine doses. Recent clinical data show 1-10 µg doses of bioconjugate vaccine candidates are well-tolerated and effective in stimulating the production of antibacterial IgGs[153-155]. To assess expression titers and for future experiments, we focused on MBP[4xDQNAT] and PD[4xDQNAT] because these carriers expressed *in vitro* with high soluble titers and without truncation products (**Figure 3.3**). We found that reactions lasting ~1 hour produced ~20 µg mL^{-1} of glycosylated MBP[4xDQNAT] and PD[4xDQNAT] as determined by ^{14}C-leucine incorporation and densitometry analysis (**Figure 3.9a**). It should be noted that vaccines are currently distributed in vials containing 1-20 doses of vaccine to minimize wastage[156]. Our yields suggest that multiple doses per mL can be synthesized in 1 hour using the iVAX platform.

To demonstrate the modularity of the iVAX approach for bioconjugate production, we sought to produce bioconjugates bearing O-PS antigens from additional pathogens including ETEC *E. coli* strain O78 and UPEC *E. coli* strain O7. *E. coli* O78 is a major cause of diarrheal disease in developing countries, especially among children, and a leading cause of traveler's diarrhea[157], while the O7 strain is a common cause of urinary tract infections[158]. Like the *Ft*O-PS, the biosynthetic pathways for *Ec*O78-PS and *Ec*O7-PS have been described previously and confirmed to produce O-PS antigens with the repeating units GlcNAc$_2$Man$_2$[159] and Qui4NAcMan(Rha)GalGlcNAc[160] (GlcNAc: *N*-acetylglucosamine; Man: mannose; Qui4NAc: 4-acetamido-4,6-dideoxy-D-glucopyranose; Rha: rhamnose; Gal: galactose), respectively. Using iVAX lysates from cells expressing *Cj*PglB and either the *Ec*O78-PS and *Ec*O7-PS pathways in reactions that were primed with PD[4xDQNAT] or sfGFP[217-DQNAT] plasmids, we observed carrier glycosylation only when both lipid-linked O-PS and *Cj*PglB were present in the reactions (**Figure 3.9b, c**). Our results demonstrate modular production of bioconjugates against multiple bacterial pathogens, enabled by compatibility of multiple heterologous O-PS pathways with *in vitro* carrier protein synthesis and glycosylation.
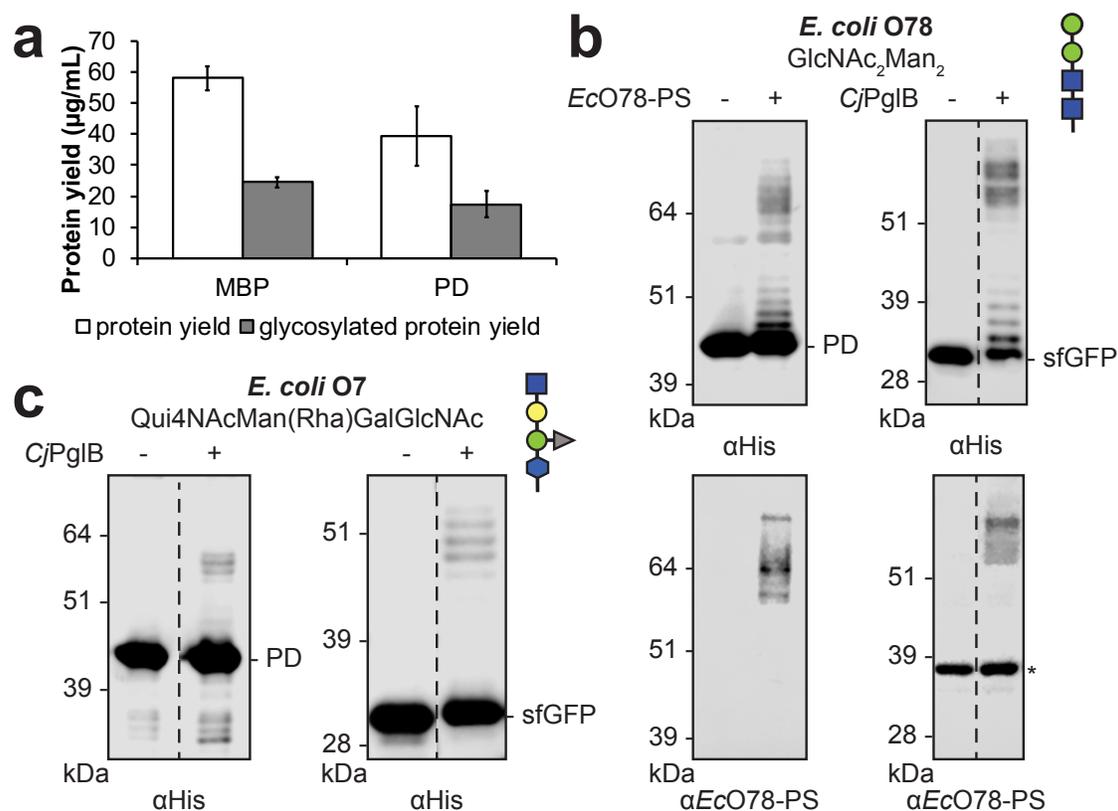
Figure 3.9 The iVAX platform is modular and can be used to synthesize clinically relevant yields of diverse bioconjugates. (**a**) Protein synthesis and glycosylation with *Ft*O-PS were measured in iVAX reactions producing MBP[4xDQNAT] and PD[4xDQNAT]. After ~1 h, reactions produced ~40 µg mL^{-1} protein, as measured via ^{14}C-leucine incorporation, of which ~20 µg mL^{-1} was glycosylated with *Ft*O-PS, as determined by densitometry. Values represent means and error bars represent standard errors of biological replicates (*n* = 2). To demonstrate modularity, iVAX lysates were prepared from cells expressing *Cj*PglB and biosynthetic pathways for either (**b**) the *E. coli* O78 antigen or (**c**) the *E. coli* O7 antigen and used to synthesize PD[4xDQNAT] (**left**) or sfGFP[217-DQNAT] (**right**) bioconjugates. The structure and composition of the repeating monomer unit for each antigen is shown. Both polysaccharide antigens are compositionally and, in the case of the O7 antigen, structurally distinct compared to the *F. tularensis* O antigen. Blots show signal from probing with anti-hexa-histidine antibody (αHis) to detect the carrier protein. If a commercial anti-O-PS serum or antibody was available, it was used to confirm the identity of the conjugated O antigen (α-*Ec*O78 blots, panel **b**). Asterisk denotes bands resulting from non-specific serum antibody binding. Images are representative of at least three biological replicates. Dashed lines indicate samples are from the same blot with the same exposure. Molecular weight ladders are shown at the left of each image.

3.4.3 Endotoxin editing and freeze-drying yield iVAX reactions that are safe and portable

A key challenge inherent in using any *E. coli*-based system for biopharmaceutical production is the presence of lipid A, or endotoxin, which is known to contaminate protein products and can cause lethal septic shock at high levels[161]. As a result, the amount of endotoxin in formulated

biopharmaceuticals is regulated by the United States Pharmacopeia (USP), US Food and Drug Administration (FDA), and the European Medicines Agency (EMEA)[162]. Because our iVAX reactions rely on lipid-associated components, such as *Cj*PglB and *Ft*O-PS, standard detoxification approaches involving the removal of lipid A[163] could compromise the activity or concentration of our glycosylation components in addition to increasing cost and processing complexities.

To address this issue, we adapted a previously reported strategy to detoxify the lipid A molecule through strain engineering[105,136]. In particular, the deletion of the acyltransferase gene *lpxM* and the overexpression of the *F. tularensis* phosphatase LpxE in *E. coli* has been shown to result in the production of nearly homogenous pentaacylated, monophosphorylated lipid A with significantly reduced toxicity but retained activity as an adjuvant[105]. This pentaacylated, monophosphorylated lipid A was structurally identical to the primary component of monophosphoryl lipid A (MPL) from *Salmonella minnesota* R595, an approved adjuvant composed of a mixture of monophosphorylated lipids[164]. To generate detoxified lipid A structures in the context of iVAX, we produced lysates from a Δ*lpxM* derivative of CLM24 that co-expressed *Ft*LpxE and the *Ft*O-PS glycosylation pathway (**Figure 3.10a**). Lysates derived from this strain exhibited significantly decreased levels of toxicity compared to wild type CLM24 lysates expressing *Cj*PglB and *Ft*O-PS (**Figure 3.10b**) as measured by human TLR4 activation in HEK-Blue hTLR4 reporter cells[136]. Importantly, the structural remodeling of lipid A did not affect the activity of the membrane-bound *Cj*PglB and *Ft*O-PS components in iVAX reactions (**Figure 3.11a**). By engineering the chassis strain for lysate production, we produced iVAX lysates with endotoxin levels <1,000 EU mL^{-1}, within the range of reported values for commercial protein-based vaccine products (0.288-180,000 EU mL^{-1})[162].
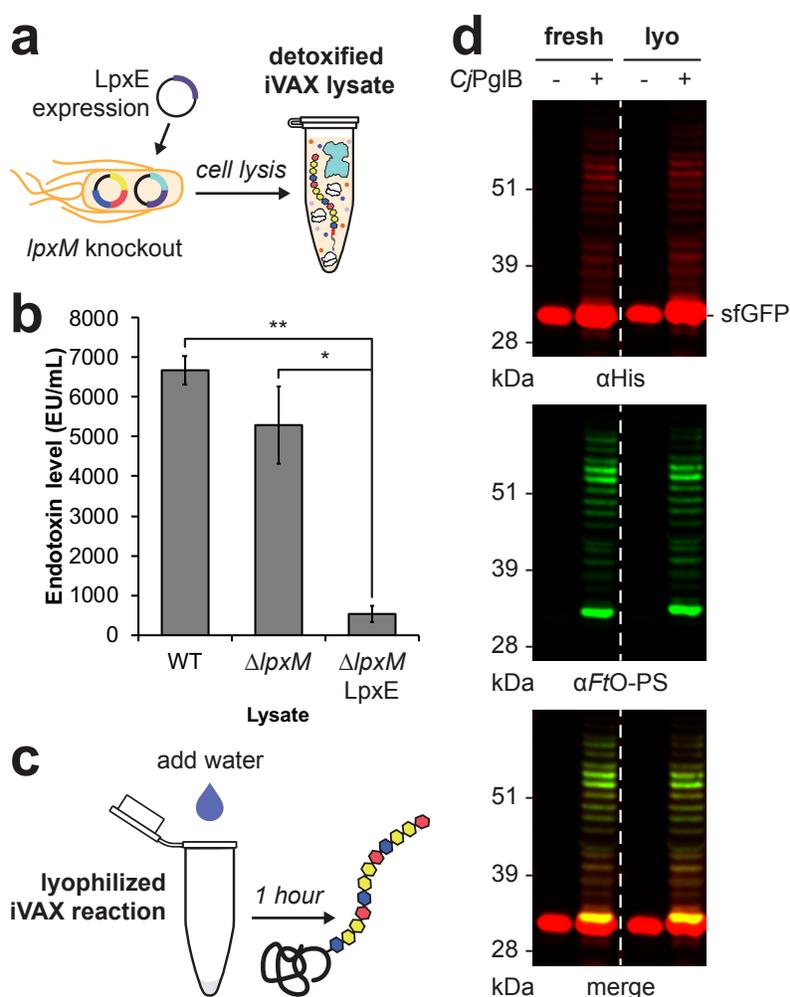
Figure 3.10 Detoxified, lyophilized iVAX reactions produce bioconjugates. (**a**) iVAX lysates were detoxified via deletion of *lpxM* and expression of *F. tularensis* LpxE in the source strain for lysate production. (**b**) The resulting lysates exhibited significantly reduced endotoxin activity. *p = 0.019 and **p = 0.003, as determined by two-tailed *t*-test. (**c**) iVAX reactions producing sfGFP[217-DQNAT] were run immediately or following lyophilization and rehydration. (**d**) Glycosylation activity was preserved following lyophilization, demonstrating the potential of iVAX reactions for portable biosynthesis of bioconjugate vaccines. Top panel shows signal from probing with anti-hexa-histidine antibody (αHis) to detect the carrier protein, middle panel shows signal from probing with commercial anti-*Ft*O-PS antibody (α*Ft*O-PS), and bottom panel shows αHis and α*Ft*O-PS signals merged. Images are representative of at least three biological replicates. Molecular weight ladder is shown at the left of each image.

A major limitation of traditional conjugate vaccines is that they must be refrigerated[38], making it difficult to distribute these vaccines to remote or resource-limited settings. The ability to freeze-dry iVAX reactions for ambient temperature storage and distribution could alleviate the logistical challenges associated with refrigerated supply chains that are required for existing vaccines. To investigate this

possibility, detoxified iVAX lysates were used to produce *Ft*O-PS bioconjugates in two different ways: either by running the reaction immediately after priming with plasmid encoding the sfGFP[217-DQNAT] target protein or by running after the same reaction mixture was lyophilized and rehydrated (**Figure 3.10c**). In both cases, conjugation of *Ft*O-PS to sfGFP[217-DQNAT] was observed when *Cj*PglB was present, with modification levels that were nearly identical (**Figure 3.10d**). We also showed that detoxified, freeze-dried iVAX reactions can be scaled to 5 mL for production of *Ft*O-PS-conjugated MBP[4xDQNAT] and PD[4xDQNAT] in a manner that was reproducible from lot to lot and indistinguishable from production without freeze-drying (**Figure 3.11b, c**). The ability to lyophilize iVAX reactions and manufacture bioconjugates without specialized equipment highlights the potential for portable, on-demand vaccine production.
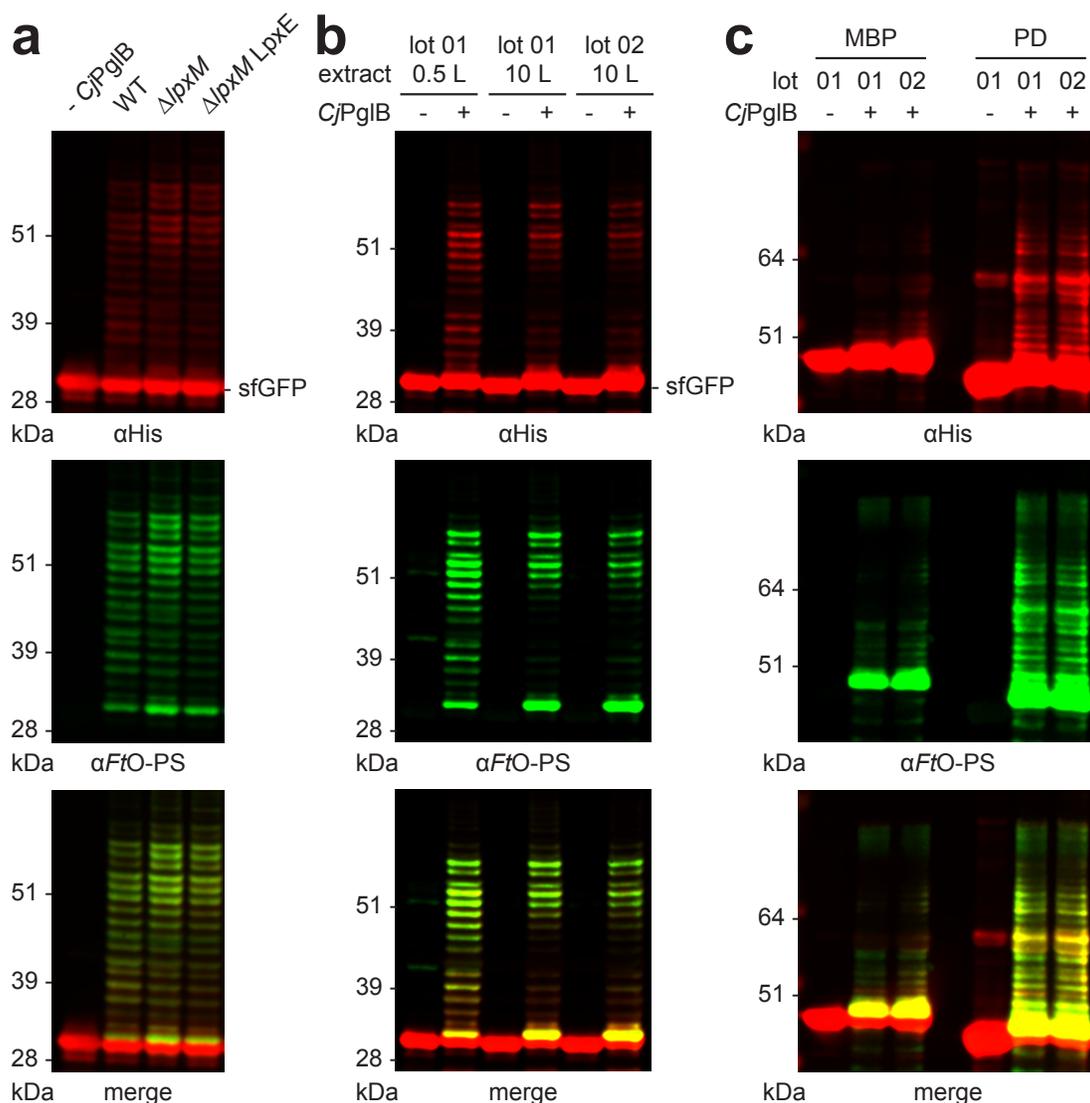
Figure 3.11 Detoxified iVAX lysates synthesize bioconjugates and both lysate production and freeze-dried reactions scale reproducibly. (**a**) iVAX lysates containing *Cj*PglB and *Ft*O-PS were prepared from wild-type CLM24, CLM24 Δ*lpxM*, or CLM24 Δ*lpxM* cells expressing *Ft*LpxE. Nearly identical sfGFP²¹⁷⁻ᴰᵠᴺᴬᵀ glycosylation was observed for each of the lysates derived from the engineered strains. (**b**) To generate material for immunizations, fermentations to produce endotoxin-edited iVAX lysates were scaled from 0.5 L to 10 L. We observed similar levels of sfGFP²¹⁷⁻ᴰᵠᴺᴬᵀ glycosylation for lysates derived from 0.5 L and 10 L cultures, and across different batches of lysate produced from 10 L fermentations. (**c**) For immunizations, we prepared two lots of *Ft*O-PS-conjugated MBP⁴ˣᴰᵠᴺᴬᵀ and PD⁴ˣᴰᵠᴺᴬᵀ from 5 mL freeze-dried iVAX reactions. We observed similar levels of purified protein (~200 µg) and *Ft*O-PS modification (>50%, measured by densitometry) across both carriers and lots of material. Top panels show signal from probing with anti-hexa-histidine antibody (αHis) to detect the carrier protein, middle panels show signal from probing with commercial anti-*Ft*O-PS antibody (α*Ft*O-PS), and bottom panels show αHis and α*Ft*O-PS signals merged. Images are representative of at least three biological replicates. Molecular weight ladders are shown at left.

3.4.4 *In vitro* synthesized bioconjugates elicit pathogen-specific antibodies in mice

To validate the efficacy of bioconjugates produced using the iVAX platform, we next evaluated the ability of iVAX-derived bioconjugates to elicit anti-*Ft*LPS antibodies in mice (**Figure 3.12a**). Importantly, we found that BALB/c mice receiving iVAX-derived *Ft*O-PS-conjugated MBP^{4xDQNAT} or PD^{4xDQNAT} produced high titers of *Ft*LPS-specific IgG antibodies, which were significantly elevated compared to the titers measured in the sera of control mice receiving PBS or aglycosylated versions of each carrier protein (**Figure 3.12b, Figure 3.13**). Interestingly, the IgG titers measured in sera from mice receiving glycosylated MBP^{4xDQNAT} derived from PGCT were similar to the titers observed in the control groups (**Figure 3.12b, Figure 3.13**), in line with the markedly weaker glycosylation of this candidate relative to its iVAX-derived counterpart (**Figure 3.8**). Notably, both MBP^{4xDQNAT} and PD^{4xDQNAT} bioconjugates produced using iVAX elicited similar levels of IgG production and neither resulted in any observable adverse events in mice, confirming the modularity and safety of the technology for production of bioconjugate vaccine candidates.
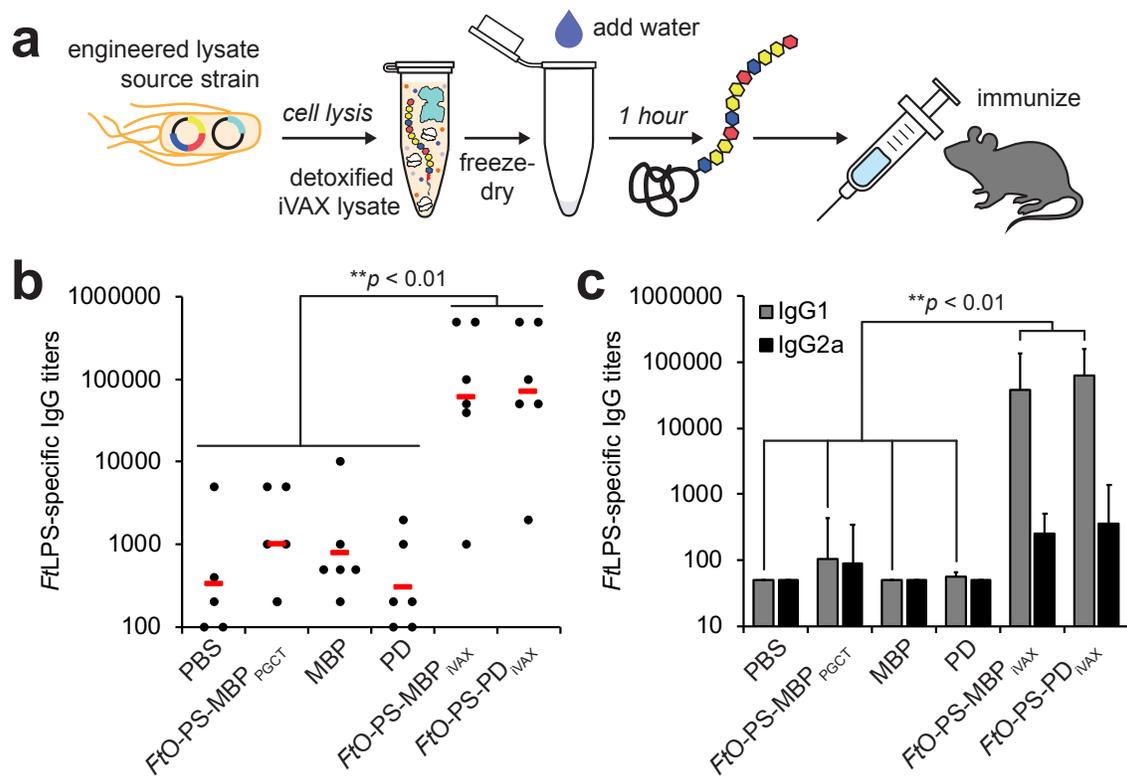
Figure 3.12 iVAX-derived bioconjugates elicit *Ft*LPS-specific antibodies in mice. (**a**) Freeze-dried iVAX reactions assembled using detoxified lysates were used to synthesize anti-*F. tularensis* bioconjugates for immunization studies. (**b**) Six groups of BALB/c mice were immunized subcutaneously with PBS or 7.5 µg of purified, cell-free synthesized aglycosylated MBP^{4xDQNAT}, *Ft*O-PS-conjugated MBP^{4xDQNAT}, aglycosylated PD^{4xDQNAT}, or *Ft*O-PS-conjugated PD^{4xDQNAT}. *Ft*O-PS-conjugated MBP^{4xDQNAT} prepared in living *E. coli* cells using PGCT was used as a positive control. Each group was composed of six mice except for the PBS control group, which was composed of five mice. Mice were boosted on days 21 and 42 with identical doses of antigen. *Ft*LPS-specific IgG titers were measured by ELISA in endpoint (day 70) serum of individual mice (black dots) with *F. tularensis* LPS immobilized as antigen. Mean titers of each group are also shown (red lines). iVAX-derived bioconjugates elicited significantly higher levels of *Ft*LPS-specific IgG compared to all other groups (**p < 0.01, Tukey-Kramer HSD). (**c**) IgG1 and IgG2a subtype titers measured by ELISA from endpoint serum revealed that iVAX-derived bioconjugates boosted production of *Ft*O-PS-specific IgG1 compared to all other groups tested (**p < 0.01, Tukey-Kramer HSD). These results indicate that iVAX bioconjugates elicited a Th2-biased immune response typical of most conjugate vaccines. Values represent means and error bars represent standard errors of *Ft*LPS-specific IgGs detected by ELISA.

We further characterized IgG titers by analysis of IgG1 and IgG2a subtypes and found that both iVAX-derived *Ft*O-PS-conjugated MBP^{4xDQNAT} and PD^{4xDQNAT} boosted production of IgG1 antibodies by >2 orders of magnitude relative to all control groups as well as to glycosylated MBP^{4xDQNAT} derived from PGCT (**Figure 3.12c**). This analysis also revealed that both iVAX-derived bioconjugates elicited a strongly Th2-biased (IgG1 >> IgG2a) response, which is characteristic of most conjugate vaccines[165].

Taken together, these results provide evidence that the iVAX platform supplies vaccine candidates that are capable of eliciting strong, pathogen-specific humoral immune responses and recapitulate the Th2 bias that is characteristic of licensed conjugate vaccines.
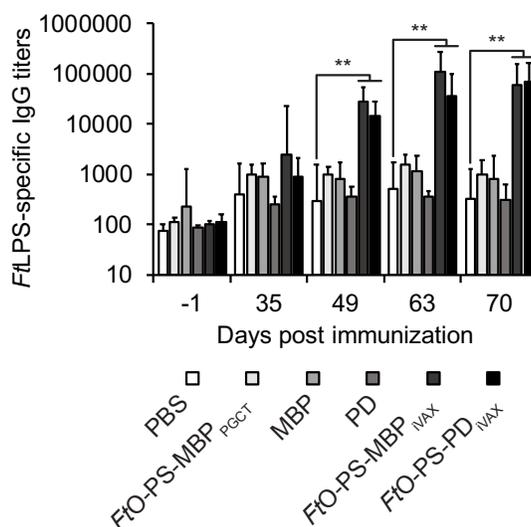
Figure 3.13 *Ft*LPS-specific antibody titers in vaccinated mice over time. Six groups of BALB/c mice were immunized subcutaneously with PBS or 7.5 µg of purified, cell-free synthesized aglycosylated MBP^{4xDQNAT}, *Ft*O-PS-conjugated MBP^{4xDQNAT}, aglycosylated PD^{4xDQNAT}, or *Ft*O-PS-conjugated PD^{4xDQNAT}. *Ft*O-PS-conjugated MBP^{4xDQNAT} prepared in living *E. coli* cells using PCGT was used as a positive control. Each group was composed of six mice except for the PBS control group, which was composed of five mice. Mice were boosted on days 21 and 42 with identical doses of antigen. *Ft*LPS-specific IgG titers were measured by ELISA in serum collected on day -1, 35, 49, 63, and 70 following initial immunization. iVAX-derived bioconjugates elicited significantly higher levels of *Ft*LPS-specific IgG compared to compared to the PBS control group in serum collected on day 35, 49, and 70 of the study (**$p < 0.01$, Tukey-Kramer HSD). Values represent means and error bars represent standard errors of *Ft*LPS-specific IgGs detected by ELISA.

3.5   Discussion

In this work we have established iVAX, a cell-free platform for portable, on-demand production of bioconjugate vaccines. We show that iVAX reactions can be detoxified to ensure the safety of bioconjugate vaccine products, freeze-dried for cold chain-independent distribution, and re-activated for high-yielding bioconjugate production by simply adding water. As a model vaccine candidate, we show that anti-*F. tularensis* bioconjugates derived from freeze-dried, endotoxin-edited iVAX reactions elicited

pathogen-specific IgG antibodies in mice as part of a Th2-biased immune response characteristic of licensed conjugate vaccines.

The iVAX platform has several important features. First, iVAX is modular, which we have demonstrated through the interchangeability of (i) carrier proteins, including those used in licensed conjugate vaccines, and (ii) bacterial O-PS antigens from *F. tularensis* subsp. *tularensis* (type A) Schu S4, ETEC *E. coli* O78, and UPEC *E. coli* O7. To our knowledge, this represents the first demonstration of oligosaccharyltransferase-mediated O-PS conjugation to authentic FDA-approved carrier proteins, likely due to historical challenges associated with the expression of licensed carriers in living *E. coli*[141-143]. Further expansion of the O-PS pathways used in iVAX should be possible given the commonly observed clustering of polysaccharide biosynthesis genes in the genomes of pathogenic bacteria[166]. This feature could make iVAX an attractive option for rapid, *de novo* development of bioconjugate vaccine candidates in response to a disease outbreak or against emerging drug-resistant bacteria.

Second, iVAX reactions are inexpensive, costing ~$12 mL^{-1} (**Table 3.4**) with the ability to synthesize ~20 µg bioconjugate mL^{-1} in one hour (**Figure 3.9a**). Assuming a dose size of 10 µg, our iVAX reactions can produce a vaccine dose for ~$6. For comparison, the CDC cost per dose for conjugate vaccines ranges from ~$9.50 for the *H. influenzae* vaccine ActHIB® to ~$75 and ~$118 for the meningococcal vaccine Menactra® and pneumococcal vaccine Prevnar 13®, respectively[167].

Third, and rather interestingly, we observed that iVAX-derived bioconjugates were significantly more effective at eliciting *Ft*LPS-specific IgGs than a bioconjugate derived from living *E. coli* cells using PGCT (**Figure 3.12**). One possible explanation for this increased effectiveness is the more extensive glycosylation that we observed for the *in vitro* expressed bioconjugates, with greater carbohydrate loading and decoration with higher molecular weight *Ft*O-PS species compared to their PGCT-derived counterparts. The reduced presence of high molecular weight O-PS species observed on bioconjugates produced using PGCT could be due to competition between the O-antigen polymerase Wzy and PglB *in vivo*. In contrast, our *in vitro* approach decouples O-PS synthesis, which occurs *in vivo* before lysate production, from glycosylation, which occurs *in vitro* as part of iVAX reactions. Our results are consistent with previous reports of PGCT-derived anti-*F. tularensis* bioconjugates which show that increasing the

ratio of conjugated *Ft*O-PS to protein in bioconjugates yields enhanced protection against *F. tularensis* Schu S4 in a rat inhalation model of tularemia[131]. In addition, a recent study showed that a conjugate vaccine made with a high molecular weight (80 kDa) *Ft*O-PS coupled to TT conferred better protection against intranasal challenge with *F. tularensis* live vaccine strain compared to a conjugate made with a low molecular weight (25 kDa) polysaccharide[152]. These results as well as our own point to the fact that a deeper understanding of important immunogen design features such as glycan loading and chain length could enable the production of more effective conjugate vaccines in the future.

Importantly, by enabling portable production of bioconjugate vaccines, iVAX addresses a key gap in both cell-free and decentralized biomanufacturing technologies. Production of glycosylated products has not yet been demonstrated in cell-based decentralized biomanufacturing platforms[10,11] and existing cell-free platforms using *E. coli* lysates lack the ability to synthesize glycoproteins[12,13,168,169]. While glycosylated human erythropoietin has been produced in a cell-free biomanufacturing platform based on freeze-dried Chinese hamster ovary cell lysates[14], this and the vast majority of other eukaryotic cell-free and cell-based systems rely on endogenous protein glycosylation machinery. As a result, these expression platforms offer little control over the glycan structures produced or the underlying glycosylation reactions, and significant optimization is often required to achieve acceptable glycosylation profiles[14]. In contrast, the iVAX platform is enabled by lysates derived from *E. coli* that lack endogenous protein glycosylation pathways, allowing for bottom-up engineering of desired glycosylation activity[134]. The ability to engineer desired glycosylation activity in iVAX uniquely enables the production of an important class of antibacterial vaccines at the point-of-care.

In summary, iVAX provides a new approach for rapid development and portable, on-demand biomanufacturing of bioconjugate vaccines. The iVAX platform alleviates cold chain requirements, which could enhance delivery of medicines to regions with limited infrastructure and minimize vaccine losses due to spoilage. In addition, the ability to rapidly produce vaccine candidates *in vitro* provides a unique means for rapidly responding to pathogen outbreaks and emergent threats. As a result, we believe that the iVAX platform, along with an emerging set of technologies with the ability to synthesize biomedicines

on-demand[10-14,168,169], has the potential to promote increased access to complex, life-saving drugs through decentralized production.

3.6   Additional information

3.6.1 Acknowledgements

3.6.2 Author contributions

J.C.S. designed research, performed research, analyzed data, and wrote the paper. T.J. and T.M. designed research, performed research, and analyzed data. R.S.D. and K.J.H. performed research. T.C.S. aided in research design. M.C.J. and M.P.D. directed research, analyzed data, and wrote the paper.

3.6.3 Publishing and patent information

Manuscript under review

Stark, J. C. †, Jaroentomeechai, T. †, Moeller, T., Dubner, R. S., Hsu, K. J., … DeLisa, M. P. *, & Jewett, M. C. * (2019) On-demand, cell-free biomanufacturing of conjugate vaccines at the point-of-care. *In review*. Preprint available on *bioRxiv*, doi: 10.1101/681841.

† contributed equally; * corresponding author

Patent applications

Jewett, M. C., Stark, J. C., DeLisa, M. P., & Jaroentomeechai, T. (2018) Method for rapid *in vitro* synthesis of bioconjugate vaccines with FDA-approved protein carriers in detoxified prokaryotic cell lysates. US Provisional Patent Application No. 62/644,811.

Jewett, M. C., Stark, J. C., DeLisa, M. P., & Jaroentomeechai, T. (2016) Method for rapid *in vitro* synthesis of bioconjugate vaccines via recombinant production of *N*-glycosylated proteins in prokaryotic cell lysates. US Patent Application Serial No. 15/650,127.

3.7   Tables

Table 3.1 Primers used to generate CLM24 Δ*lpxM*. Primers used to construct and verify the CLM24 Δ*lpxM* strain are listed below. KO primers were used for amplification of the kanamycin resistance cassette from pKD4 with homology to *lpxM*. Seq primers were used for colony PCRs and sequencing confirmation of knockout strains.

Primer Name	DNA Sequence (5' to 3')
lpxM KO for	TACACTATCACCAGATTGATTTTTGCCTTATCCGAAACTGGAAAAGCATGGTG TAGGCTGGAGCTGCTTC
lpxM KO rev	GCGAAGGCCTCTCCTCGCGAGAGGCTTTTTTATTTGATGGGATAAAGATCCAT ATGAATATCCTCCTTAGTTCCTATTC
lpxM seq for	AGTACCGGCTTTTTTTATTTGG
lpxM seq rev	CTAATACCACGCGTATTTTAACG

Table 3.2 Plasmids used in this study.

Plasmid	Description	Source
pSF-*Cj*PglB	*C. jejuni* PglB with a C-terminal 1xFLAG epitope tag in pSF, a modified pBAD expression vector	[123]
pGAB2	*F. tularensis* O-PS antigen gene cluster in pLAFR1	[34]
pMW07-O78	*E. coli* O78 antigen gene cluster in pMW07	[170]
pJHCV32	*E. coli* O7 antigen gene cluster in pVK102	[171]
pKD46	Encodes λ red system for recombineering	[95]
pKD4	Encodes kanamycin resistance cassette with upstream and downstream FRT sites	[95]
pCP20	Encodes *flp* for Flp-FRT recombination	[95]
pSF-*Cj*PglB-LpxE	*C. jejuni* PglB with a C-terminal 1xFLAG epitope tag and *F. tularensis* LpxE in pSF	This work; Addgene 128389
pJL1-sfGFP$^{217\text{-}DQNAT}$	Superfolder green fluorescent protein variant modified after residue T216 with 21 amino acid insertion containing the *C. jejuni* AcrA N123 glycosylation site but with an optimal DQNAT glycosylation sequence and a C-terminal 6xHis tag	[134]
pJL1-sfGFP$^{217\text{-}AQNAT}$	Same as pJL1 sfGFP$^{217\text{-}DQNAT}$, but with an AQNAT glycosylation sequence that is not modified by *Cj*PglB	[134]
pJL1-MBP^{4xDQNAT}	*E. coli* maltose binding protein with a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pJL1, a T7-driven *in vitro* expression vector	This work; Addgene 128390
pJL1-PD^{4xDQNAT}	*H. influenzae* protein D with a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pJL1	This work; Addgene 128391
pJL1-PorA^{4xDQNAT}	*N. meningitidis* PorA porin protein with a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pJL1	This work; Addgene 128392
pJL1-TTc^{4xDQNAT}	Fragment C domain of *C. tetani* toxin with a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pJL1	This work; Addgene 128393
pJL1-TTlight^{4xDQNAT}	Light chain variant of *C. tetani* toxin containing an inactivating E234A mutation in the enzyme active site with a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pJL1	This work; Addgene 128394
pJL1-CRM197^{4xDQNAT}	*C. diphtheriae* toxin variant with an inactivating G52E mutation in the enzyme active site with a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pJL1	This work; Addgene 128395
pJL1-TT^{4xDQNAT}	*C. tetani* toxin variant containing an inactivating E234A mutation in the enzyme active site with a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pJL1	This work; Addgene 128396
pJL1-EPA$^{DNNNS\text{-}DQNRT}$	*P. aeruginosa* exotoxin A containing a DNNNS glycosylation site at residue 242 and a DQNRT glycosylation site at residue 384 and a C-terminal 6xHis tag in pJL1	This work; Addgene 128397

Table 3.2 (continued). Plasmids used in this study.

Plasmid	Description	Source
pTrc99s-ssDsbA-MBP^{4xDQNAT}	*E. coli* maltose binding protein with an N-terminal DsbA signal sequence for periplasmic translocation and a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pTrc99s	This work; Addgene 128398
pTrc99s-ssDsbA-PD^{4xDQNAT}	*H. influenzae* protein D with an N-terminal DsbA signal sequence for periplasmic translocation and a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in Trc99s	This work; Addgene 128399
pTrc99s-ssDsbA-PorA^{4xDQNAT}	*N. meningitidis* PorA porin protein with an N-terminal DsbA signal sequence for periplasmic translocation and a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pTrc99s	This work; Addgene 128400
pTrc99s-ssDsbA-TTc^{4xDQNAT}	Fragment C domain of *C. tetani* toxin with an N-terminal DsbA signal sequence for periplasmic translocation and a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pTrc99s	This work; Addgene 128401
pTrc99s-ssDsbA-TTlight^{4xDQNAT}	Light chain variant of *C. tetani* toxin containing an inactivating E234A mutation in the enzyme active site with an N-terminal DsbA signal sequence for periplasmic translocation and a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pTrc99s	This work; Addgene 128402
pTrc99s-ssDsbA-CRM197^{4xDQNAT}	*C. diphtheriae* toxin variant with an inactivating G52E mutation in the enzyme active site with an N-terminal DsbA signal sequence for periplasmic translocation and a C-terminal 4xDQNAT glycosylation tag and a 6xHis tag in pTrc99s	This work; Addgene 128403
pTrc99s-ssDsbA-EPA$^{DNNNS-DQNRT}$	*P. aeruginosa* exotoxin A containing a DNNNS glycosylation site at residue 242 and a DQNRT glycosylation site at residue 384 with an N-terminal DsbA signal sequence for periplasmic translocation and a C-terminal 6xHis tag in pTrc99s	This work; Addgene 128404

Table 3.3 Antibodies and antisera used in this study.

Target	Source	Dilution
Rabbit pAb to 6xHis epitope tag	Abcam	1:7500
Mouse mAb FB11 to *F. tularensis* LPS	Abcam	1:5000
Rabbit pAb to *E. coli* O78 antigen	Abcam	1:2500
Rabbit pAb to *C. diphtheriae* toxin	Abcam	1:2000
Rabbit pAb to *C. tetani* toxin	Abcam	1:2000
Goat anti-rabbit IgG IR dye 680	LI-COR	1:15000-1:10000
Goat anti-rabbit IgG IR dye 800	LI-COR	1:15000-1:10000
Goat anti-mouse IgG IR dye 800	LI-COR	1:15000-1:10000
Goat anti-mouse IgG HRP	Abcam	1:25,000
Goat anti-mouse IgG1 HRP	Abcam	1:25,000
Goat anti-mouse IgG2a HRP	Abcam	1:25,000

Table 3.4 Cost analysis for iVAX reactions. The total cost to assemble iVAX reactions is calculated below. A 1 mL iVAX reaction produces two 10 μg vaccine doses and can be assembled for $11.75. In the table, amino acid cost accounts for 2 mM each of the 20 canonical amino acids purchased individually from Sigma. Lysate cost is based on a single employee making 50 mL lysate from a 10 L fermentation, assuming 30 lysate batches per year and a 5-year equipment lifetime. Component source is also included in the table if it is available to purchase directly from a supplier. Homemade components cannot be purchased directly and must be prepared according to procedures described in the Methods section.

Component	Cost ($/mL rxn)	Supplier	Product No
Mg(Glu)$_2$	<0.00	Sigma	49605
NH$_4$Glu	<0.00	MP	02180595
KGlu	<0.00	Sigma	G1501
ATP	0.01	Sigma	A2383
GTP	0.27	Sigma	G8877
UTP	0.23	Sigma	U6625
CTP	0.20	Sigma	C1506
Folinic acid	0.02	Sigma	47612
tRNA	0.21	Roche	10109541001
Amino acids	<0.00	homemade	
PEP	1.79	Roche	10108294001
NAD	0.07	Sigma	N8535-15VL
CoA	0.34	Sigma	C3144
Oxalic acid	<0.00	Sigma	P0963
Putrescine	<0.00	Sigma	P5780
Spermidine	<0.00	Sigma	S2626
HEPES	<0.00	Sigma	H3375
MnCl$_2$	<0.00	Sigma	63535
DDM	0.36	Anatrace	D310S
Plasmid	0.88	homemade	
Lysate	7.37	homemade	
Total	**11.75**	**$/mL rxn**	
	5.88	**$/dose**	

4. Energizing eukaryotic cell-free protein synthesis with glucose metabolism

4.1 Abstract

Eukaryotic cell-free protein synthesis (CFPS) is limited by the dependence on costly high-energy phosphate compounds and exogenous enzymes to power protein synthesis (*e.g.*, creatine phosphate and creatine kinase, CrP/CrK). Here, we report the ability to use glucose as a secondary energy substrate to regenerate ATP in a *Saccharomyces cerevisiae* crude extract CFPS platform. We observed synthesis of 3.64 ± 0.35 µg mL^{-1} active luciferase in batch reactions with 16 mM glucose and 25 mM phosphate, resulting in a 16% increase in relative protein yield (µg protein/$ reagents) compared to the CrP/CrK system. Our demonstration provides the foundation for development of cost-effective eukaryotic CFPS platforms.

4.2 Introduction

Cell-free protein synthesis (CFPS) is an emerging field that allows for the production of proteins without intact cells[39,40]. Crude cell lysates, or extracts, are employed instead. Supplying chemical energy (in the form of ATP) for the aminoacylation of tRNAs and peptide bond formation has been a grand challenge for CFPS development[39]. Historically, high-energy phosphate bond donors; such as phosphoenolpyruvate (PEP), creatine phosphate (CrP) (**Figure 4.1a**), and acetyl phosphate have been used[39,55,139,172-174]. In these cases, ATP regeneration requires the addition of pyruvate kinase, creatine kinase, or acetate kinase, respectively, or the endogenous presence of these enzymes in the cell extract. Unfortunately, rapid production of phosphate from these high-energy compounds has been shown to be inhibitory to CFPS (*e.g.*, *Escherichia coli*[175] and yeast[176]). Furthermore, batch reactions using these secondary energy substrates typically provide only a brief burst of ATP. In addition, phosphorylated energy compounds are costly, which limits industrial applications[46,47,177]. To address these limitations, new cost-effective secondary energy regeneration systems are sought.
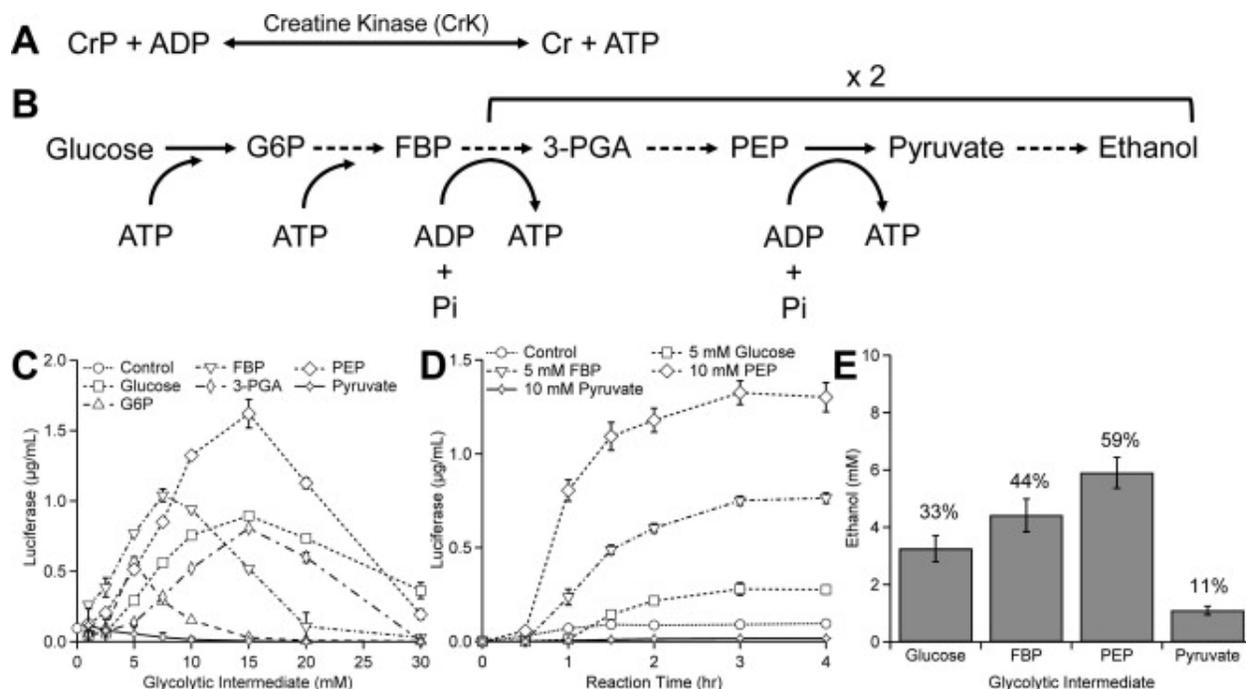
Figure 4.1 Glycolysis is active in yeast crude extract CFPS. (**a**) Schematic of creatine phosphate (CrP)/creatine kinase (CrK) energy regeneration system. (**b**) Proposed glycolytic energy regeneration system in yeast crude extracts. (**c**) To assess the possibility of using glycolytic intermediates to fuel CFPS, six glycolytic intermediates (FBP, PEP, glucose, 3-PGA, pyruvate, and G6P) were added as the sole secondary energy substrate to different yeast CFPS reactions in concentrations ranging from 0 mM to 30 mM and compared to a control containing no secondary energy substrate (circle). Of the non-phosphorylated secondary energy substrates assessed, glucose is the highest yielding for yeast CFPS. (**d**) Active luciferase is reported for time course reactions containing equivalent of 30 mM carbon for select glycolytic intermediates (*e.g.*, 5 mM glucose or 10 mM PEP) and (**e**) HPLC analysis of ethanol production after 4-h incubation for reactions performed in panel (**d**). The numbers above each column denote the percentage of theoretical conversion of each secondary energy substrate to ethanol. Values shown are means with error bars representing the standard deviation of at least three independent experiments.

Within the last decade, the *E. coli* CFPS platform has been able to activate natural metabolism within the lysate to fuel highly active CFPS from non-phosphorylated energy substrates by replacing PEP with glucose[47,111,177]. Mainly enabled by advances from Swartz and colleagues, glucose drives CFPS with a much lower cost and generates more ATP per secondary energy substrate molecule[47,111,177]. For example, glucose has a 2:1 molar ratio of secondary energy metabolite to ATP, compared to 1:1 ratio for both CrP and PEP (**Figure 4.1a, b**)[178]. As an extension of the pioneering works above, many groups have

turned to use of slowly metabolized glucose polymers to fuel *E. coli* based CFPS, including starch[179], maltodextrin[180,181], and maltose[45].

While *E. coli* based CFPS systems have been developed from non-phosphorylated energy substrates, making possible many new applications in industrial biotechnology and rapid prototyping[49,182-188], most eukaryotic CFPS platforms have been limited to the use of high-energy phosphate secondary energy substrates. This includes, for example, a yeast-based CFPS system we developed that leverages creatine phosphate and creatine phosphokinase (CrP/CrK) to power protein synthesis[172,176,189,190]. Here, we sought to assess the possibility to activate glycolysis in crude yeast cell extracts to regenerate cofactors and energy to provide the support system necessary to fuel highly active protein synthesis. The ability to use glucose to fuel CFPS is not only important for CFPS applications, but also can expand the impact of cell-free synthetic biology by joining a rapidly growing number of reports highlighting the ability to co-activate multiple biochemical systems in an integrated cell-free platform[45-47,101,111,180,191-194]. We demonstrate that it is indeed possible to power yeast CFPS reactions with glucose, as well as other glycolytic intermediates and non-phosphorylated energy sources, and have reached synthesis yields of 1.05 ± 0.12 µg mL^{-1} active luciferase with 16 mM glucose. After demonstrating synthesis of luciferase from glucose as the sole secondary energy substrate, we optimized our glucose energy system with the addition of cyclic AMP (cAMP) and exogenous phosphate, reaching batch yields of 3.64 ± 0.35 µg mL^{-1} active luciferase. To the best of our knowledge, our work is the first example of powering a eukaryotic CFPS reaction from the native glycolytic pathway. This opens the way to development of cost-effective eukaryotic CFPS platforms from multiple host organisms for a variety of applications.

4.3   Materials and Methods

Yeast extract preparation, CFPS reactions, and luciferase quantification were performed as previously described[172,176,189], with the exception that the energy regeneration system (CrP/CrK) was replaced with glycolytic intermediates. The concentration of magnesium glutamate added to CFPS reactions was optimized for each extract, as CFPS yields are known to be sensitive to magnesium[172]. We tested glucose, glucose-6-phosphate (G6P), 3-phosphoglyceric acid (3-PGA), phosphoenolpyruvate

(PEP), fructose-1,6-bisphosphate (FBP), and pyruvate in concentrations ranging from 0 to 30 mM. We also tested CFPS reactions containing 0–25 mM glucose in combination with the CrP/CrK energy regeneration system. When denoted, 0.15 mM cAMP and phosphate (in the form of potassium phosphate, pH 7.4) were included in the reaction mixture. Reaction conditions can be found in **Table 4.1**. HPLC analysis of ethanol was performed as previously described[189]. Nucleotide analysis was performed as previously described[176] except the gradient for buffer B was adjusted to: 0 min, 0%; 10 min, 30%; 50 min, 80%; 55 min, 100%; 60 min, end.

4.4   Results

We sought to fuel yeast CFPS by activating glycolysis and central metabolism with non-phosphorylated energy substrates. We expect this metabolism to be active given the fact that Eduard Büchner discovered in 1897 that yeast extract could convert sugar to ethanol and carbon dioxide[195]. Initially, we screened for the ability of six different glycolytic intermediates to fuel combined transcription and translation in 15 μL batch CFPS reactions for 4 h at 21 °C (**Figure 4.1c**). The six intermediates included fructose 1,6-bisphosphate (FBP), phosphoenolpyruvate (PEP), glucose, 3-phosphglyceric acid (3-PGA), pyruvate, and glucose 6-phosphate (G6P) at concentrations ranging from 0 to 30 mM. The CFPS reaction was programmed to synthesize luciferase as a model reporter protein and combined transcription and translation was enabled by the use of the Ω cap-independent translation initiation leader sequence[190]. Strikingly, our results demonstrated that it is indeed possible to activate yeast CFPS reactions from glycolytic intermediates upstream of pyruvate, reaching 1.04 ± 0.45 and 1.62 ± 0.10 μg mL^{-1} when powering the reaction with FBP and PEP, respectively. Of the six glycolytic intermediates, only pyruvate was unable to function as a secondary energy source (**Figure 4.1c**). The inability of pyruvate to power CFPS was expected due to the lack of ATP regenerating power of pyruvate alone in fermentation metabolic processes.

In order to more carefully understand the system dynamics, we subsequently performed time course CFPS reactions with the three highest-yielding intermediates (FBP, glucose, and PEP). This revealed that the choice of glycolytic intermediate impacted the rate of protein synthesis but not the

reaction duration; in all cases protein synthesis had terminated after 4 h (**Figure 4.1d**). Negative control reactions performed with pyruvate or no secondary energy substrate produced little to no luciferase (**Figure 4.1d**). The carbon from the glycolytic intermediates is expected to produce ethanol through fermentation, as shown in previous works[195,196]. Thus, we measured ethanol production to confirm glycolysis was active for each carbon source. As expected, we found that ethanol is synthesized when glucose, FBP, and PEP are used to power protein synthesis (**Figure 4.1e**). Ethanol is also produced in the presence of pyruvate, but no protein is synthesized due to limited ATP availability as described above (**Figure 4.1e**).

With the goal of increasing protein synthesis yields, we next tested a dual system, in which glucose is used in combination with CrP/CrK. Previously, such a system was demonstrated by Kim et al. to enhance yields in an *E. coli* CFPS platform[178]. Unexpectedly, we found that the addition of glucose to the CrP/CrK system severely inhibits CFPS, with 10 mM glucose addition resulting in an 89% reduction in protein synthesis (**Figure 4.2a**). We reasoned that this could result from a decrease in pH, as seen previously in *E. coli* CFPS platforms powered by glucose, or a toxicity effect from ethanol accumulation[47]. However, we observed no change in pH during the course of the reaction (**Figure 4.2b**), and showed that ethanol is not toxic in our reactions at concentrations of up to 25 mM (**Figure 4.2c**), which far exceeded the expected ethanol produced (**Figure 4.1e**). Historically, non-productive energy consumption has been identified as one of the primary reasons for early termination of CFPS. Thus, we used quantitative HPLC analysis to track the ATP pool over time. Nucleotide analysis revealed that the decrease in protein synthesis yields when glucose is added to the reaction is due to rapid ATP consumption. For example, in the presence of 25 mM glucose, ATP is fully consumed within the first 15 min of reaction (**Figure 4.2d**), constraining the ability to produce protein.
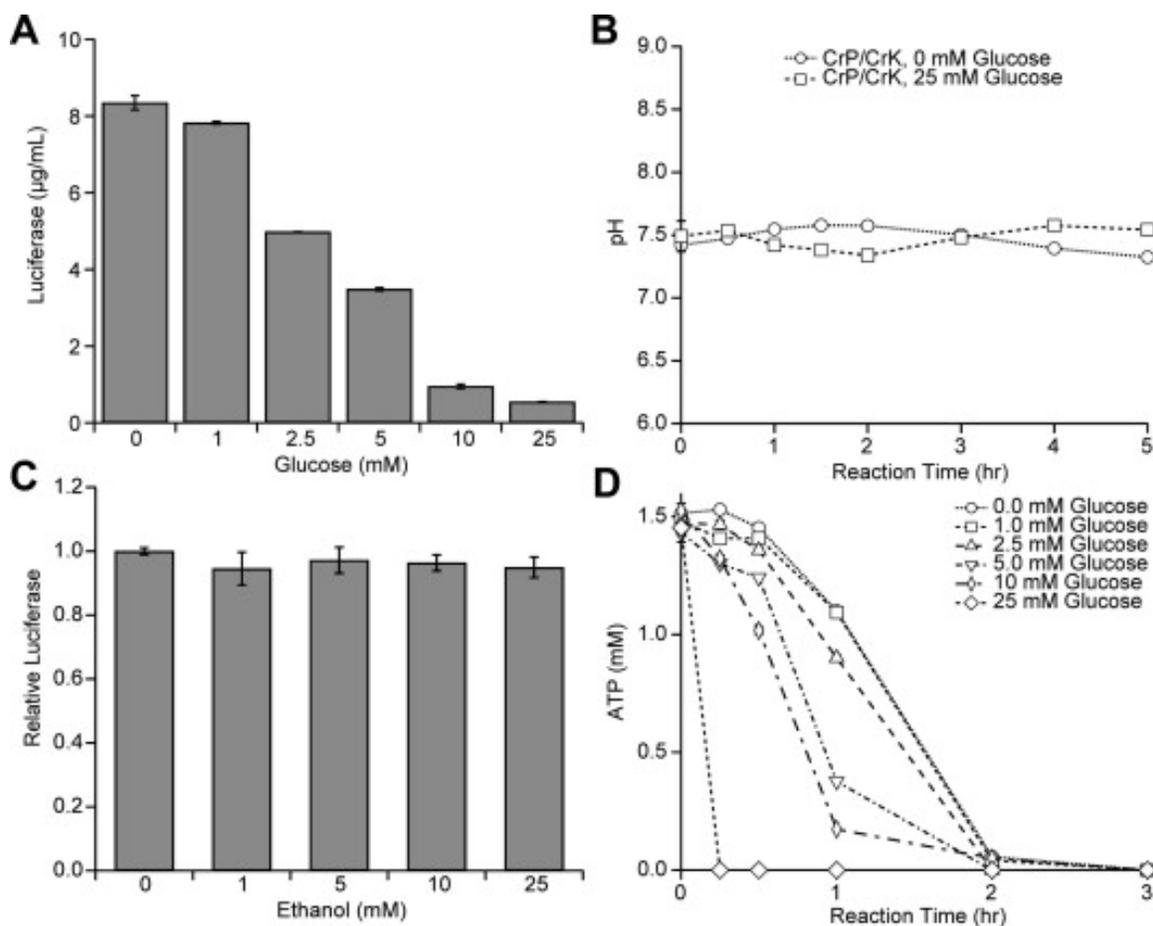
Figure 4.2 Yeast CFPS CrP/CrK plus glucose dual system for energy regeneration does not improve CFPS yields. (**a**) 0 to 25 mM glucose was added to CFPS reactions containing 25 mM creatine phosphate (CrP) and 0.27 mg/mL creatine kinase (CrK). Increasing the starting glucose concentration decreases luciferase yields. (**b**) The pH of CFPS reactions containing 25 mM CrP, 0.27 mg/mL CrK, and either 0 mM or 25 mM glucose was measured at regular intervals. Reaction pH remains approximately constant over 5 h. (**c**) To assess possible ethanol inhibition, various concentrations of ethanol, ranging from 0 mM to 25 mM, were added to CFPS reactions. Active luciferase yields are reported relative to the 0 mM ethanol condition, showing that inhibition was not observed. (**d**) The concentration of ATP was measured at intervals during CFPS reactions including 25 mM CrP, 0.27 mg/mL CrK, and 0–25 mM glucose. ATP is rapidly depleted as the starting glucose concentration is increased. Values shown in (**a**)–(**c**) are means with error bars representing the standard deviation of at least three independent experiments. Data from panel (**d**) traces are individual measurements.

Given the inability to activate a dual energy regeneration system, we returned to the glucose-only system, and determined through an initial optimization that 16 mM glucose is the optimal substrate concentration (**Figure 4.3a**). We subsequently carried out a series of additional optimization experiments to try to increase CFPS. We explored the effects of reaction temperature, magnesium glutamate

concentration, potassium glutamate concentration, spermidine concentration and additives such as cyclic AMP (cAMP) (**Table 4.2**). Despite a rigorous search, we only observed that addition of cAMP increased yields, suggesting that our original conditions for yeast CFPS captured a maximum. The addition of 0.15 mM cAMP increased our yields 1.5-fold, bringing our yields to approximately 1 μg mL^{-1}. The kinetics of protein synthesis follows an interesting trajectory when using glucose and cAMP. Specifically, protein synthesis is delayed when using glucose as the energy source (**Figure 4.3b**), which we attribute to ATP availability. ATP is rapidly consumed in the first 30 min of the reaction, but more than 50% is regenerated after 90 min (**Figure 4.3c**).
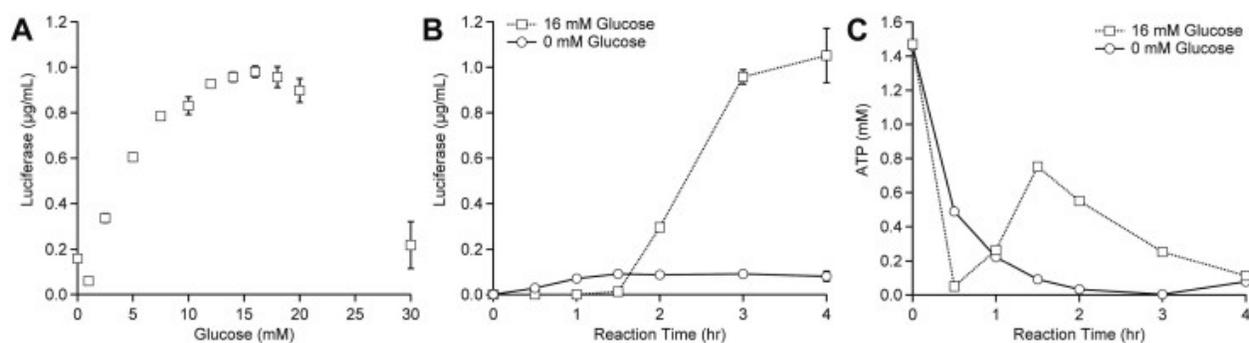
Figure 4.3 Optimizing yeast CFPS reaction conditions with glucose as a secondary energy substrate. (**a**) The optimal starting concentration of glucose was determined via addition of 0–30 mM of glucose to CFPS reactions containing 0.15 mM cAMP. The optimum was observed at 16 mM glucose. (**b**) Luciferase and (**c**) ATP concentrations were measured at regular intervals over time in CFPS reactions containing 16 mM glucose or 0 mM glucose. Values shown are means with error bars representing the standard deviation of at least three independent experiments.

With the ability to fuel CFPS by glycolysis at hand, we next investigated the use of slowly metabolized carbon polymers to slow the initial consumption of ATP. We demonstrated that soluble starch can fuel CFPS, though at much lower yields than the glucose system, reaching only ~0.3 μg mL^{-1} with 1.4% (w/v) starch (**Figure 4.4a, b**). Using starch did not reduce initial consumption of ATP, with only 0.2 mM left after 30 min of the reaction (**Figure 4.4c**). Our data suggest that ATP regeneration limits the use of starch when compared to glucose alone. Specifically, the regeneration of ATP when using starch is lower than with 16 mM glucose, leading to a lower protein yield. Supplying α-

glucosidase and amyloglucosidase enzymes did not improve protein synthesis yields, suggesting the activity of our crude lysates is sufficient to metabolize starch (**Figure 4.4d**).
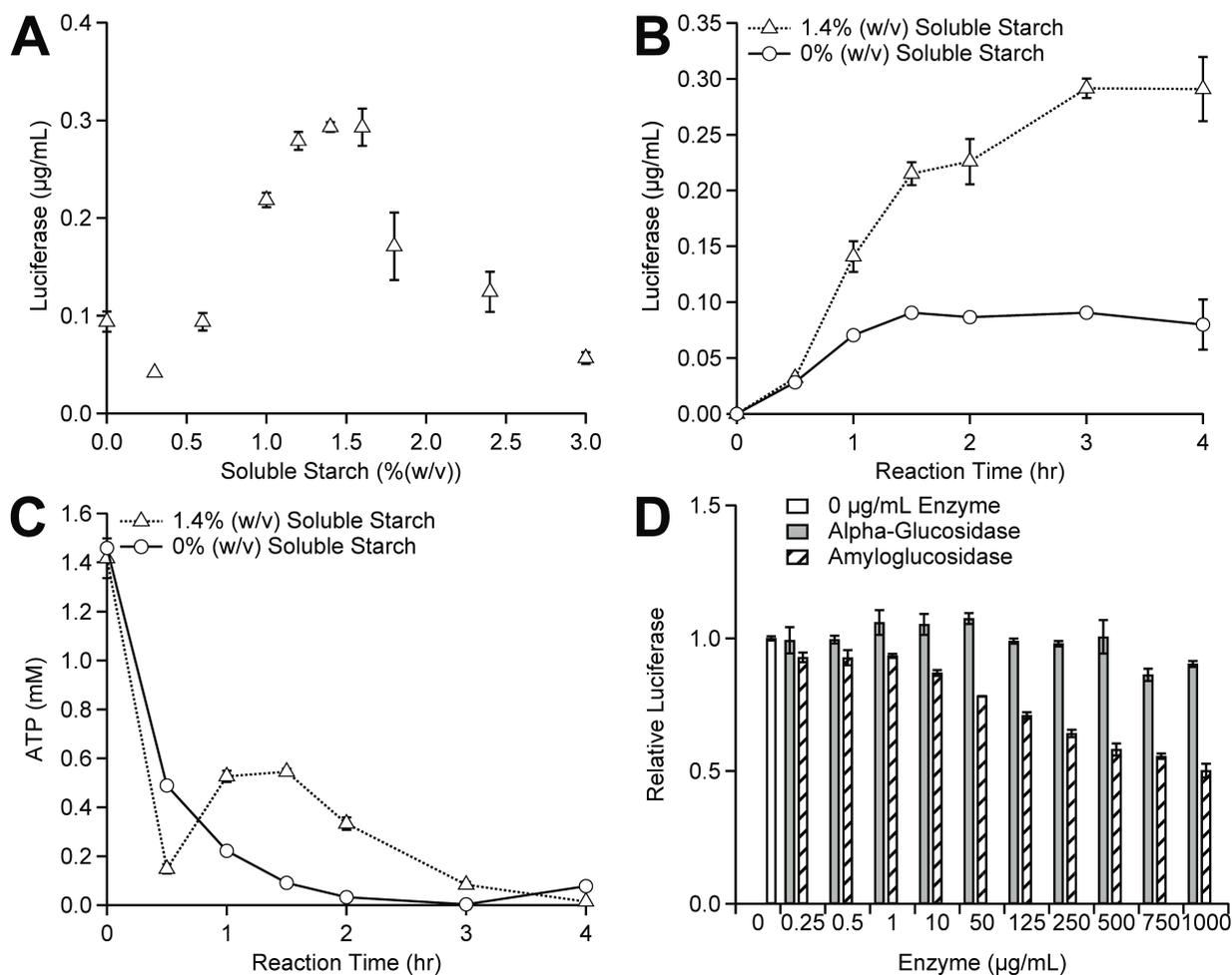
Figure 4.4 Optimizing yeast CFPS reactions with starch. (**a**) Soluble starch was added to the CFPS reaction in concentrations ranging from 0% to 3% weight starch/volume reaction (w/v). The optimal concentration of starch in the CFPS reactions was 1.4% (w/v). Concentrations of (**b**) luciferase and (**c**) ATP were measured at regular intervals during CFPS reactions with 1.4% (w/v) starch or 0% (w/v) soluble starch. (**d**) Varying concentrations of alpha-glucosidase, amyloglucosidase, or no exogenous enzymes were added to CFPS reactions containing 1.4% (w/v) starch. Luciferase yields are reported relative to the 0 µg/mL enzyme condition. Values shown are means with error bars representing the standard deviation of at least three independent experiments.

Although we demonstrated proof of principle with starch as an energy substrate, yields remained higher with the glucose energy regeneration system. Therefore, we returned to the glucose system to

search for parameters that could increase the level of luciferase synthesized. Previously, Calhoun and

Swartz showed that the use of non-phosphorylated energy substrates can result in phosphate limitation

during energy regeneration. They observed that the addition of 10 mM inorganic phosphate provided a 3-

fold increase in CFPS yields compared to their glucose-driven *E. coli* CFPS system alone[47]. Building off of

this advance, we evaluated the addition of 0–50 mM inorganic phosphate in the form of potassium

phosphate to our glucose-driven yeast CFPS system (**Figure 4.5a**). With the addition of 25 mM inorganic

phosphate, CFPS yields increased almost 3.5-fold, reaching 3.64 ± 0.35 µg mL^{-1} (**Figure 4.5a**). **Figure**

4.5b shows luciferase accumulation over time.
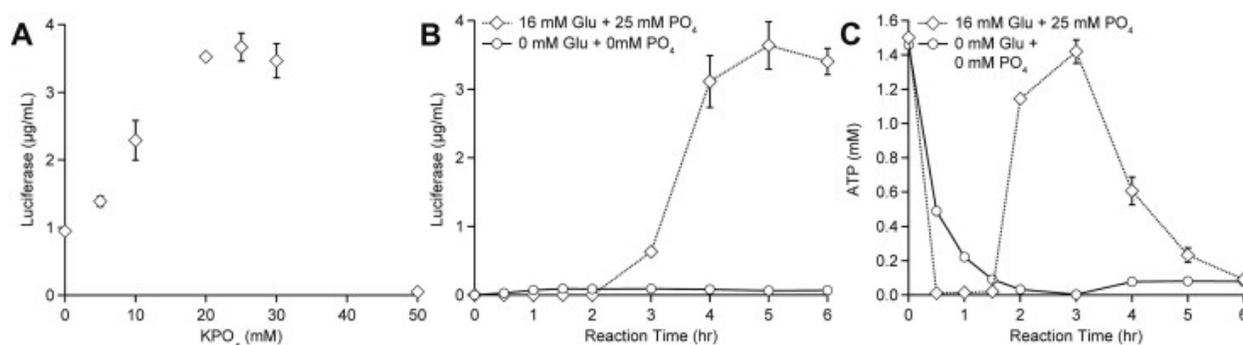
Figure 4.5 CFPS reactions with glucose are phosphate-limited: increasing phosphate concentration increases protein yields and prolongs the CFPS reaction. (**a**) The optimal amount of exogenous phosphate was determined via addition of 0–50 mM of phosphate to CFPS reactions containing 16 mM glucose. The optimum was observed at 25 mM phosphate. (**b**) Luciferase and (**c**) ATP concentration were measured at regular intervals in CFPS reactions containing 16 mM glucose and 25 mM phosphate or 0 mM glucose and 0 mM phosphate. Values shown are means with error bars representing the standard deviation of at least three independent experiments.

As reported for the glucose and starch systems, protein production appears to be linked to ATP

availability, which can be described by Atkinson's adenylate energy charge (E.C.) calculation[197] (**Figure**

4.6a). *In vivo* studies have shown energy is limiting in systems with an E.C. less than 0.8[198]. In reactions

containing glucose and phosphate, we observed that ATP is rapidly consumed within the first 30 min of

the reaction, but now almost 100% is regenerated after 3 h (**Figure 4.5c**), enabling protein synthesis to

extend to 5 h (**Figure 4.5b**). The observed ATP regeneration coincides exactly with initiation of protein

synthesis and the point at which E.C. rises above 0.8, between 2 and 3 h of the reaction (**Figure 4.6b**).

Based on the adenylate energy charge calculations, we propose that this trend in ATP concentration is observed due to the activation of glucose metabolism. At the start of the reaction, ATP is consumed in the pay-in phase of glycolysis while glucose is metabolized. After all available glucose has been consumed, ATP is regenerated by glucose metabolism and accumulates until sufficient ATP is available for protein synthesis.

A

$$\text{Energy Charge (E.C.)} = \frac{[\text{ATP}] + \tfrac{1}{2}[\text{ADP}]}{[\text{ATP}] + [\text{ADP}] + [\text{AMP}]}$$

B
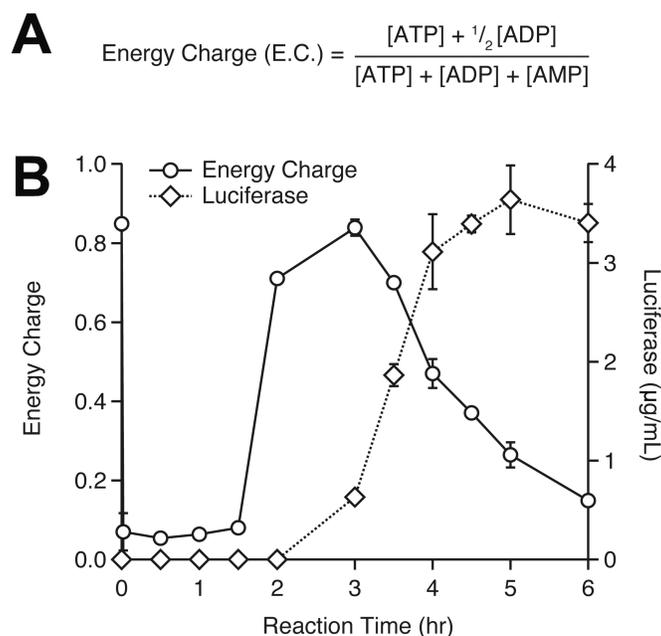
Figure 4.6 Glucose metabolism regenerates energy to fuel protein synthesis. (**a**) The definition of the adenylate energy charge (E.C.) as described by Atkinson[197]. In vivo studies have shown that energy is limiting when E.C.<0.8[198]. (**b**) Energy charge and luciferase concentration are plotted as a function of reaction time for CFPS reactions containing 16 mM glucose and 25 mM phosphate. The energy charge is >0.8 when protein synthesis begins, between t = 2 − 3 hours. Values shown are means with error bars representing the standard deviation of at least three independent experiments.

As compared to the glucose only system, ATP regeneration is improved in the glucose/phosphate system, resulting in prolonged availability of a high concentration of ATP, which manifests in higher protein synthesis yields. This is the longest reported batch yeast CFPS reaction to date, to the best of our knowledge.

4.5 Discussion

In summary, we have developed a new energy regeneration system for yeast CFPS that uses glucose and phosphate. This novel approach removes the need for an expensive phosphorylated secondary energy source and avoids inhibitory phosphate accumulation. To our knowledge, this is the first time that a eukaryotic-based CFPS system has been powered by natural energy metabolism of a non-phosphorylated energy substrate. Although our yields do not exceed those previously reported with yeast extract and the CrP/CrK system[189], we have increased the relative protein yield (μg protein/$ reagents) by 16% with our novel glucose/phosphate system (**Figure 4.7**). Further optimization of this platform through host strain engineering, as has been done in *E. coli*-based systems[47,199], holds promise to result in a cost-effective eukaryotic CFPS platform for high throughput protein expression, synthetic biology, and proteomic and structural genomic applications. We anticipate that yeast CFPS will become a major player alongside other CFPS technologies in years to come.
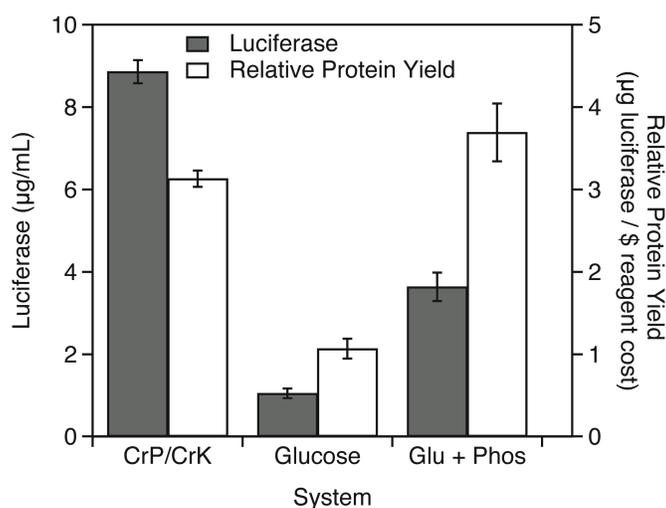
Figure 4.7 Glucose and phosphate system achieves improved relative protein yields compared to the state-of-the-art CrP/CrK system. Here we compare the traditional CrP/CrK system to the novel glucose and glucose/phosphate system reported here as measured by active protein synthesis yield (μg/mL; left axis) and relative protein yield (μg protein synthesized per $ reagent cost; right axis). Substrate cost includes all substrates used to treat the crude extract, make the genetic template, and assemble the CFPS reaction. Values shown are means with error bars representing the standard deviation of at least three independent experiments.

4.6   Additional information

4.6.1 Acknowledgements

4.6.2 Author contributions

M.J.A., C.E.H., and M.C.J. conceived and supervised the study; M.J.A., J.C.S., C.E.H., and M.C.J. designed experiments; M.J.A. and J.C.S. performed experiments; M.J.A. and J.C.S. analyzed data; M.J.A., J.C.S., C.E.H. and M.C.J. wrote and revised the manuscript.

4.6.3 Publishing and patent information

Published manuscript

Anderson, M. J. †, Stark, J. C. †, Hodgman, C. E., & Jewett, M. C. * (2015) Energizing eukaryotic cell-free protein synthesis with glucose metabolism. *FEBS Letters*, doi: 10.1016/j.febslet.2015.05.045.

† contributed equally; * corresponding author

Patent application

Jewett, M. C., Anderson, M. J., Stark, J. C., & Hodgman, C. E. (2014) Methods for activating natural energy metabolism for improved yeast cell-free protein synthesis. PCT International Patent Application Serial No. PCT/IB2015/059960.

4.7   Tables

Table 4.1 Final concentrations of components used for CrP/CrK- and glucose-powered CFPS systems. These values do not include the concentrations of small molecules in the yeast extract. Notably, optimal magnesium glutamate concentrations depend heavily on the amount of magnesium in the extract. Each extract is tested individually to determine optimal [Mg(Glu)$_2$].

Reagent	CrP/CrK System		Glucose System	
Salts and polyamines:				
Magnesium glutamate (Mg(Glu)$_2$)	4 – 6	mM	4 – 6	mM
Potassium glutamate (KGlu)	120	mM	120	mM
Spermidine	0.50	mM	0.50	mM
Putrescine	2	mM	2	mM
NTPs (ATP, GTP, UTP, CTP) [individual concentration]	1.50	mM	1.50	mM
20 amino acids [individual concentration]	80	µM	80	µM
DTT	1.20	mM	4	mM
Creatine phosphate	25	mM	0	mM
Creatine kinase	0.27	mg/mL	0	mg/mL
Transcriptional and translational components:				
Yeast extract	2.80	mg/mL	2.80	mg/mL
Reporter PCR template: ΩLucA$_{50}$	6.67	µg/mL	6.67	µg/mL
T7 RNA polymerase	0.027	mg/mL	0.027	mg/mL
Other components:				
HEPES-KOH, pH 7.6 (total in reaction)	22	mM	22	mM
Glucose	0	mM	16	mM
Phosphate (potassium phosphate)	0	mM	25	mM
Glycerol	11	%	11	%
Cyclic AMP (cAMP)	0	mM	0.15	mM

Table 4.2 Parameters optimized during development of CFPS platforms powered by glucose metabolism. These values do not include the concentrations of small molecules in the yeast extract. The optimal values for each parameter (right column) were used in all subsequent CFPS reactions (**Table 4.1**).

Reagent	Conditions Tested		Optimal Condition	
Magnesium glutamate (Mg(Glu)$_2$)	4 – 6	mM	4 – 6	mM
Potassium glutamate (KGlu)	80 – 160	mM	120	mM
Spermidine (Spe)	0 – 2	mM	0.50	mM
Cyclic AMP (cAMP)	0 – 0.40	mM	0.15	mM
Reaction temperature	21 – 30	°C	21	°C

5.   BioBits^{TM}: Hands-on kits for teaching molecular and synthetic biology

5.1   Abstract

Synthetic biology offers opportunities for experiential educational activities at the intersection of the life sciences, engineering, and design. However, implementation of hands-on biology activities in classrooms is challenging because of the need for specialized equipment and expertise to grow living cells. To address this limitation, we developed three synthetic biology education kits enabled by freeze-dried cell-free reactions, which can be activated by just adding water. The BioBits^{TM} Bright kit includes activities and supporting curricula for teaching the central dogma, tunable protein expression, and design-build-test cycles through *in vitro* expression of fluorescent proteins. The BioBits^{TM} Explorer kit includes activities designed to teach enzymatic catalysis, biomaterial formation, and biosensors with protein or small molecule outputs that students can see, smell, and touch. The BioBits^{TM} Health kit contains lab activities and supporting curricula for teaching antibiotic resistance mechanisms and CRISPR-Cas9 gene editing using reactions with visual, fluorescent readouts. We report data generated by K-12 students and teachers using the BioBits^{TM} Bright and Health kits, demonstrating the utility of these resources for use by untrained operators without sophisticated laboratory equipment. Finally, we developed inexpensive custom incubators and imagers, enabling the production of comprehensive kits costing between $100-$200 per 30-person classroom. Together, the BioBits^{TM} kits represent user-friendly resources that promise to enhance biology education both inside and outside the classroom.

Note: Supporting curricula for the BioBits^{TM} Bright and Health kits can be found in the Appendix and additional supporting materials for the BioBits^{TM} Bright kit can be found online:

http://advances.sciencemag.org/content/suppl/2018/07/30/4.8.eaat5107.DC1

5.2   BioBits^{TM} Bright: A fluorescent synthetic biology education kit

5.2.1 Introduction

Synthetic biology aims to program biological systems to carry out useful functions. As a field, synthetic biology has made meaningful progress toward biomanufacturing of medicines[200,201], sustainable chemicals[202,203], and advanced fuels[204], as well as cellular diagnostics and therapies[205-208]. At the core of these advances is the ability to control and tune the processes of transcription and translation, offering a point of entry for teaching fundamental biology topics through cutting-edge biological technologies. Synthetic biology also offers rich educational opportunities, as it requires students to confront real-world, interdisciplinary problems at the intersection of diverse disciplines including chemistry, biology, engineering, computer science, design, policy, and ethics. Such cross-cutting educational activities align closely with the objectives of K-12 STEAM (science, technology, engineering, the arts, and mathematics) education and priorities identified by the National Academy of Engineering to enable students to apply, adapt, and connect fundamental principles across multiple disciplines[209].

Synthetic biology–based educational efforts such as the BioBuilder Educational Foundation[210-213] and the International Genetically Engineered Machines competition[214,215] have made great strides toward incorporating synthetic biology into high school and university education. These programs have resulted in student-reported academic gains, high student engagement, and increased self-identification as biological engineers[216-218]. However, efforts to incorporate a hands-on molecular or synthetic biology curriculum have been limited by (i) the number of robust systems that can be converted into teaching materials; (ii) the need for expensive, specialized equipment to store, grow, and transport cells; and (iii) biosafety considerations that limit the ability to work with cells outside of a laboratory setting[219]. Addressing these limitations would help expand educational opportunities for students in classrooms, as well as inform and promote public engagement in synthetic biology.

Freeze-dried, cell-free (FD-CF) systems represent an emerging technology with exciting potential as a chassis for educational tools. FD-CF systems harness an ensemble of catalytic components [for example, RNA polymerases, ribosomes, aminoacyl–transfer RNA (tRNA) synthetases, translation

initiation, and elongation factors, etc.] from cell lysates to synthesize proteins in vitro[39]. Hence, FD-CF reactions do not use intact organisms; thus, they circumvent many of the biosafety and biocontainment regulations that exist for living cells. Further, FD-CF systems are stable at room temperature for more than 1 year[50] and can be run simply by adding water and DNA template to a freeze-dried pellet of reagents, eliminating the need for specialized equipment or expertise to run reactions. Finally, FD-CF systems are robust, with demonstrated utility for point-of-use biosynthesis of sophisticated diagnostics, protein therapeutics, vaccines, small molecules, and molecular biology reagents[12,13,50,51,89,134,220,221]. If FD-CF technology could be used to develop safe, portable, and easy-to-use educational tools, it would significantly lower the barrier to entry for teaching synthetic biology.

Here, we describe BioBits™ Bright, a portable, just-add-water educational kit and accompanying hands-on laboratory modules designed for use outside of the laboratory by untrained operators (**Figure 5.1**). To facilitate kit construction, we developed a library of fluorescent proteins that express at high yields (≥600 µg ml^{-1}) in FD-CF reactions. We report data for each module from workshops with Chicago K-12 students and teachers to demonstrate robustness and ease of use. Laboratory modules are designed to (i) synergize with fundamental biology education, as evidenced by the supporting curriculum developed by Chicago middle and high school teachers (**Curricula 1-5**); (ii) be run independently or in sequence; and (iii) be adapted for use with students at various educational levels. Notably, to make BioBits™ Bright laboratory activities accessible to resource-limited classrooms, we have also developed low-cost incubators and imagers. Separately, we describe BioBits™ Explorer[66], a next-generation BioBits™ kit developed to illustrate an even wider range of biological concepts (for example, enzymatic catalysis and genetic circuits). We anticipate that the availability of our BioBits™ kits and the data reported here will encourage teaching and broaden participation in the field of synthetic biology.
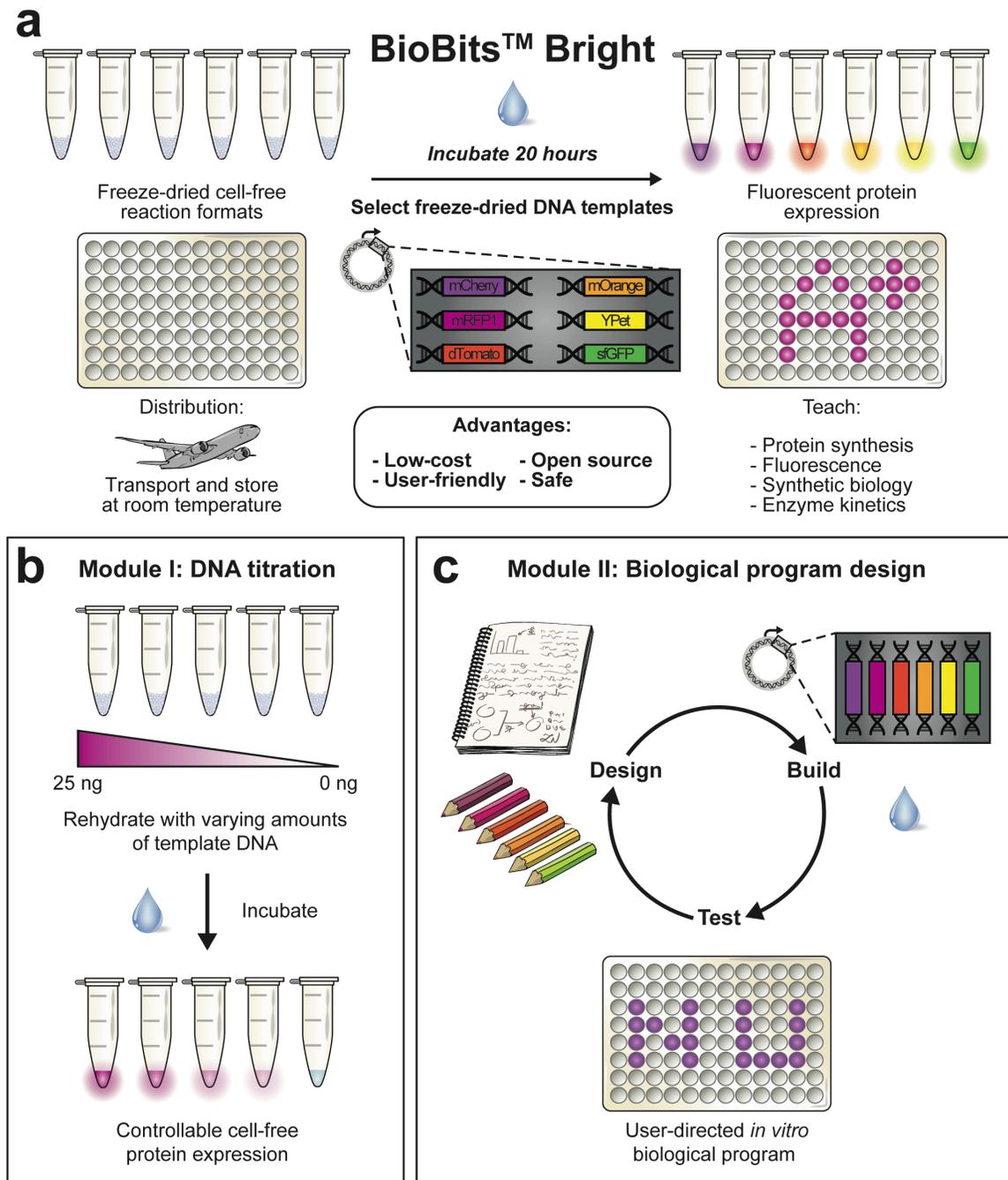
Figure 5.1 BioBits™ Bright: A portable, cell-free synthesized fluorescent protein library for teaching the central dogma of molecular biology and synthetic biology. (**a**) We describe here the development of an educational kit containing two laboratory modules using FD-CF reactions and a library of in vitro–synthesized fluorescent proteins. (**b**) In module I, students investigate how biological systems can be engineered by adding varying amounts of DNA template to FD-CF reactions. Titrating the amount of DNA template results in varying levels of fluorescent protein production, which are visible to the naked eye and under a blue or black light. (**c**) In module II, users design their own in vitro program using DNA encoding the fluorescent protein library and any of the DNA template concentrations investigated in module I. This module offers the opportunity to go through a user-directed design-build-test (DBT) cycle. All reagents used in these activities (freeze-dried reactions and plasmids) can be stored and transported without refrigeration, making them highly portable for use outside of the laboratory.

5.2.2 Methods

5.2.2.1 Bacterial strains and plasmids

Escherichia coli NEB 5-alpha (New England BioLabs) was used in plasmid cloning transformations and for plasmid preparation. *E. coli* BL21 Star (DE3) (Thermo Fisher Scientific) was used for preparation of CFPS extracts. Gibson assembly was used for seamless construction of plasmids used in this study (**Table 5.1**). For cloning, the pJL1 vector (Addgene, 69496) was digested using restriction enzymes *Nde*I and *Sal*I–HF (NEB). Each gene was amplified via polymerase chain reaction (PCR) using Phusion High-Fidelity DNA polymerase (NEB) with forward and reverse primers designed with the NEBuilder Assembly Tool (nebuilder.neb.com) and purchased from IDT (Integrated DNA Technologies). PCR products were gel-extracted using the EZNA Gel Extraction Kit (Omega Bio-Tek), mixed with Gibson assembly reagents, and incubated at 50°C for 1 hour. Plasmid DNA from the Gibson assembly reactions was transformed into *E. coli* NEB 5-alpha cells, and circularized constructs were selected on LB agar supplemented with kanamycin (50 µg ml^{-1}; Sigma-Aldrich). Sequence-verified clones were purified using the EZNA Plasmid Midi Kit (Omega Bio-Tek) for use in FD-CF reactions.

5.2.2.2 CFPS extract preparation

CFPS extract was prepared by sonication, as previously reported[102]. Briefly, *E. coli* BL21 Star (DE3) (Thermo Fisher Scientific) was grown in 2× YTPG media at 37°C. T7 polymerase expression was induced at an OD$_{600}$ (optical density at 600 nm) of 0.6 to 0.8 with 1 mM isopropyl-β-D-1-thiogalactopyranoside. Cells were grown at 30°C to a final OD$_{600}$ of 3.0, at which point cells were pelleted by centrifugation at 5000g for 15 min at 4°C. Cell pellets were then washed three times with cold S30 buffer [10 mM tris-acetate (pH 8.2), 14 mM magnesium acetate, and 60 mM potassium acetate] and pelleted at 5000g for 10 min at 4°C. After the final wash, cells were pelleted at 7000g for 10 min at 4°C, weighed, flash-frozen in liquid nitrogen, and stored at −80°C. For lysis, cell pellets were suspended in 1 ml of S30 buffer per 1 g of wet cell mass, and cells were transferred into 1.5-ml microcentrifuge tubes and placed in an ice-water bath to minimize heat damage during sonication. The cells were lysed using a Q125 Sonicator (Qsonica) with a 3.175-mm-diameter probe at 20 kHz and 50% amplitude. The input energy was monitored, with 640 J used to lyse 1 ml of suspended cells. The lysate was then centrifuged

once at 12,000g at 4°C for 10 min. Cell extract was aliquoted, flash-frozen on liquid nitrogen, and stored at −80°C. Alternatively, for classroom settings where it is not practical to generate or obtain FD-CF reactions, similar cell-free systems are available commercially from companies such as Promega (L1130).

5.2.2.3   Cell-free protein synthesis

FD-CF reactions were carried out in PCR tubes or plates (5 µl reactions). The CFPS reaction mixture consisted of the following components: 1.2 mM adenosine 5′-triphosphate; 0.85 mM each of guanosine 5′-triphosphate, uridine 5′-triphosphate, and cytidine 5′-triphosphate; L-5-formyl-5,6,7,8-tetrahydrofolic acid (34.0 µg ml^{-1}; folinic acid); *E. coli* tRNA mixture (170.0 µg ml^{-1}); 130 mM potassium glutamate; 10 mM ammonium glutamate; 8 mM magnesium glutamate; 2 mM each of 20 amino acids; 0.4 mM nicotinamide adenine dinucleotide; 0.27 mM coenzyme A; 1.5 mM spermidine; 1 mM putrescine; 4 mM sodium oxalate; 33 mM phosphoenolpyruvate; 57 mM HEPES; plasmid (13.3 µg ml^{-1}; unless otherwise noted); and 27% (v/v) of cell extract[101]. For quantification of fluorescent protein yields via radioactive leucine incorporation, 10 µM L-^{14}C-leucine (11.1 gigabecquerel mmol^{-1}, PerkinElmer) was added to the CFPS mixture.

5.2.2.4   Lyophilization of cell-free reactions

FD-CF reactions were prepared according to the recipe above, but without plasmid added. CFPS reactions and plasmids were separately lyophilized using a VirTis BenchTop Pro lyophilizer (SP Scientific) at 100 mtorr and −80°C overnight or until fully freeze-dried. Following lyophilization, plasmids were rehydrated with nuclease-free water (Ambion) and added to FD-CF reaction pellets at a final concentration of 13.3 µg mL^{-1}, unless otherwise noted. CFPS reactions were carried out at 30°C for 20 hours after rehydration, unless otherwise noted. In a classroom setting, reactions can be incubated in our portable incubator at 30°C or in a 30°C water bath in an insulated container (Styrofoam, plastic cooler, etc.) for 20 hours. Alternatively, reactions can be run in a room temperature water bath or on a tabletop for 40 hours.

5.2.2.5   Quantification of *in vitro*-synthesized protein

Active full-length protein synthesis was measured continuously via fluorescence using the CFX96 Touch Real-Time PCR Detection System (Bio-Rad). If fluorescence saturated the real-time PCR detector,

then endpoint fluorescence was measured in 96-well half-area black plates (CoStar 3694; Corning Incorporated) using a Synergy2 plate reader (BioTek). Excitation (ex) and emission (em) wavelengths used to measure fluorescence of each protein construct were as follows: mCherry, eforRed, mRFP1, and dTomato: ex, 560 to 590 nm; em, 610 to 650 nm; mOrange: ex, 515 to 535 nm; em, 560 to 580 nm; YPet, sfGFP, mTFP1, CyPet, Aquamarine, mTagBFP2, mKalama1, and eBFP2: ex, 450 to 490; em, 510 to 530 nm. Following CFPS, reactions were centrifuged at 20,000g for 10 min to remove insoluble or aggregated protein products before further analysis. To quantify the amount of protein synthesized, two approaches were used. For assessing yields of the full 13-member library, reaction samples were analyzed directly by incorporation of ^{14}C-leucine into trichloroacetic acid–precipitable radioactivity using a liquid scintillation counter, as described previously[111]. These reactions were also run on a Coomassie-stained SDS-PAGE gel and exposed by autoradiography. Autoradiographs were imaged with Typhoon 7000 (GE Healthcare Life Sciences). Following selection of the smaller six-member library, standard curves were generated for mCherry, mRFP1, dTomato, mOrange, and YPet constructs via serial dilution of CFPS reactions containing ^{14}C-leucine and correlating protein yields with measured fluorescence using a standard curve. Fluorescence units of sfGFP were converted to concentrations using a standard curve, as previously described[138].

For quantification without a spectrophotometer, reactions can be semiquantitatively analyzed via imaging using one of our portable, low-cost imagers and subsequent fluorescence analysis in ImageJ, a free image-processing program (imagej.nih.gov/ij). Images of FD-CF reactions were taken with a digital single-lens reflex (DSLR) camera and arranged in Adobe Illustrator. Protein production can also be qualitatively assessed with the naked eye under white light or blue or black light using our portable blue light imagers or others [for example, Bio-Rad ultraviolet (UV) pen lights #1660530EDU, Walmart black light bulb with fixture #552707607, Home Science Tools portable UV black light #OP-BLKLITE, and miniPCR blueBox transilluminator #QP-1700-01].

5.2.2.6   Construction of portable imagers and incubators

To design our portable laboratory equipment, we used the open-source three-dimensional CAD modeling software FreeCAD. Open-source tutorials for FreeCAD are also available on their website

(freecadweb.org). Designed acrylic or wood components were laser-cut to desired specifications (folder S1) and assembled using adhesive (SCIGRIP Weld-On 16 for acrylics or Gorilla Wood Glue for wood components). Individual acrylic or wood parts were gently pressed together by hand for about a minute and left to cure overnight. Electronic components were soldered, and heat shrink was applied as necessary. Once the incubator circuit was assembled, it was mounted onto the incubator with 0.25-inch screws through laser-cut and/or predrilled pilot holes.

After the incubator was assembled, the set temperature was calibrated. For the switch version of the incubator, various resistors or resistor combinations were tested to achieve the two desired temperature set points (30° and 37°C). For the dial version of the incubator, the potentiometer position was adjusted to reach the desired set points. In both cases, the temperature was monitored using an Arduino and, once determined, the set positions were labeled and temperatures were verified through additional temperature monitoring.

5.2.2.7   Statistical analysis

Statistical parameters including the definitions and values of *n*, SDs, and/or SEs are reported in the figures and corresponding figure legends.

5.2.3 Results

5.2.3.1   High-yielding in vitro expression of a diverse fluorescent protein library

Based on the success of colorimetric chemistry kits, we sought to create synthetic biology classroom modules for BioBits™ Bright with simple, visual readouts. We reasoned that the ability to link a visual output to abstract concepts such as the central dogma of molecular biology would increase student engagement and understanding. Fluorescent proteins are routinely used as reporters in synthetic biology and represent an attractive readout for an educational kit for two main reasons. First, a wide variety of fluorescent protein variants have been discovered or engineered[222-227], which produce an array of colors visible to the naked eye. Second, these variants are well studied and documented in freely available databases such as the Protein Data Bank (PDB) (**Table 5.2**), making them ideal instructional tools.

To build BioBits™ Bright, we initially designed a diverse 13-member fluorescent protein library based on existing fluorescent protein variants (**Table 5.2**) and cloned this library into the pJL1 cell-free expression vector. As an open-source kit, we have made these constructs available through Addgene (constructs 102629 to 102640, 106285, and 106320). The library was chosen to include red, orange, yellow, green, cyan, and blue fluorescent proteins. The selected library members represent a diversity of amino acid sequences, with sequence homology to our standard cell-free protein synthesis (CFPS) reporter, a superfolder green fluorescent protein (sfGFP) variant[98], ranging from 90 to 22%. Because of this diversity, and because many of the library members were evolved in the laboratory from naturally occurring fluorescent proteins, the fluorescent protein library could be used to teach evolution, a required subject according to Next Generation Science Standards (NGSS) for K-12 education[228,229]. Plasmids encoding each of the selected library members were used as templates in 5 µl FD-CF reactions lasting 20 hours at 30°C. Yields and full-length expression of all 13 fluorescent proteins were assessed using ^{14}C-leucine incorporation.

a

b

c
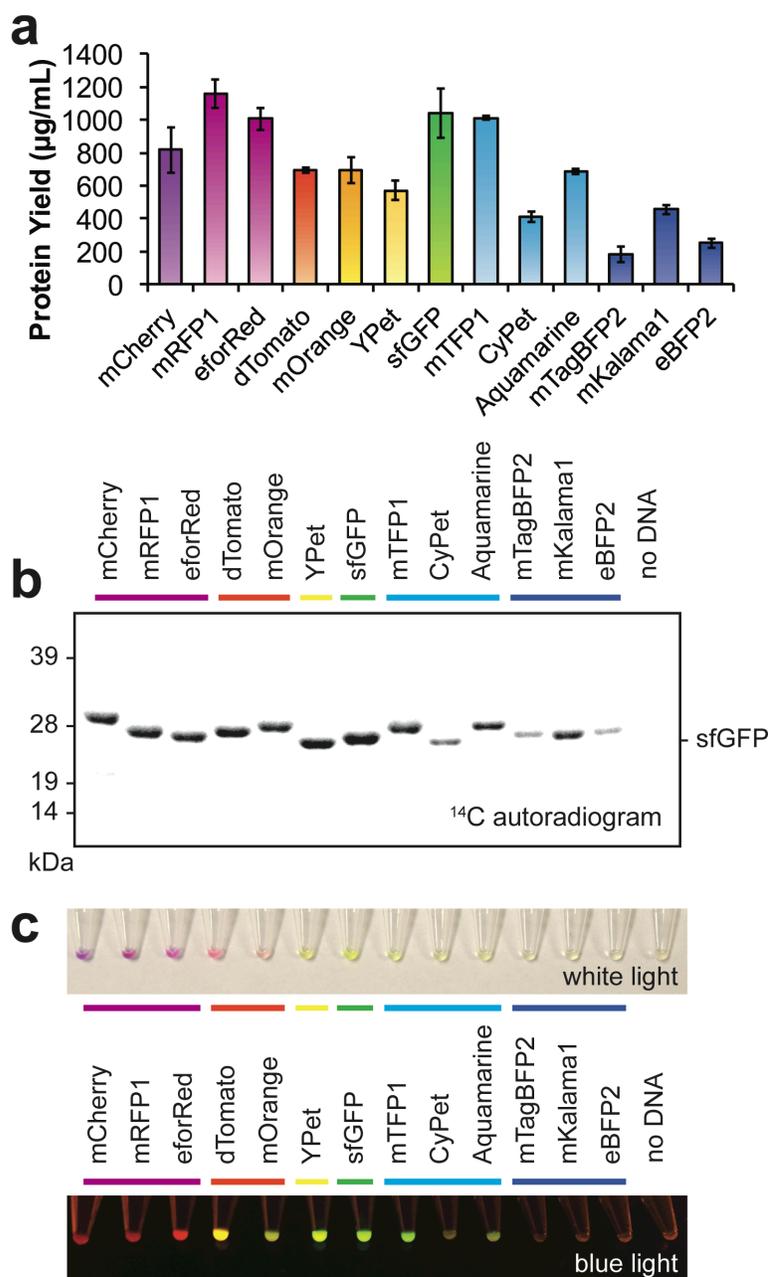
Figure 5.2 High-yielding cell-free production of fluorescent protein library enables development of BioBits™ Bright. A 13-member fluorescent protein library was designed to include red, orange, yellow, green, cyan, and blue fluorescent protein variants and cloned into the cell-free expression vector pJL1. (**a**) Following CFPS for 20 hours at 30°C, soluble yields of the fluorescent protein library were measured via ^{14}C-leucine incorporation. Values represent averages, and error bars represent SDs of $n \geq 3$ biological replicates. (**b**) Soluble fractions were analyzed by SDS-PAGE and ^{14}C autoradiogram. All library members expressed with exclusively full-length products observable by autoradiogram. (**c**) Images of FD-CF reactions expressing the fluorescent protein library under white light (**top**) and blue light (**bottom**).

We observed that all proteins expressed with high soluble yields (between 160 and >1100 μg ml^{-1}) (**Figure 5.2a**) with exclusively full-length products observed by autoradiogram (**Figure 5.2b**). In particular, six fluorescent protein constructs (mCherry, mRFP1, dTomato, mOrange, YPet, and sfGFP) expressed at yields of ≥600 μg ml^{-1} and generated distinct colors and fluorescence visible to the naked eye (**Figure 5.2c**). These results make these six proteins ideal candidates for educational tools, especially in resource-limited classrooms or other nonlaboratory settings. While expression is optimal at 30°C, the six-member library expresses with similar yields (~60% or higher) in reactions incubated at 21°C (room temperature) for 40 hours (**Figure 5.3**). These results indicate that precise temperature control is not required for CFPS, demonstrating that these reactions can be run without an incubator, water bath, or other specialized equipment. Notably, these proteins represent a diversity of amino acid sequences to facilitate evolution curriculum, with between 24 and 89% amino acid sequence homology to sfGFP. For these reasons, these six proteins were selected to form the core set of reagents for BioBits™ Bright, which we next used to develop two educational modules.
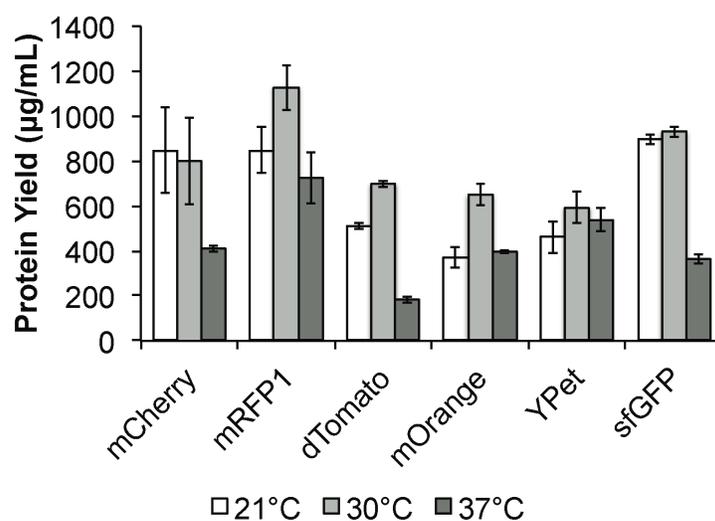
Figure 5.3 FD-CF reactions tolerate a range of incubation temperatures. FD-CF reactions containing DNA template encoding mCherry, mRFP1, dTomato, mOrange, YPet, sfGFP were incubated at 37°C, 30°C, or 21°C. Reactions incubated at 37°C and 30°C were run for 20 hours, while reactions incubated at 21°C were run for 40 hours. Values represent averages and error bars represent standard deviations of n = 3 biological replicates.

5.2.3.2   Module I: Tunable in vitro expression of fluorescent proteins

The first laboratory module demonstrates the ability to control protein synthesis titers by varying the amount of DNA template present in FD-CF reactions, essentially limiting the in vitro transcription and translation reaction for one of its essential substrates. This activity teaches students fundamental biology and synthetic biology concepts such as (i) information flow in the central dogma of molecular biology and (ii) how synthetic biologists can engineer biological systems in predictable ways. Freeze-dried DNA templates encoding mCherry, mRFP1, dTomato, mOrange, and YPet were rehydrated, added to FD-CF reactions in varying amounts (25, 10, 5, 2.5, or 0 ng of DNA), and incubated at 30°C for 20 hours. The sixth library member, sfGFP, exhibited protein synthesis rates between 2 and 10 times faster than the other library members. This relatively high rate of protein synthesis is expected because sfGFP was evolved to exhibit enhanced folding and rapid fluorescence[230]; however, after 20 hours, we were unable to observe discrete variations in protein synthesis with varying amounts of DNA template (**Figure 5.4**). This is not ideal for a typical classroom setting, where teachers will not see students for 24 to 48 hours after reactions are set up. For this reason, sfGFP was excluded from this module.
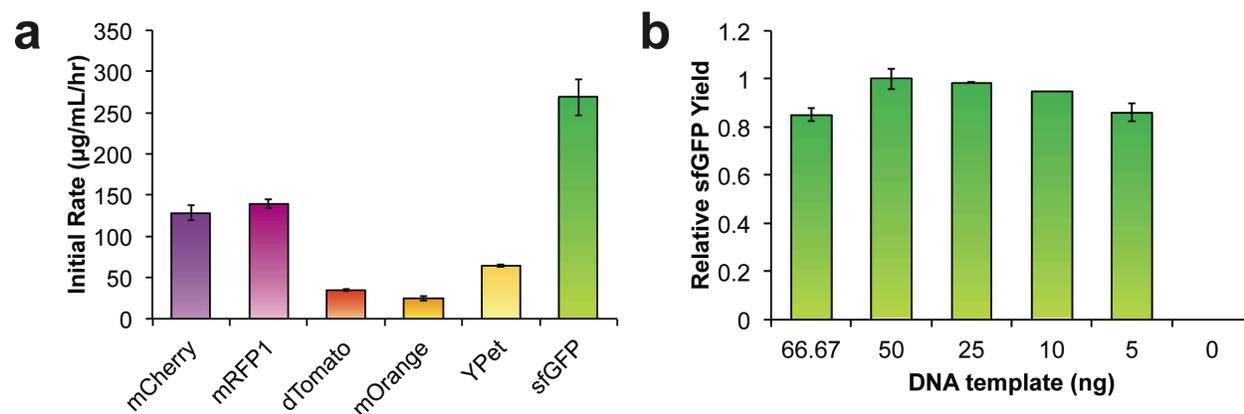
Figure 5.4 DNA template is not limiting for in vitro sfGFP synthesis due to relatively high initial rates of protein synthesis. FD-CF reactions containing DNA template encoding mCherry, mRFP1, dTomato, mOrange, YPet, sfGFP were incubated at 30°C for 20 hours. (**a**) Initial rates of protein synthesis from reactions containing 66.67 ng DNA template were measured by fluorescence. (**b**) Endpoint yields for sfGFP synthesis measured via fluorescence at 20 hours show that protein synthesis is not limited by DNA template concentration. Values represent averages and error bars represent average errors of $n \geq 2$ biological replicates.

FD-CF reactions primed with varying concentrations of the five selected DNA templates were assembled by a graduate student (expert) and compared to those assembled by Chicago middle and high school students and teachers. In all cases, we observed that reducing the concentration of DNA template led to a concomitant decrease in total protein expression, even in reactions assembled by users who were running the BioBits™ Bright laboratory for the first time (**Figure 5.5a**). Visible differences in color and fluorescence showing these trends were observable in all samples under both white and blue light (**Figure 5.5b**). The ability to easily perceive variations in reaction color with the naked eye makes it possible to qualitatively assess protein synthesis yields from this module without a spectrophotometer. Through its easy, visual outputs, this laboratory module helps students understand how proteins are synthesized, as well as some of the key biochemical factors that affect this process (for example, DNA as the instructions that guide protein synthesis). As an extension of the activity presented here, students could investigate factors other than DNA concentration that affect protein synthesis, such as ion concentration, amino acid concentration, or energy substrate concentration, among others[231]. As examples of these activities, we worked with Chicago public high school teachers to develop a set of inquiry-based curricula for this module with emphasis on student-driven experimental design to satisfy NGSS requirements for high school biology (**Curricula 1 and 2**).
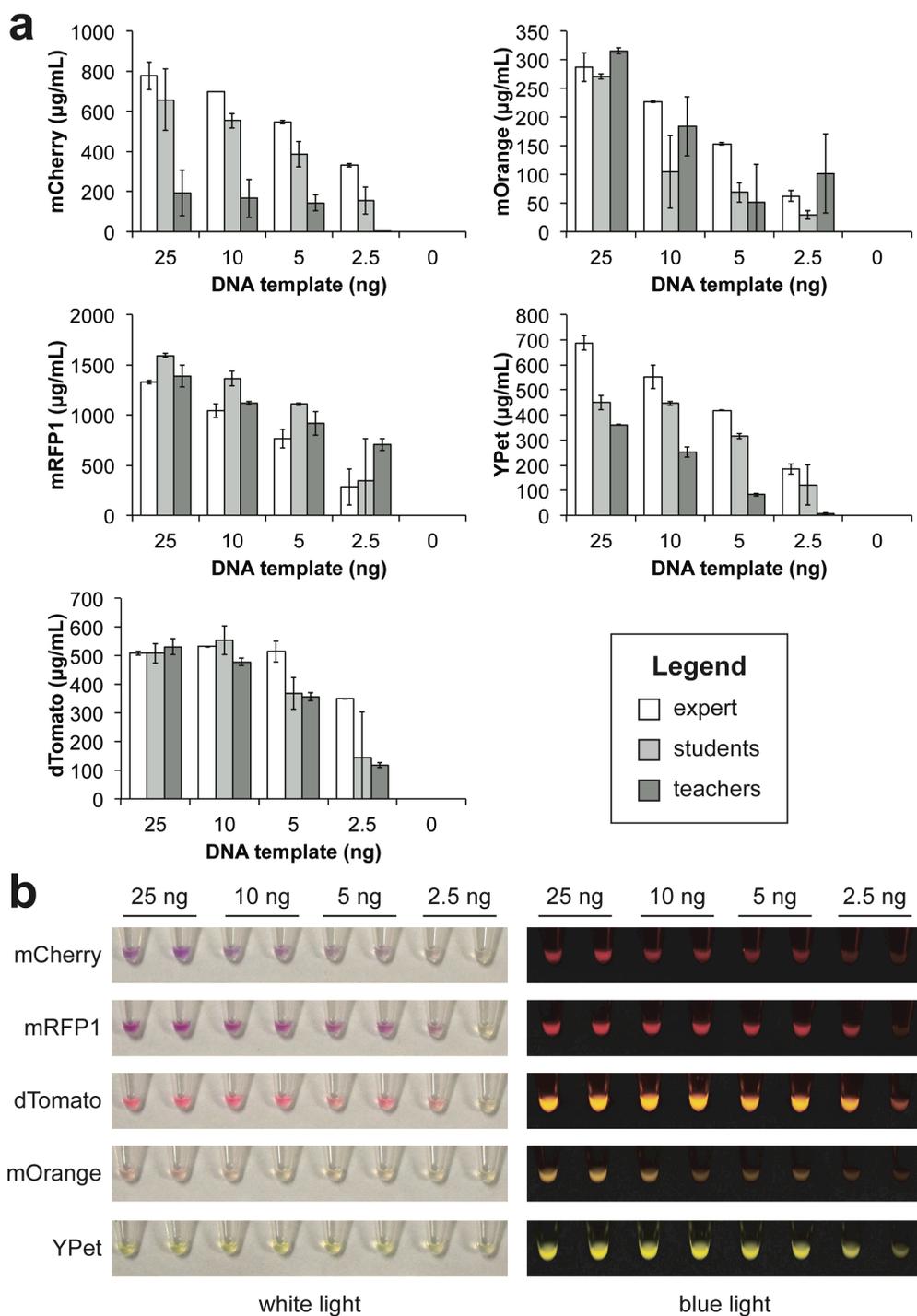
Figure 5.5 Controllable in vitro expression of diverse fluorescent proteins. FD-CF reactions were rehydrated with 25, 10, 5, 2.5, or 0 ng of template DNA encoding mCherry, mRFP1, dTomato, mOrange, or YPet and run for 20 hours at 30°C. (**a**) Results from experiments run by graduate students (experts), high school students, or middle and high school teachers are shown. In all cases, we observed a concomitant decrease in protein synthesis as the amount of DNA template was decreased. Values represent averages, and error bars represent average errors of $n \geq 2$ biological replicates. (**b**) The variation in protein expression was marked enough to be observed qualitatively with the naked eye under

both white light and blue light. Images are representative examples of experiments prepared by high school students.

5.2.3.3 Module II: Design, build, and test an in vitro biological program

The second laboratory module engages participants in a design, build, test (DBT) cycle wherein they create their own in vitro program with DNA. This laboratory recapitulates the idea of controllable protein expression from module I, introduces the DBT cycle as a key synthetic biology and engineering concept, and could pair with a brief research project to introduce students to the broader field of synthetic biology (for example, **Curriculum 3**). Specifically, participants were given a 96-well PCR plate containing 5 µl FD-CF reactions and separately freeze-dried plasmid templates. Programs could be constructed by rehydrating FD-CF reactions with any of the six-member fluorescent protein library members at any of the concentrations tested in the first laboratory module (0 to 25 ng of total template DNA). Participants designed, built, and tested their in vitro programs by carrying out protein synthesis for 20 hours at 30°C.

We ran this activity with students of varying ages, ranging from preschool-aged students to high school teachers, and observed a number of successful designs (**Figure 5.6**). This module's educational merit is twofold. First, this activity engages students in the engineering process, helping them go beyond simple pipetting and reagent handling for a self-directed, independent learning experience. Second, this module bridges the gap between science and art, offering an opportunity for incorporation of emerging interdisciplinary STEAM ideologies into biology curriculum, which have reported improved educational outcomes[232]. One participant described this laboratory as a "biological Lite Brite," highlighting the design component of this module and the potential for students' creative innovation within this laboratory activity. Of note, sample curricula for high school math (**Curriculum 4**) and middle school science classes (**Curriculum 5**) were developed in partnership with Chicago area teachers, emphasizing the laboratory's cross-cutting nature and the value of this activity at various educational levels.
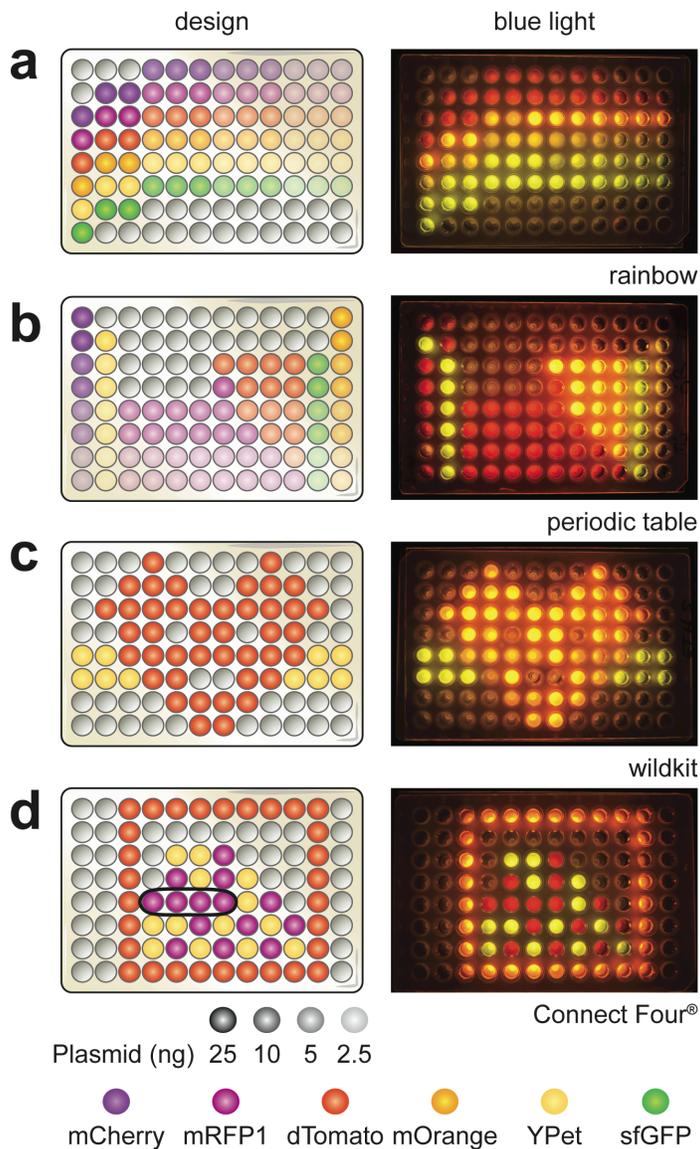
Figure 5.6 Design and execution of in vitro programs. Participants were asked to design, build, and test their own in vitro program with DNA in a 96-well PCR plate. Designs could include the mCherry, mRFP1, dTomato, mOrange, YPet, or sfGFP plasmids at concentrations between 0 and 25 ng (same template concentrations tested in module I), denoted with corresponding colors and opacity in the pictured designs (legend, bottom left). Successful designs included (**a**) a rainbow, (**b**) a periodic table, (**c**) a wildkit (the Evanston Township High School mascot), and (**d**) a game of Connect Four®. These biological programs were designed, built, and tested by untrained operators, demonstrating the potential of this laboratory for use in a classroom setting.

5.2.3.4   Portable, low-cost imagers and incubators for taking BioBits™ beyond the laboratory

Recognizing that a vast majority of classrooms will not have laboratory-grade fluorescent imagers or incubators to run FD-CF reactions, we developed affordable and portable versions to make the BioBits™ Bright laboratory activities accessible to resource-limited classrooms. Specifically, we developed two compact, battery-powered imagers for visualizing FD-CF reactions producing fluorescent proteins. One imager is designed to accommodate eight-strip PCR tubes for imaging DNA titration experiments, while the second is designed for imaging 96-well plates containing in vitro biological programs. Both systems faithfully image the fluorescent protein library and have the same key components: a single 450-nm light-emitting diode (LED) light, colored acrylic plates to filter out the inherent color of the LED for fluorescence visualization, and a laser-cut casing to house the system (**Figure 5.7a, b**). The initial prototypes for the 8-well and 96-well imagers cost about $15 and $32, respectively, to build (**Table 5.3**). We also developed two versions of a USB (universal serial bus)–powered incubator: one in which temperature is controlled by a switch calibrated to two temperature settings, 30° or 37°C (**Figure 5.7c**), and one with a dial to enable any temperature setting between 30° and 37°C. Both versions perform similarly and can be built in schools with fabrication workshops for less than $20 (**Table 5.3**).

We tested the expression of our six-member fluorescent protein library at 30°C in our portable incubator and observed at least 50% of protein yields achieved using a thermocycler, with fluorescence easily observable in our handheld eight-well imager (**Figure 5.7c**). As an example of cross-cutting STEAM education integrating engineering, fabrication, electronics, and synthetic biology, the BioBits™ Bright computer-aided design (CAD) files (**Folder S1**) can be used with the open-source FreeCAD software and accompanying circuit diagrams (**Folder S1**) to enable students to manufacture their own portable imager or incubator for use in subsequent experiments.
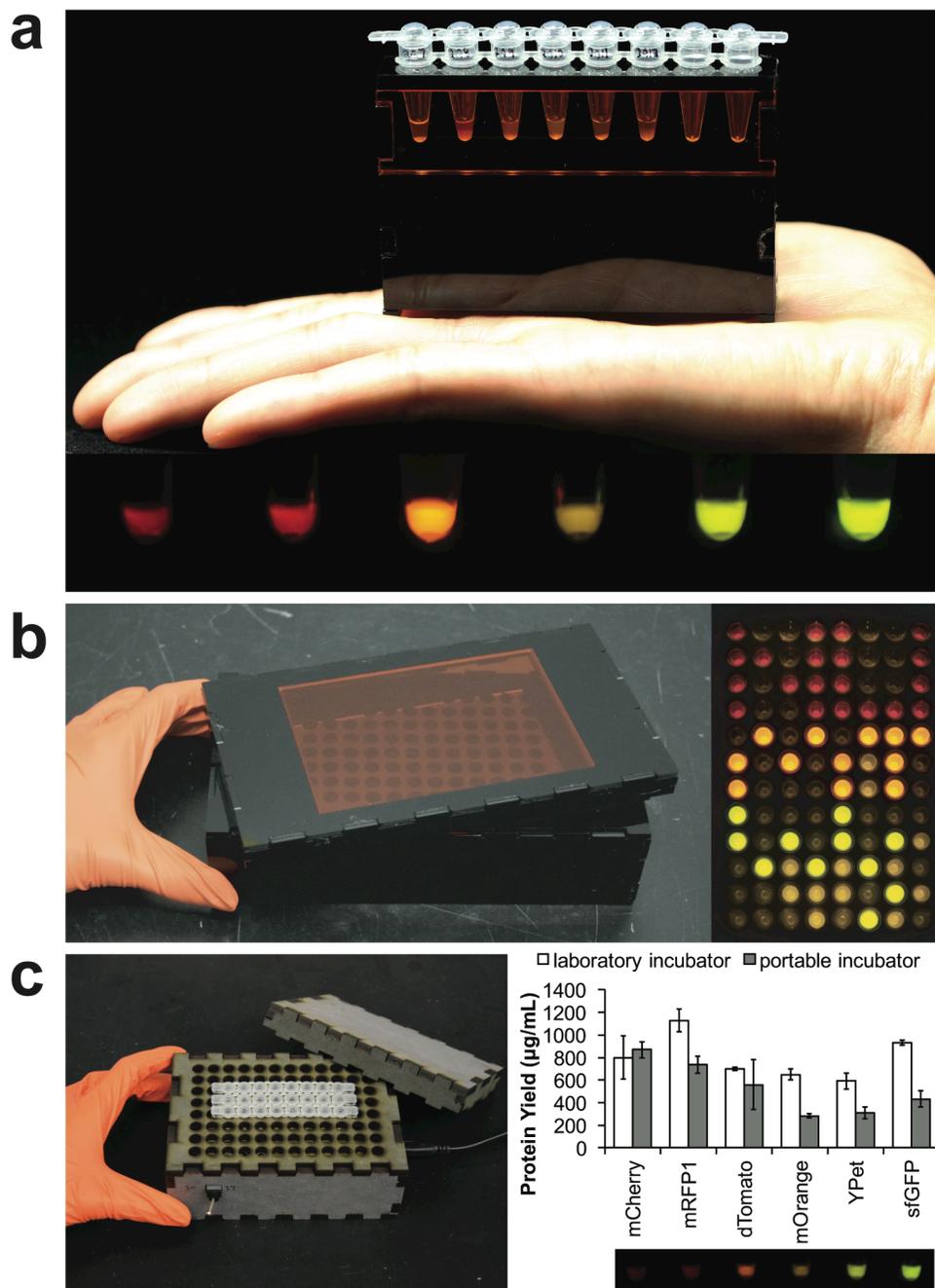
Figure 5.7 Portable, low-cost equipment for teaching outside of the laboratory. (**a**) The eight-well imager is handheld and battery-operated for easy use (**top**) and can be used to image the six-member fluorescent library (**bottom**). We show FD-CF reactions expressing, from left to right, mCherry, mRFP1, dTomato, mOrange, YPet, and sfGFP. (**b**) The 96-well imager is also battery-powered and has a removable lid for easy use (**left**). In vitro biological programs can be imaged using our custom 96-well imager with similar performance as a laboratory imager (**right**). (**c**) The portable incubator accommodates up to 96 standard PCR tubes and has a removable, insulating lid for maintaining reaction temperature at its two set points, 30° and 37°C (**left**). Fluorescent protein yields using our incubator set at 30°C are at least 50% of those achieved using a laboratory incubator (**top right**) and produce fluorescence that is

visible in our handheld eight-well imager (**bottom right**). Values represent averages, and error bars represent average errors of *n* = 2 biological replicates.

With the portable imagers and incubators at hand, we were able to demonstrate that FD-CF reactions can be run in a "laboratory-free" environment, using our portable incubator, imager, and disposable exact-volume transfer pipettes (VWR 89497-718) to rehydrate the reaction. Reactions run in the laboratory (with laboratory pipettes, incubators, and imagers) are comparable to those run with our kit components and are visually consistent across different experiments and different operators (**Figure 5.8**).
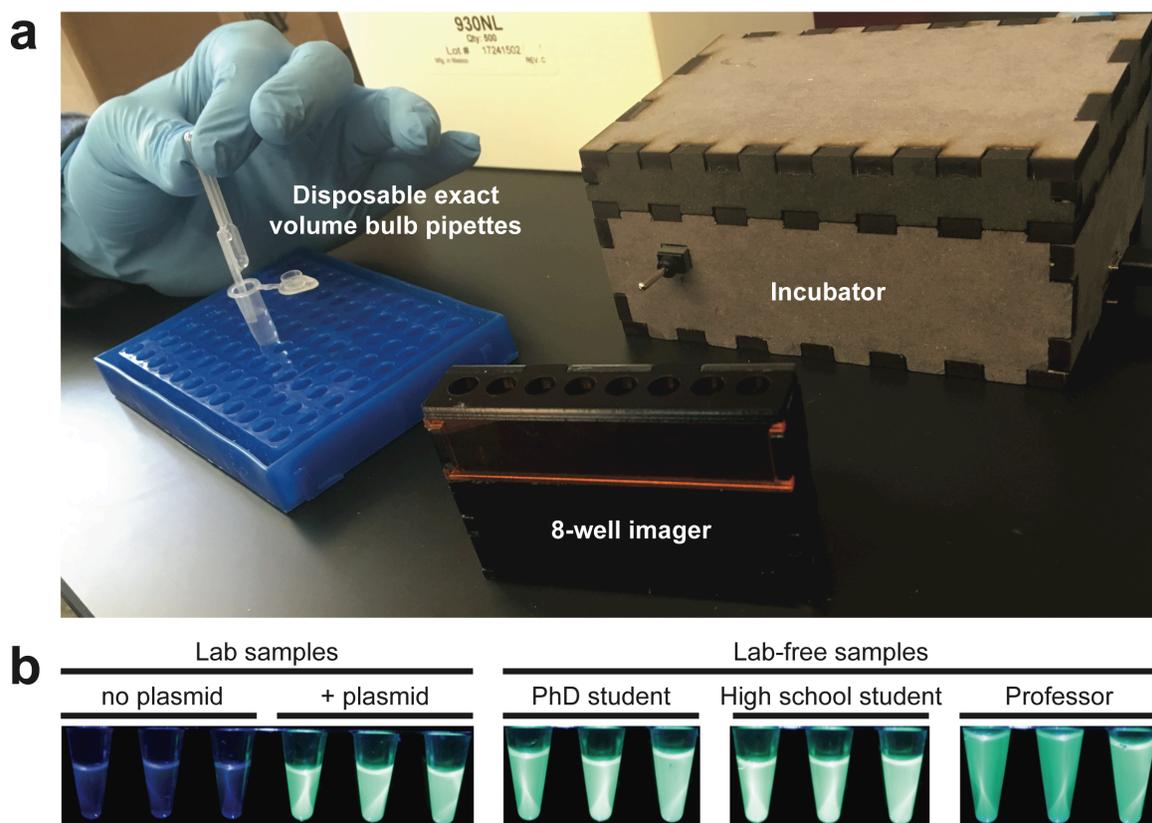
Figure 5.8 FD-CF reactions can be run in a laboratory-free environment using low-cost, portable imagers and incubators. (**a**) Equipment used in "lab-free" experiments, including disposable 50 μL transfer pipettes, a portable imager, and a portable incubator. (**b**) sfGFP expression is visually consistent across different experiments and different operators. All images of reactions are scaled identically; variations in the volume of the reactions are due to pipetting differences across individual operators.

5.2.4 Discussion

We present here the BioBits™ Bright educational kit and an accompanying collection of resources and data for teaching synthetic biology outside of the laboratory. To develop the fluorescent reagents, we assembled a fluorescent protein library that expresses at high yields in FD-CF reactions. We further demonstrated that both DNA templates encoding this library and cell-free reactions could be freeze-dried and reconstituted by just adding water, providing the necessary reagents for portable educational tools. Furthermore, we developed two laboratory modules designed to teach students about synthetic biology and successfully tested these modules with Chicago K-12 teachers and students. For both laboratory modules, we report data generated by both teachers and students, demonstrating the utility of these resources for use by untrained operators without sophisticated laboratory equipment.

In the first laboratory module, participants investigate how protein expression in FD-CF can be tuned by adding varying amounts of DNA template. This activity can be used to introduce the central dogma of molecular biology or the idea of tunable protein expression (for example, **Curricula 1 and 2**). This module also reinforces basic biology concepts by demonstrating how variations in gene/protein sequence can affect protein function, since differences in protein sequence result in distinct protein properties (visible differences in protein color and fluorescence).

For more advanced groups, differences in protein synthesis rates and final titers can be measured and quantified to investigate how protein synthesis can be modeled as an enzymatic reaction and how kinetics can be controlled by changing the amount of substrate (DNA template). Alternatively, students can carry out the same investigation using sample kinetic data we collected from student-assembled reactions (**Data S1**). Long-term independent science projects can also be conceived by incorporating complementary biochemistry and molecular biology experiments, such as one project we recently designed with a high school synthetic biology after-school club. In this example, students used FD-CF reactions to synthesize the human leptin hormone as a potential treatment for obesity and quantified the amount produced using a commercial enzyme-linked immunosorbent assay (ELISA)[233].

In the second laboratory module, participants design, build, and test their own in vitro program with DNA. This laboratory demonstrates how in vitro biological systems can be engineered to produce

outputs of interest. This module primes students for discussion of synthetic biology and potential application areas (for example, therapeutic protein production, sustainable chemical production, and cellular/organismal engineering) and the ethics involved in the field (for example, **Curriculum 3**). In addition, by engaging participants in a self-directed DBT cycle, this module offers a straightforward way to incorporate engineering principles into biology curriculum. Finally, the simple framework of this module encourages creative innovation through STEAM principles. The potential for such opportunities are highlighted by the complementary design activity (**Curriculum 5**) and math curriculum piece we have developed (**Curriculum 4**), as well as the availability of FreeCAD and our open-source design files to enable students to build their own portable fluorescence imagers and incubators (**Folder S1**).

Importantly, BioBits™ Bright makes even more educational resources possible, perhaps through the formation of an open-source community. For example, next-generation iterations of these kits could incorporate antibiotic ribosome inhibitors for tuning protein expression, offering opportunities for educators to discuss health-related themes in class. In addition, coexpression of two or more fluorescent proteins or incorporation of synthetic genetic circuits[234] to control fluorescent protein expression would introduce students to more complex examples of biological regulation. Further, engagement of students through different sensory outputs could improve student engagement and understanding, which will empower them to make informed decisions about cutting-edge synthetic biology topics [for example, clustered regularly interspaced short palindromic repeats (CRISPR)–Cas9 genome editing][235]. We have addressed some of these needs through the development of a next-generation kit: BioBits™ Explorer[66]. The Explorer kit expands the toolbox of educational materials for teaching synthetic biology and provides additional opportunities for student-driven, independent synthetic biology investigations. Beyond this, future work could expand the parallels between engineering, biology, and design, such as through the integration of a novel phone application and LED display to aid the design of in vitro biological programs in module II of the BioBits™ Bright kit[236]. We also plan to launch a website where students can share their data and biological program designs with other users of these kits from around the world.

In sum, BioBits™ Bright represents a comprehensive set of educational resources for synthetic biology akin to the "chemistry set" that brought chemistry education to the masses and inspired

generations of scientists. We have purposely designed our kit to be economically accessible, priced at less than $100 per 30-person classroom (**Table 5.3**). This is made possible by our in-house freeze-dried reactions, which are two orders of magnitude more affordable than existing commercial cell-free kits, at just ~$0.01 per microliter of reaction volume (**Table 5.5**) compared to ~$1 per microliter (Promega L110; NEB E6800S). Our custom imagers and incubators are included in BioBits™ Bright, making reaction analysis accessible for resource-limited classrooms. Because of the highly portable, cost-effective, and user-friendly nature of the reagents and laboratory activities, the BioBits™ Bright and Explorer kits have utility both inside and outside of a formal classroom or laboratory setting. In sum, these resources promise to increase access to cell-free technologies, enhance basic biology education, and increase participation and teaching in the field of synthetic biology.

5.3 BioBits™ Explorer: A modular synthetic biology education kit

5.3.1 Introduction

Many of us can trace our initial fondness for the sciences to formative experiences with hands-on exploratory kits, such as traditional chemistry sets. This trend has expanded today to include a spectrum of educational kits that teach subjects such as physics, electronics, programming, or robotics[237-239]. However, there are few successful and engaging systems for teaching advanced molecular or synthetic biology concepts in a hands-on manner[240,241]. This absence is largely due to the particularities of traditional biology experimentation, which requires a cold chain to prevent the biological components from spoiling, sterile equipment and media to prevent contamination, specialized instruments such as shaking incubators, and concerns with the biocontainment of recombinant microorganisms. Here, we present the development of a synthetic biology platform that circumvents all of these challenges, resulting in a shelf-stable and affordable educational kit for demonstrating advanced biological concepts.

Synthetic biology is a rapidly advancing field that uses engineering concepts to harness the power and diversity of biology. At the foundation of this endeavor is the ability to control gene expression in a predictable manner, which is accomplished by using modular biological components to control and fine tune the processes of transcription and translation[242,243]. The resulting synthetic biology toolbox enables powerful new methods for chemical and drug manufacturing[242,244], clinical diagnostics[51,220], and

cell therapies[208,245]. Synthetic biology kits also have great potential as educational tools to teach molecular and synthetic biology concepts but are generally too expensive to implement in classrooms due to the numerous infrastructure requirements of these types of experiments.

To create an array of biology demonstrations that could be used in any classroom setting, we turned to cell-free synthetic biology. Cell-free systems use essential cellular machinery, including polymerases, ribosomes, and transcription factors, in an in vitro setting to carry out the processes of transcription and translation, which circumvents the need for specialized, sterile equipment and media to culture living cells; moreover, the lack of living cells eliminates concerns of biocontainment. There are two general types of cell-free systems: crude extracts, where the required cellular components are harvested from bacterial lysate[102], and reconstituted systems, such as the commercial protein synthesis using recombinant elements (PURE) system[246], where each individual component is produced recombinantly and then recombined *in vitro*. Both systems require supplementation with additional essential components such as nucleotides, amino acids, and energy equivalents. Cell-free systems have been used extensively to produce proteins and other biomolecules, as well as build and execute synthetic biology circuits[12,39,50,51,220].

We have shown that cell-free systems can be freeze-dried along with genetic elements to form pellets that are stable at room temperature and are highly portable[12]. The shelf-stable nature of these freeze-dried, cell-free (FD-CF) pellets eliminates the need for dedicated refrigerators or freezers. In addition, FD-CF reactions do not require any specialized equipment, making them a robust technology for using synthetic biology in low-resource environments, including classrooms. Reactivation of the FD-CF components simply requires the end user to add water. We have previously used this technology for the rapid development of inexpensive, paper-based nucleic acid diagnostics and as a portable biomanufacturing platform[12,50,51]. With the unique practicality of FD-CF technology, we also considered this platform to be highly suitable for applications in biology education, where there is a glaring lack of hands-on biology experiments[65]. Specifically, FD-CF reactions are an ideal way to bring the ever-increasing toolbox of the synthetic biology community to secondary schools and the general public (**Figure 5.9a**). Previously, this has required substantial investment in laboratory equipment and

infrastructure, resulting in the lack of formative STEM experiences in poorly funded schools[247]. We believe that the innovative approach we present here and in Stark *et al.*[248] will have a significant impact on lowering the barriers to explore advanced synthetic biology concepts and reduce inequalities in public science education.
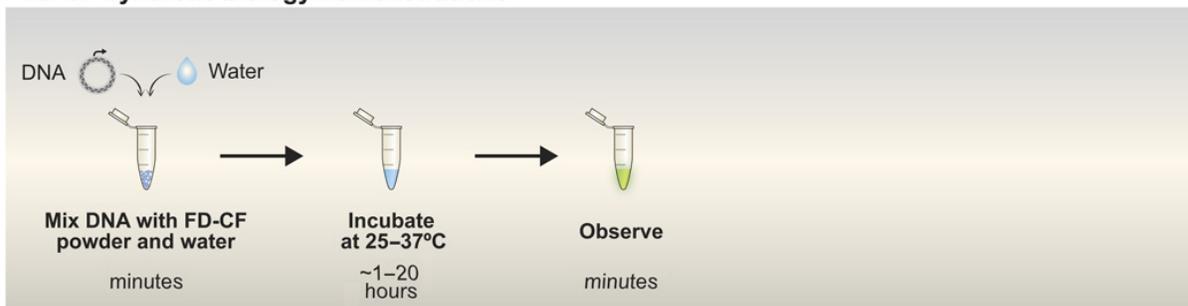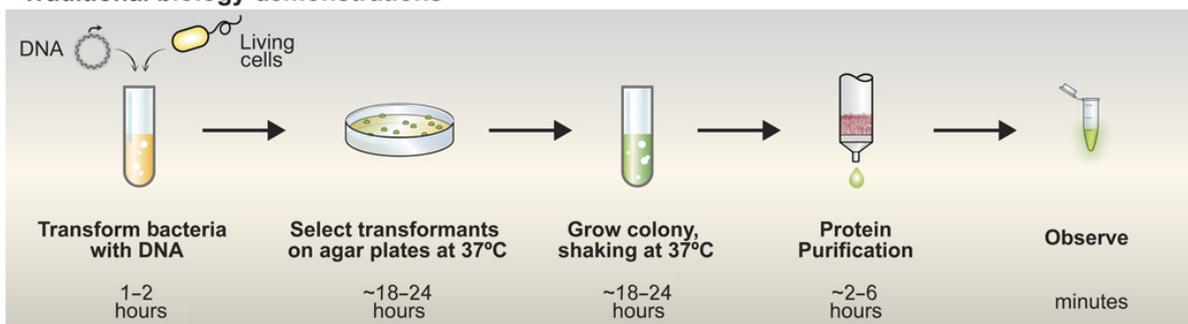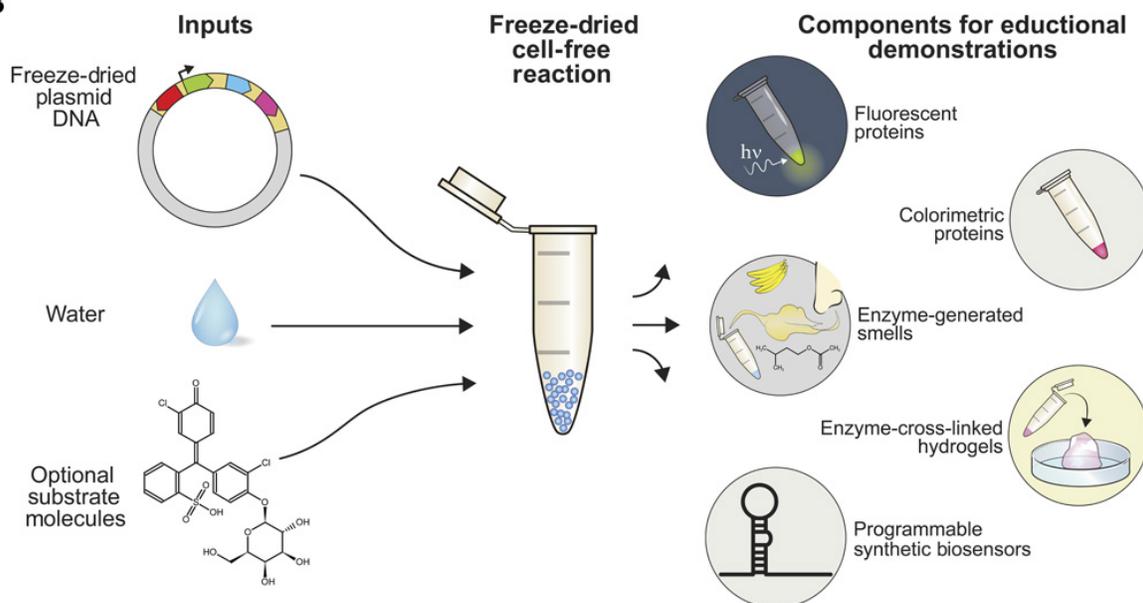
Figure 5.9 BioBits™ kits: Freeze-dried educational kits. (**a**) FD-CF demonstrations require only the addition of water to the supplied reactions and incubation for 1 to 20 hours at 25° to 37°C for observation and analysis by students. In contrast, traditional biology experiments require substantial time, resources, and specialized equipment. (**b**) With the DNA template and any substrate molecules provided with the FD-CF reaction, the students just have to add water to run a number of bioscience activities and demonstrations.

Here, we introduce BioBits™ Explorer, a low-cost modular educational kit that uses FD-CF technology to teach synthetic biology concepts through sensory engagement and provides opportunities for inquiry-based learning. We have developed a set of demonstrations designed to engage three of the five senses—sight, smell, and touch (**Figure 5.9b**)—through the expression of proteins in FD-CF reactions that produce fluorescence, enzyme-generated fragrances, and large-scale hydrogels, respectively (**Figure 5.10**; **Table 5.1**). This was made possible through the development of functionally robust synthetic cell-free programs—several of which are original. We discuss here how these outputs can be used to create activities to teach the fundamentals of protein expression, enzyme catalysis, and properties of biomaterials. In addition, we incorporate modular biosensing components that can be used to control gene expression—specifically, RNA toehold switches—to develop a demonstration that allows students to discriminate between species of different fruits using extracted DNA. These activities can be run on their own or in sequence with additional laboratory activities that we developed using fluorescent protein outputs, which we pair with low-cost, portable laboratory equipment and supporting curricula in a kit we call BioBits™ Bright[248]. Together, the BioBits™ kits demonstrate both the breadth of synthetic biology activities that can be developed with FD-CF technology and how these platforms can increase student involvement, illustrate core concepts in molecular and synthetic biology, and provide opportunities for independent, student-directed research projects (for example, synthetic biology after school clubs and science fair research teams) in the life sciences.
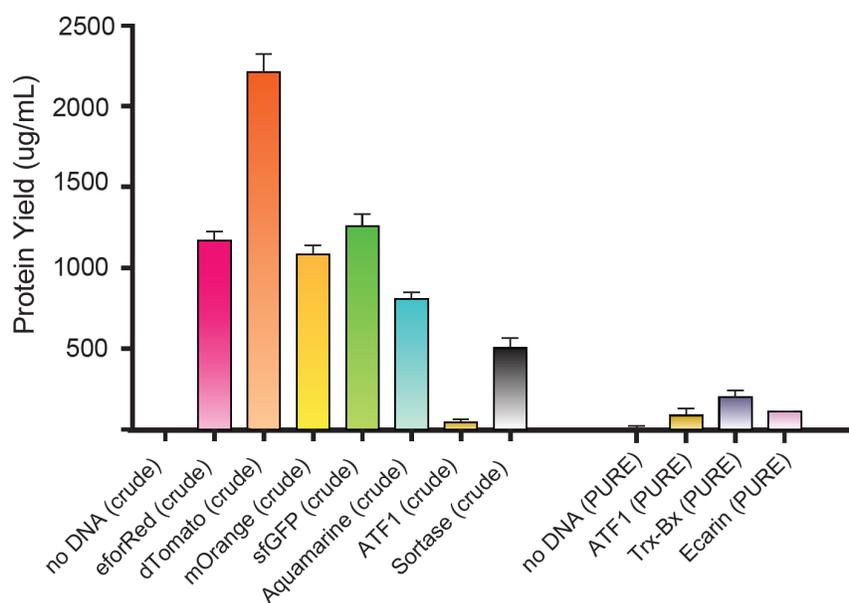
Figure 5.10 Quantification of all proteins expressed in FD-CF. All of the FD-CF expressed proteins used in the demonstration experiments had high soluble yields (between 100 and >1000 µg/mL), as measured by ^{14}C leucine incorporation. Values represent averages and error bars represent standard deviations of *n* = 3 biological replicates.

5.3.2 Methods

5.3.2.1 General template design and preparation

DNA sequences encoding eforRed, dTomato, mOrange, ATF1, Ecarin, and Trx-Bx (batroxobin fused with thioredoxin as a solubility domain) genes were derived from the literature, codon-optimized for *Escherichia coli*, and synthesized as gBlocks or oligonucleotides by Integrated DNA Technologies. pPROEX-Aquamarine was a gift from F. Merola (plasmid #42889, Addgene), and pET29-sortaseA-penta-mutant was a gift from L. Griffith. We previously reported the pJL1-sfGFP plasmid (plasmid #69496, Addgene). Cloning and plasmid propagation were performed using either Mach1 (C862003, Thermo Fisher Scientific) or NEB Turbo (C2984H, New England Biolabs) competent *E. coli* cells. All templates were cloned into a T7 expression plasmid system—the PURExpress control vector from New England Biolabs, called pNP1 in the text, pJL1, or pCOLADuet-1 (71406-3, Novagen)—using Gibson assembly[249]. All template plasmid DNA preps of the plasmids were performed with the E.Z.N.A. Plasmid Midi Kit (#D6904, Omega Bio-Tek) for crude extract reactions or the QIAprep Spin Miniprep Kit (#27106, Qiagen) for PURE reactions. All sequences are available on Addgene (**Table 5.1**).

5.3.2.2   PURE cell-free reaction preparation and lyophilization protocol

For cell-free reactions performed in the PURExpress In Vitro Protein Synthesis Kit (E6800S, New England Biolabs), the reactions consisted of the following: NEB Solution A (40%) and B (30%), ribonuclease inhibitor (0.5%; 03335402001, Roche), and the template DNA (10 to 50 nM). For the FD-CF expression of ATF1, the Disulfide Bond Enhancer (E6820S, New England Biolabs) was added into the reactions, as per the manufacturer's instructions, before lyophilization. The reactions were then flash-frozen in liquid nitrogen, lyophilized overnight to obtain the freeze-dried reaction, and stored at room temperature. The reactions were reconstituted with nuclease-free water to the original reaction volume and incubated at 30° or 37°C.

5.3.2.3   In-house crude cell-free extract preparation and lyophilization protocol

Cell extract was prepared as described previously[102]. Briefly, *E. coli* BL21 Star (DE3) cells (Thermo Fisher Scientific) or a BL21 variant called RARE[250] was grown in 150 ml of LB at 37°C at 250 rpm. Cells were harvested in mid-exponential growth phase [OD_{600} (optical density at 600 nm) = ~2 to 3], and cell pellets were washed three times with ice-cold Buffer A containing 10 mM tris-acetate (pH 8.2), 14 mM magnesium acetate, 60 mM potassium glutamate, and 2 mM dithiothreitol, flash-frozen, and stored at −80°C. Briefly, cell pellets were thawed and resuspended in 1 ml of Buffer A per 1 g of wet cells and sonicated in an ice water bath. Total sonication energy to lyse cells was determined by using the sonication energy equation for BL21-Star (DE3) cells, [Energy] = [[volume (μl)] − 33.6]*1.8^{-1}. A Q125 Sonicator (Qsonica) with 3.174-mm-diameter probe at a frequency of 20 kHz was used for sonication. An amplitude of 50% in 10-s on/off intervals was applied until the required input energy was met. Lysate was then centrifuged at 12,000 relative centrifugal force (rcf) for 10 min at 4°C, and the supernatant was incubated at 37°C at 300 rpm for 1 hour. The supernatant was centrifuged again at 12,000 rcf for 10 min at 4°C, flash-frozen, and stored at −80°C until use.

The reaction mixture consists of the following components: 1.2 mM adenosine 5′-triphosphate; 0.85 mM each of guanosine-5′-triphosphate, uridine 5′-triphosphate, and cytidine 5′-triphosphate; L-5-formyl-5,6,7,8-tetrahydrofolic acid (34.0 μg ml^{-1}; folinic acid); *E. coli* transfer RNA mixture (170.0 μg ml^{-1}); 130 mM potassium glutamate; 10 mM ammonium glutamate; 12 mM magnesium glutamate; 2 mM each

of 20 amino acids; 0.33 mM nicotinamide adenine dinucleotide; 0.27 mM CoA; 1.5 mM spermidine; 1 mM putrescine; 4 mM sodium oxalate; 33 mM phosphoenolpyruvate; plasmid (13.3 µg ml^{-1}); T7 RNA polymerase (100 µg ml^{-1}); and 27% (v/v) of cell extract[101,111]. The reactions were then flash-frozen in liquid nitrogen, lyophilized overnight to obtain the freeze-dried reaction, and stored at room temperature. The reactions were reconstituted with nuclease-free water to the original reaction volume and incubated at 30° or 37°C.

5.3.2.4   Fluorescent protein production and characterization

The FD-CF synthesized fluorescent proteins were expressed at 30°C (for the constitutively expressed outputs) or 37°C (for the toehold-encoded designs) overnight and visualized using a Safe Imager 2.0 Blue-Light Transilluminator (Thermo Fisher Scientific), white light, or the inexpensive imager developed as part of the BioBits™ kit. Images were taken with a DSLR camera and adjusted and cropped in Adobe Photoshop. For quantitative analysis, cell-free reactions were transferred to a 384-well clear-bottom, black-walled plate, and relative fluorescent units were read on a SpectraMax M3 Multi-Mode Microplate Reader (Molecular Devices).

5.3.2.5   Smell production and characterization

FD-CF reactions for the expression of ATF1 enzyme were incubated at 37°C for 20 hours in the cell-free reaction. The completed FD-CF reaction containing the enzymes was then added into a separate freshly prepared catalysis reaction. The total catalysis reaction volume was 300 µl and included 50 mM HEPES (pH 7.5), 100 mM KCl, 5 mM EDTA, the relevant substrates (25 mM isoamyl alcohol for ATF1), and freshly prepared cofactor (5 mM acetyl-CoA for ATF1), and 10% of the volume was the FD-CF reaction containing the enzyme. These reactions were allowed to proceed 20 hours in capped vials at room temperature. For GC-MS analysis, the stir bar sorptive extraction method[251] was used. Polydimethylsiloxane stir bars (GERSTEL 011222-001-00) were held in the headspace of the reaction vial by a magnet during the catalysis reaction to absorb volatile components. After the completion of the reaction, the stir bar was added to a headspace vial containing 100 µl of dodecane/ethanol (10:1) and analyzed on a GC-MS headspace sampler (Agilent 7697A) to confirm the identity of the converted product. The GC-MS total ion count signal was converted to parts per million by generating a standard

curve using the same process described above, but the FD-CF reactions did not contain DNA template or substrates but were spiked instead with known parts per million concentrations of the product isoamyl acetate.

5.3.2.6   Hydrogel production and characterization

For the sortase hydrogel peptides, eight-arm PEG vinyl sulfone MW 20,000 Da (PEG-VS) was purchased from JenKem Technology. The cross-linking peptides GCRELPRTGG and GGGSGRC were custom-synthesized by CPC Scientific. Each peptide (8 mM) was conjugated separately to 1 weight % (wt %) PEG, dialyzed, lyophilized, and then reconstituted to 30 wt %. FD-CF reactions were used to generate enzymatically cross-linked hydrogels using a two-step process. First, FD-CF reactions containing a sortase-, ecarin-, or Trx-Bx–encoding template were reconstituted with nuclease-free water and incubated at 37°C. Following incubation, the hydrated sortase reaction was added to a solution of 0 to 8% PEG-GCRELPRTGG and 0 to 8% PEG-GGGSGRC in a reaction buffer [50 mM HEPES, 150 mM NaCl, and 10 mM CaCl$_2$ (pH 7.9)] and incubated at 37°C for 30 min. The hydrated ecarin or Trx-Bx reactions were added separately to a solution of bovine fibrinogen (17.5 mg/ml), aprotinin [2.3 TIU (trypsin inhibitor unit)/ml], and 20 mM CaCl$_2$ and incubated overnight at room temperature. The hydrogels were transferred to glass vials and inverted to demonstrate the hydrogel properties. Images were taken with a DSLR camera and adjusted for size and contrast in Adobe Photoshop. For ultrastructural analysis of the resulting hydrogels, the samples were extensively washed with nuclease-free water, snap-frozen in liquid nitrogen, and lyophilized to remove all water. The hydrogel samples were then sputter-coated with 5 nm of Pt/Pd before imaging using a Zeiss Supra55VP FE-SEM.

5.3.2.7   DNA extraction and processing from fruit

Household dish soap was diluted 1:10 in water along with 1 g of table salt and then added to a plastic bag containing chopped fruit (banana, kiwi, or strawberry). The fruit was then gently crushed in the soap and salt mixture by hand until a homogeneous mixture was obtained. The resulting mixture was strained through a household coffee filter into a cup. A prechilled 25-ml volume of 91% isopropyl alcohol (rubbing alcohol) was added to the strained liquid. The mixture was left undisturbed for 5 min to allow phase separation to occur. The upper white layer containing extracted DNA was removed, placed on a

clean coffee filter, and washed with 70% ethanol (ethyl rubbing alcohol). The resulting extracted DNA was then patted with paper towels to remove any excess extraction liquid. The DNA was then diluted in water until it dissolved and added to an isothermal RPA, according to the manufacturer's protocol (TwistAmp Basic RT, TwistDx; **Figure 5.11**), with primers that were complementary to one section of the banana or kiwi genome. The primers also incorporated a T7 promoter for transcription in FD-CF. The resulting RPA product was then added 1:3.75 to a rehydrated FD-CF reaction containing a linearized toehold complementary to the amplified RPA product and run according to the FD-CF methods described above.
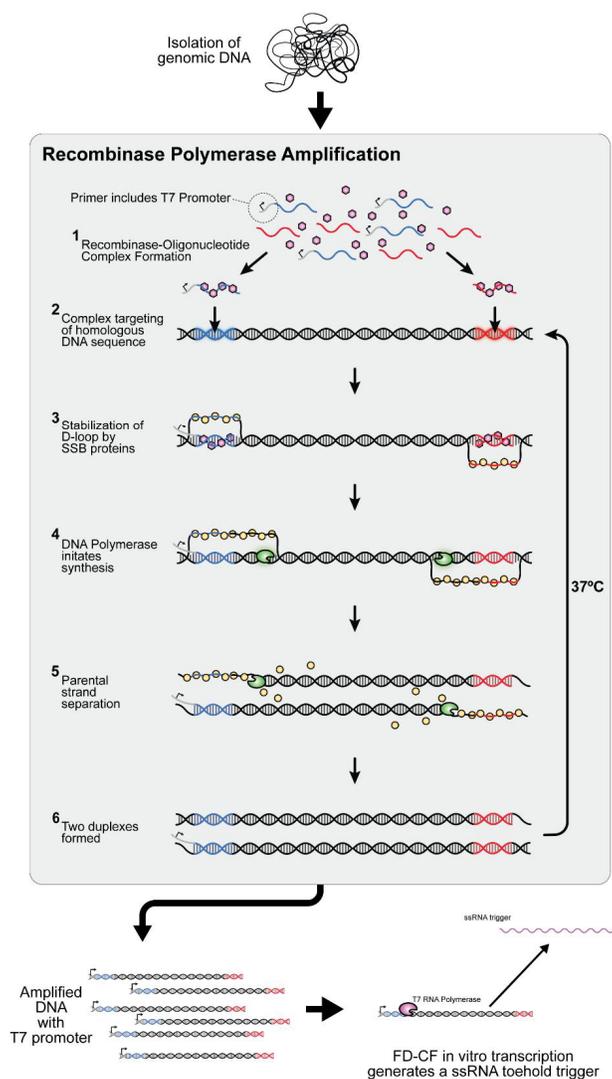
Figure 5.11 Schematic of RPA reaction. From a genomic DNA sample, a specific region is isothermally amplified using Recombinase Polymerase Amplification. The primer includes a T7 promoter, such that the amplicons act as a template to generate a large amount of RNA trigger molecules when added to a FD-CF reaction. This results in signal amplification for toehold sensor activation.

5.3.2.8   Statistical analysis

Statistical parameters including the definitions and values of n, SDs, and/or SEs are reported in the figures and corresponding figure legends.

5.3.3 Results

5.3.3.1   Fluorescent proteins as visual outputs

Our first goal was to develop a set of outputs that would engage as many of the five senses as possible to pique students' interest in the activities. As a visual output, we used FD-CF crude extract reactions to express fluorescent proteins. We selected a set of five fluorescent proteins[224,226,252-256] that cover a spectrum of colors, a subset from the BioBits™ Bright kit composed of red (eforRed), orange (dTomato), yellow (mOrange), green (sfGFP), and cyan (Aquamarine). FD-CF pellets, including DNA templates encoding the five proteins, were rehydrated and incubated overnight (20 hours) at 30°C. The fluorescent proteins expressed robustly and were easily visible to the eye even without fluorescent excitation. The fluorescent colors were also vivid when viewed using a custom low-cost, portable fluorescence illuminator we developed[248] (**Figure 5.12a**).
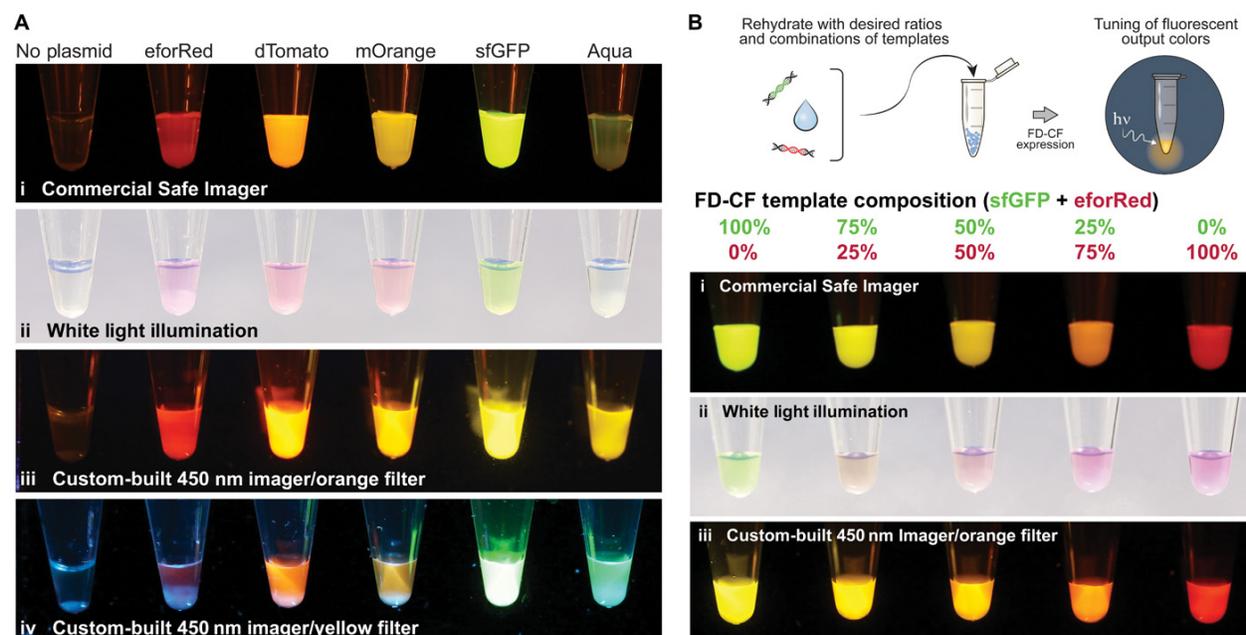
Figure 5.12 Fluorescent proteins as visual outputs. (**a**) A set of fluorescent proteins were expressed by FD-CF expression in crude extract and visualized with (i) a laboratory transilluminator (Safe Imager at 470-nm excitation), (ii) white light epi-illumination, (iii) a portable, inexpensive (<US$15) 450-nm classroom illuminator with an orange acrylic filter, or (iv) a yellow acrylic filter. (**b**) sfGFP and eforRed fluorescent proteins were expressed at a range of different combinations (by ratio of template DNA added) in FD-CF crude extract and visualized with (i) the Safe Imager, (ii) white light, and (iii) the classroom illuminator with the orange acrylic filter to demonstrate tunable protein expression.

As an example of how these fluorescent outputs can be used to teach advanced biological concepts, we created a demonstration designed to convey the concept of tuning gene expression, a key aspect of synthetic biology. This activity builds on an activity in the BioBits™ Bright kit, where protein expression was tuned by varying the input DNA concentration. Here, we used FD-CF crude extract reactions to coexpress two different fluorescent proteins, sfGFP and eforRed, simultaneously in a single reaction. The FD-CF pellets, containing different ratios of each DNA template, were rehydrated with water to achieve a range of intermediate colors from green to red that are visible to the eye under both white light and fluorescence (**Figure 5.12b**; **Figure 5.13**). This coexpression, which, to our knowledge, has not been carried out before in this FD-CF format, can be replicated with DNA templates for any other fluorescent protein pairs in the kit, providing students the freedom to choose the combination of visual outputs they would like to engineer.
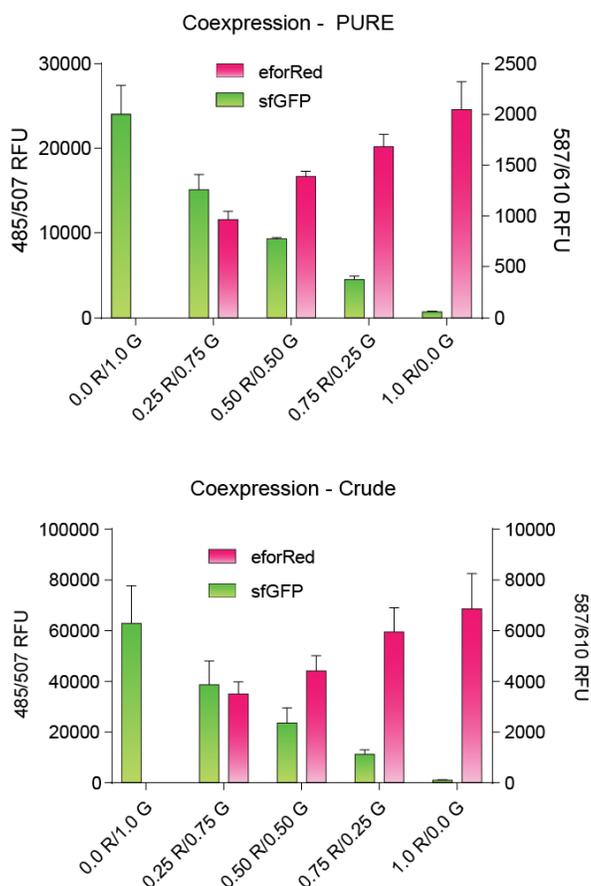
Figure 5.13 Quantitative analysis of fluorescent proteins. Endpoint fluorescent readouts of coexpressed sfGFP and eforRed proteins in the PURE or crude extract system. Values represent averages and error bars represent standard deviations of *n* = 3 biological replicates.

This activity also provides the opportunity to teach students the concept of the design-build-test cycle, a common paradigm used by synthetic biologists when developing new genetic circuits[257]. Once students choose the visual output they would like to engineer, they can design an experiment to mix fluorescent proteins in different ratios to achieve their goal. In this example, the build step would involve obtaining FD-CF pellets with the appropriate DNA concentrations. Students would then test their experimental design, evaluate the results, and iterate the process, as desired. The application of these fluorescent modules and educational demonstrations, paired with the inexpensive fluorescent imager that we developed, provides simple and cost-effective alternatives to traditional biology experiments, which are too expensive and complex to implement in an average classroom.

5.3.3.2   Fragrance-generating enzymes as olfactory outputs

Next, we sought to develop a synthetic biology circuit that would engage students' sense of smell. To achieve this, we expressed a single enzyme, alcohol acetyltransferase (ATF1), in FD-CF PURE and crude reactions overnight (20 hours) at 37°C. ATF1 is a key enzyme in aroma biochemistry[211] that converts isoamyl alcohol to isoamyl acetate, which imparts a strong banana fragrance (**Figure 5.14a**). We then mixed FD-CF reactions expressing ATF1 at a 1:10 dilution into a buffered reaction containing 25 mM isoamyl alcohol and 5 mM acetyl–coenzyme A (CoA). We allowed the enzymatic reactions to proceed 20 hours at room temperature, after which we were able to detect strong banana scents by smell.
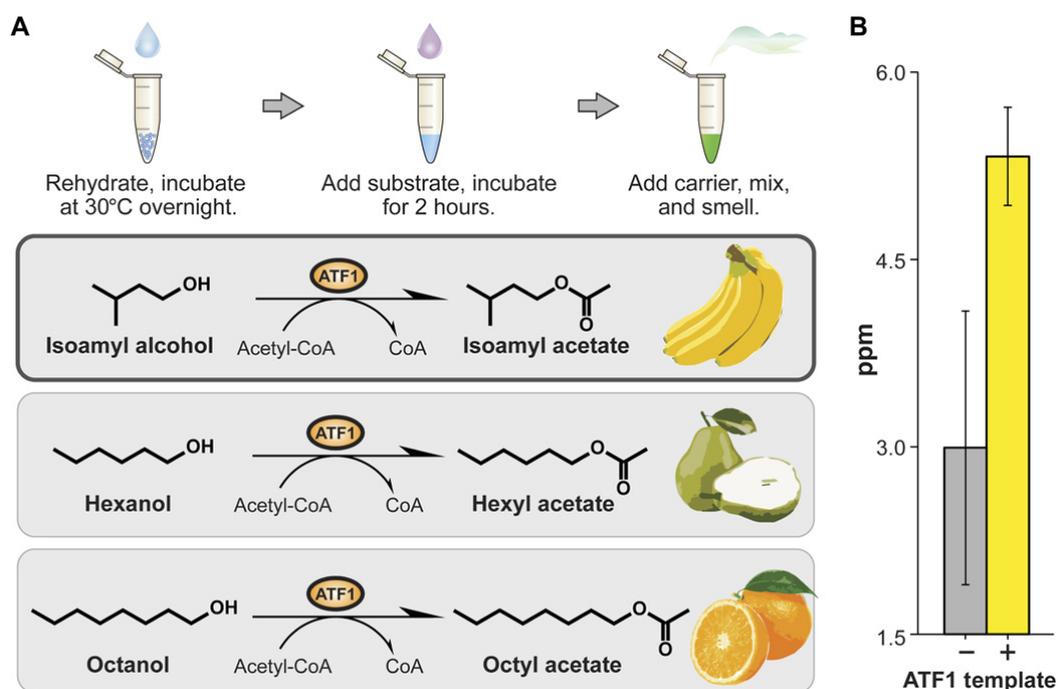
Figure 5.14 Fragrance-generating enzymes as olfactory outputs. (**a**) Using FD-CF reactions, we manufactured enzymes that can generate various smells from the *Saccharomyces cerevisiae* acetyltransferase ATF1. (**b**) Production of fragrance molecules after substrate addition to overnight FD-CF reactions of ATF1, as detected by headspace GC-MS. Values represent averages, and error bars represent SDs of *n* = 3 biological replicates.

To quantify the production of the volatile product, we set up identical FD-CF PURE reactions during which we collected the vapor phase of the reaction and analyzed it by gas chromatography–mass spectrometry (GC-MS). GC-MS analysis confirmed the presence of isoamyl acetate at ~5.3 parts per

million (ppm; volatile phase), which is well within the reported odor detection threshold for this compound (0.00075 to 366 ppm) but well below the permissible exposure limit of 100 ppm (**Figure 5.14b**)[258]. Thus, an average student can readily detect this FD-CF–generated aromatic in a classroom. ATF1 can convert various long-chain alcohol substrates to the corresponding acetylated esters, which have different fragrances[259]. Incubation of ATF1-expressing FD-CF reactions with the substrates hexanol and octanol, for example, could generate volatile products that encompass pear and citrus smells, respectively (**Figure 5.14a**).

To our knowledge, simple cell-free reactions that generate volatile fragrance molecules readily detectable by humans from an overnight incubation have not been previously developed. Production of these olfactory outputs can teach students about basic enzymatic reactions, provide a great connection to lessons learned in their chemistry classes, and inspire potential research projects for more advanced student groups. For example, here, we set up the enzymatic reactions containing isoamyl alcohol with and without FD-CF–produced ATF1 to show that the enzyme must be present to generate a smell. ATF1 could also be mixed with nonreactive substrates with different chemical functional groups to demonstrate that the enzyme only catalyzes a specific reaction. Moreover, these experiments can be put into a real-world context by noting that there are synthetic biology companies that work with enzymes in engineered microbes to produce fragrances and other commodity chemicals[260,261].

5.3.3.3   Hydrogel-generating enzymes as tactile outputs

Next, we sought to create a product using FD-CF reactions that students could interact with in a tactile manner. To do so, we drew inspiration from engineered hydrogel materials that have been developed for biomedical and biotechnological applications[262,263]. Like the olfactory outputs, hydrogels can be produced by enzymatic reactions (**Figure 5.15a**). Sortase is an enzyme that recognizes and covalently links specific peptide sequences (GGG and LPRT) through a transpeptidation reaction[264]. We expressed sortase in FD-CF crude and PURE reactions overnight at 37°C and used it to cross-link a solution of eight-arm polyethylene glycol (PEG) molecules modified with GGG or LPRT peptides (8%, w/v). We observed hydrogel formation within 30 min of incubation at 37°C (**Figure 5.15b, c**; **Figure 5.16**).

In contrast, PEG solutions incubated with FD-CF reactions that contained no template DNA did not exhibit a phase change and remained in liquid form.
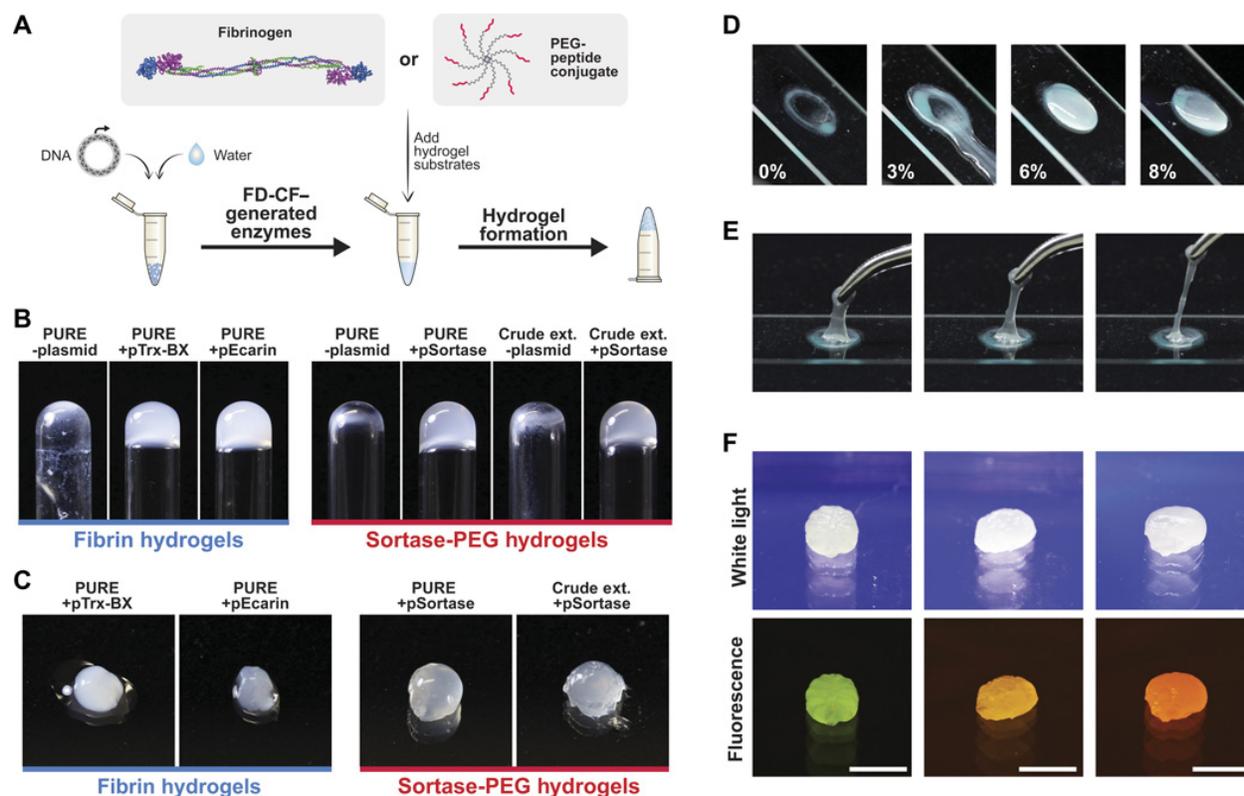
Figure 5.15 Hydrogel-generating enzymes as tactile outputs. (**a**) Schematic of fibrin hydrogels created from FD-CF–generated batroxobin/ecarin proteases that activate fibrinogen by cleavage or PEG-peptide hydrogels cross-linked by FD-CF–generated sortase enzymes that induce cross-linking by transpeptidase activity. (**b**) Inverted glass tubes to demonstrate formation of hydrogels. (**c**) Close-up images of the formed hydrogels that can be manipulated by hand. (**d**) Tuning the mechanical properties of the hydrogel by varying the % PEG to create a range of materials with varying viscosities. (**e**) An 8% crude FD-CF PEG hydrogel is highly elastic. (**f**) Casting the hydrogels into shapes using molds and mixing with crude FD-CF fluorescent protein reactions to obtain shaped fluorescent hydrogels. Scale bar, 1 cm.

In addition to the sortase-catalyzed hydrogel, we developed a method to use FD-CF PURE reactions to produce fibrin-based hydrogels. Fibrinogen is a glycoprotein found in blood that, when enzymatically converted to fibrin, leads to the formation of a blood clot[265]. Outside the context of blood, the proteases ecarin and batroxobin have been shown to cleave fibrinogen, which leads to the self-assembly of fibrin molecules into a hydrogel[266,267]. FD-CF reactions were used to produce ecarin and

batroxobin, which were added to resolubilized fibrinogen. We observed fibrin-based hydrogels after overnight incubation at room temperature, while the enzyme-free reaction remained unpolymerized (**Figure 5.15b, c**; **Figure 5.16**).
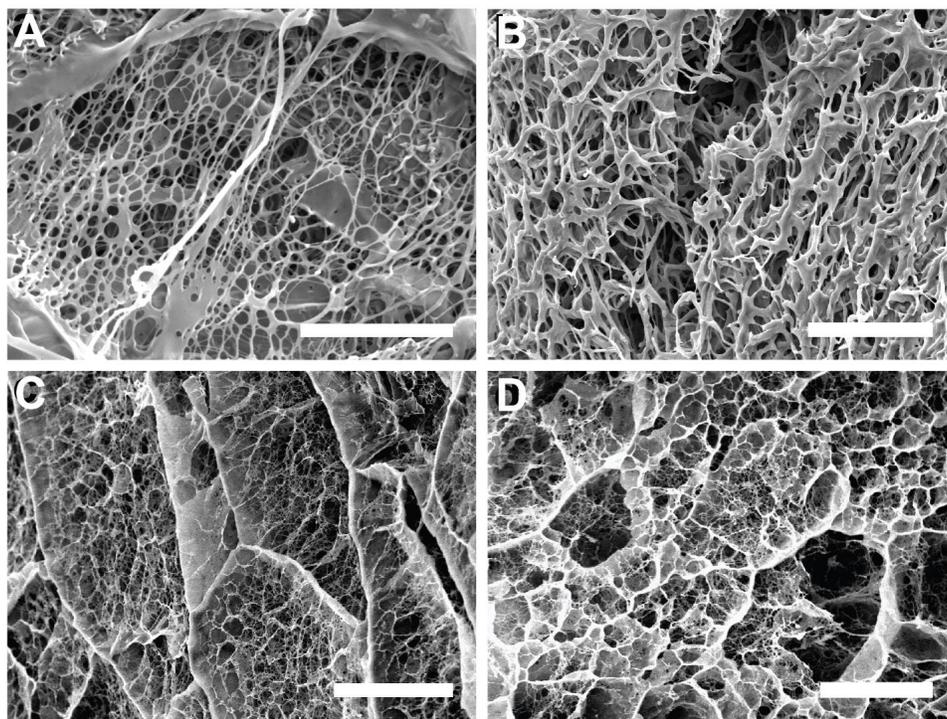
Figure 5.16 Representative scanning electron microscopy images of hydrogel ultrastructures generated with FD-CF enzymes. PEG hydrogels crosslinked by FD-CF expressed sortase in (**a**) PURE and (**b**) crude extract. Fibrin hydrogels crosslinked by FD-CF expressed (**c**) ecarin in PURE and (**d**) batroxobin in PURE. All scale bars are 10 microns.

The sortase and ecarin-mediated hydrogel formation are the first demonstrations, to our knowledge, of an engineered cell-free protein synthesis reaction that generates self-assembling macromolecular hydrogels from genetically encoded components. Once formed, one can manipulate the hydrogels by hand, allowing students to experience another enzyme-catalyzed biochemical reaction, this time resulting in an output that they can feel. These basic demonstrations can be used to introduce advanced biological concepts such as blood clotting or how cells use similar processes to form human tissue. As a follow-up project, students can experiment with the notion of engineering gene expression to

create biomaterials with tunable properties, a current aim of synthetic biology research[268]. Specifically,

the mechanical properties of the hydrogel can be tuned by varying the concentration of the substrates to

create a range of materials from a viscous slime to a stiff hydrogel (**Figure 5.15d, e**). The hydrogels can

also be cast into shapes using molds and/or combined with the fluorescent protein outputs to create

fluorescent hydrogels (**Figure 5.15f**).

5.3.3.4   Probing the environment using designer biosensors

While the first three components provide students with an FD-CF toolkit of sensory outputs from

simple DNA inputs, we wanted to expand the BioBits™ Explorer kit to inspire a proactive, inquisitive

mindset in students and provide them with the means to interrogate the world around them. In our final

demonstration, we develop FD-CF–based tools that allow students to probe real-world biological samples

using toehold switch sensors. Here, we expand on the common classroom activity of isolating DNA from

fruits to allow students to detect DNA signatures of a specific fruit and couple it to a fluorescent output.

Toehold switch sensors are programmable synthetic riboregulators that allow protein expression

only when a specific trigger RNA is present[51,269]. These sensors consist of an mRNA molecule designed

to include a hairpin structure that blocks gene translation in cis by sequestration of the ribosome binding

site and start codon. Hybridization to a complementary trigger RNA results in secondary structure

rearrangement, facilitating ribosomal translation of an output gene (**Figure 5.17a**). This technology allows

for regulatory control of the various sensory outputs described previously (fluorescence, fragrance, and

hydrogels) to be conditionally dependent on the presence of a specific nucleic acid molecule.
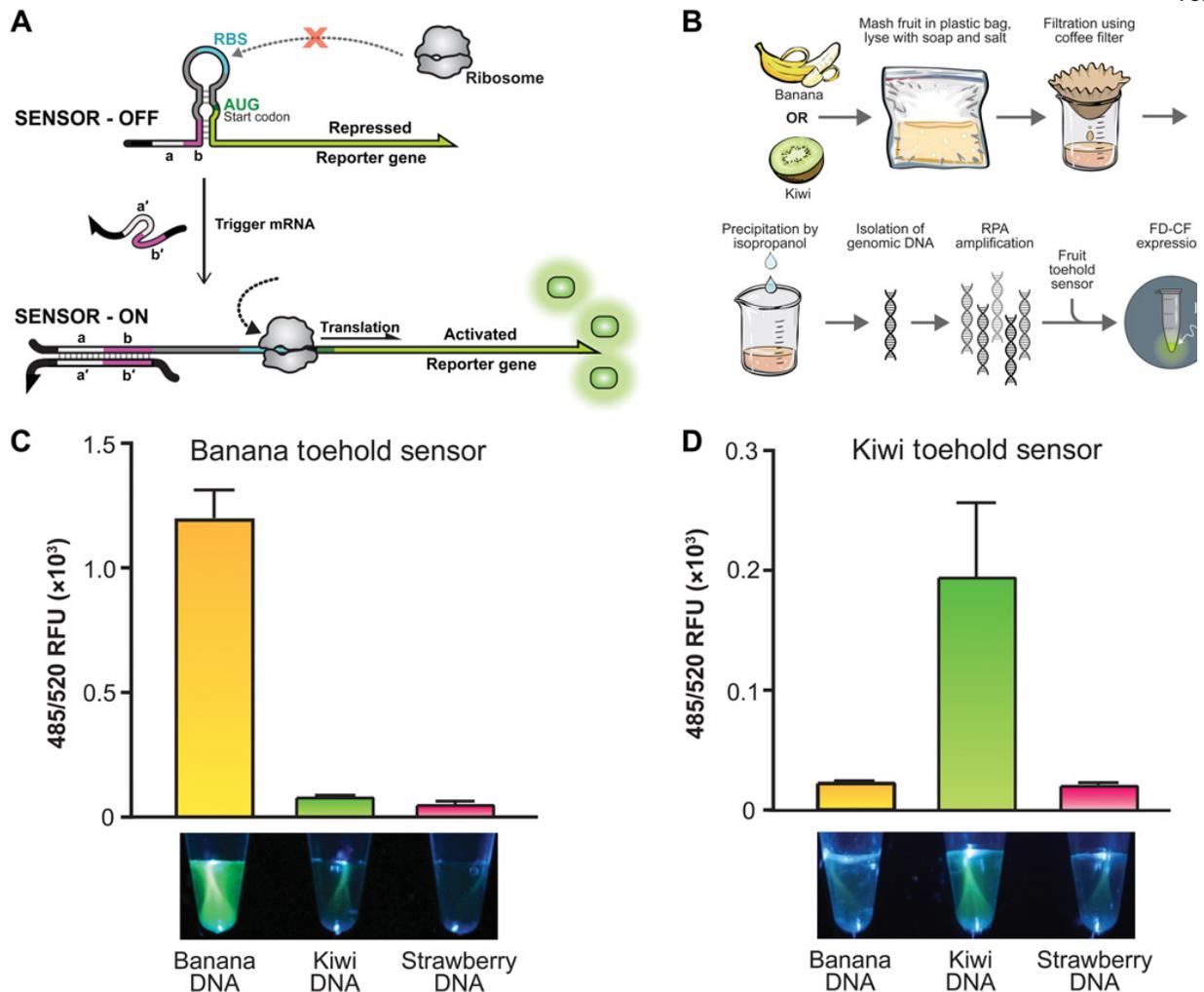
Figure 5.17 T**oehold-based environmental sensing demonstrations.** (**a**) Schematic of a toehold switch sensor. Upon the presence of a trigger RNA, strand invasion melts the secondary structure, allowing ribosomal translation to occur. (**b**) Schematic of activity that allows extracted DNA from banana or kiwi fruit to be processed and detected by a toehold switch sensor in FD-CF. (**c**) The banana toehold switch sensor or (**d**) the kiwi toehold switch sensor produces a clear fluorescence output (sfGFP) when exposed to extracted and amplified DNA of the relevant fruit but not when exposed to DNA sequences from other fruits. Images shown are from a custom-built 450-nm handheld imager with a yellow acrylic filter and quantified by a plate reader at 485-nm excitation and 520-nm emission. Values represent averages, and error bars represent SDs of n = 3 biological replicates.

For this demonstration, we designed the first toehold switch sensors that are able to discriminate between plant species. These toehold sensors are activated by sequences from either the banana genome (*rbcL* gene from *Musa acuminata* or *Musa balbisiana*) or the kiwi genome (5.8*S* ribosomal RNA gene from *Actinidia deliciosa*) and produce sfGFP as an output[270,271]. Since the toehold switch sensors recognize RNA and not DNA, we included a recombinase polymerase amplification (RPA)[272] step to

generate short DNA amplicons from fruit genomic DNA that incorporate a T7 promoter for transcription of RNA triggers in FD-CF reactions (**Figure 5.11**). The RPA reaction components can also be freeze-dried, conforming to the shelf-stable aspect of these kits. To demonstrate the specificity of the toehold switch sensors, we isolated DNA from banana, kiwi, or strawberry (**Figure 5.17b**; **Figure 5.18**), using a simple procedure often performed in classrooms that utilizes common inexpensive supplies such as dish soap, table salt, rubbing alcohol, and coffee filters[273]. Diluted DNA from each of the fruits was used to rehydrate RPA reactions and incubated overnight at 37°C. Completed RPA reactions were further diluted 1:4 in water and used to rehydrate FD-CF PURE reaction pellets containing the banana or kiwi toehold switch sensor. Upon overnight incubation at 37°C, activation of each toehold switch sensor was only observed from the reaction containing the specific fruit DNA trigger and not from the reactions containing other fruit DNA (**Figure 5.17c, d**).
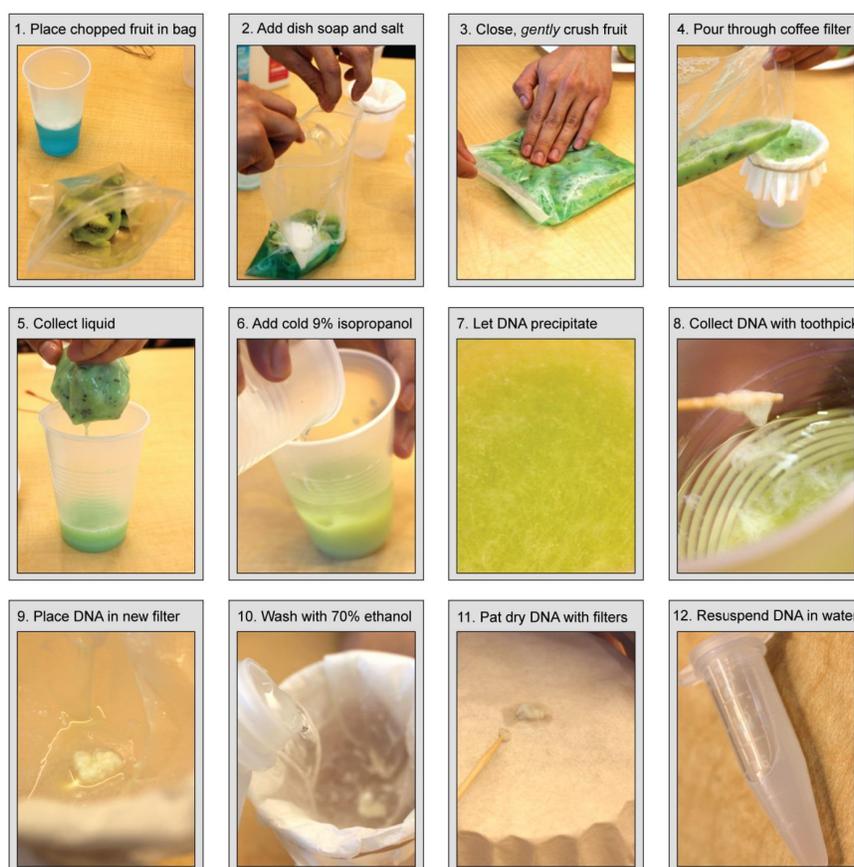
Figure 5.18 Detailed steps for isolating genomic DNA from fruits for environmental sensing activity. Photographs showing the DNA extraction process from fruit.

The biosensor module described here uses the detection capability of toehold switches to provide a hands-on introduction to the concept of nucleic acid–based diagnostics[51,220]. This module could be expanded readily to include additional toehold switch sensors designed to detect other fruits and vegetables. Sensors could be designed similarly to identify specific animal species (for example, cats and dogs), enabling students to test DNA from their pets.

5.3.4 Discussion

The next-generation synthetic biology educational kit described here addresses the need for easy-to-implement, hands-on biology demonstrations in STEM education. We used shelf-stable FD-CF reactions to bring molecular and synthetic biology experiments into a classroom setting in an affordable manner, without the need for specialized equipment or refrigeration. First, we developed a set of

genetically encoded outputs that engages the senses of sight, smell, and touch via fluorescent proteins, enzyme-generated scents, and enzyme-generated hydrogels, respectively. These outputs provide demonstrations that can be used to teach fundamental biology concepts and principles of tunable protein expression, enzyme catalysis, and material properties. Although we focus on the modules as educational demonstrations, it should be noted that two of the engineered freeze-dried synthetic biology circuits presented here, genetically encoded macroscopic hydrogellation and olfactory detection, represent the first demonstration of tactile or fragrance outputs in cell-free systems. These modules expand the repertoire of sensory outputs available for cell-free biosensors beyond visual fluorescent outputs. For the BioBits™ Explorer kit, we also created a module that allows students to take DNA extracted from fruits and analyze the samples using toehold switch sensors designed to detect specific DNA sequences from the banana and kiwi genome. The isolation of DNA from fruits is a widely used classroom activity due to its simple protocol and ease of implementation[273]. For the first time, using these toehold switches, students can actually go further and probe the DNA on the genetic level. Beyond educational demonstrations, biosensors for plant tissue discrimination such as those described here could also be used practically in agriculture to detect contamination in food crops. These toehold sensors also open up the possibility of students being able to design their own custom biosensors to probe living organisms and couple that to a wide variety of outputs.

The demonstrations presented for these toolkits were designed to be modular: Teachers can incorporate these explorations into their curricula as they see fit, depending on the content they wish to teach and classroom time available. We thus envision BioBits™ Explorer being further developed into a diverse array of kits to accommodate different grade levels and budgets, although the modules presented here are all affordable. The basic Explorer kit could include simple demonstrations involving the three sensory outputs. The kit would also include the low-cost incubator and portable fluorescent imager we developed[248]. Each FD-CF reaction costs approximately $0.15. Thus, reagents and other supplies for a 30-student classroom for the basic kit would only cost about $200 to produce (**Table 5.6**)—much less than the cost of materials and traditional equipment (shaking incubators, refrigerators, thermocyclers, etc.) needed for in vivo biological experiments. More advanced Explorer kits could also include activities

such as the toehold biosensors; the additional reagents would only add on a cost of about $200 (**Table 5.6**). We anticipate that costs associated with these kits would be further reduced as manufacturing methods are optimized and economies of scale are leveraged.

We note that some of the FD-CF reactions were carried out with the more expensive PURE system when expression was low in the crude extracts. The fruit DNA biosensors, for example, were implemented in the PURE system because of inherent autofluorescence from the crude extract that hindered visualization. In the future, the toehold switches could be optimized to increase their output expression, which would allow the use of the inexpensive crude extract. In addition, crude extracts could be optimized to improve the expression of specific enzymes by using different bacterial strains. Recent improvements to the efficiency of the reconstituted PURE production system also suggest that its cost could be reduced to that of the crude system[274,275]. These optimizations would reduce the cost of the BioBits™ kits even further.

We are currently developing a companion website to facilitate the formation of an open-source community around the BioBits™ kits. This online community would provide users with a forum that would facilitate discussion and development of new ideas for lessons and demonstrations using the provided components. We also plan to add a software component that would allow students to design their own sensors (specific to other environmental samples) and other synthetic biology components and request the DNA for them online. In this way, the open-source community could design, build, and test additional genetic constructs to add to the BioBits™ parts library.

The activities demonstrated here engage students by appealing to their senses through diverse genetic outputs using simple just-add-water FD-CF pellets. These illustrative demonstrations can be used to introduce a wide range of molecular and synthetic biology concepts in classrooms. Our kits contain just a few examples of the potential activities that can be developed using FD-CF reactions; by mining the available library of synthetic biology parts and developing novel genetic circuit combinations, a plethora of additional modules could be created to teach advanced biology concepts. Together, our BioBits™ Bright and Explorer kits provide a new paradigm for bringing affordable life sciences and biotechnology experiments into any classroom, making quality biology education accessible to all students.

5.4   BioBits^{TM} Health: Classroom activities exploring engineering, biology, and human health with fluorescent readouts

5.4.1 Introduction

Synthetic biological technologies promise to enable paradigm-shifting advances in human health and disease. To date, the field of synthetic biology has made meaningful progress toward biomanufacturing of antimicrobials and other medicines[134,200,201], cellular diagnostics and therapies[205-208], and human gene editing technologies[276-279]. These potentially transformative technologies offer rich opportunities for hands-on biology education, as they require students to confront real-world problems at the intersection of diverse disciplines, including biology, chemistry, engineering, math, design, policy, and ethics. Such cross-cutting educational activities align closely with the objectives of K-12 STEM (Science, Technology, Engineering, and Mathematics) education and priorities identified by the National Academy of Engineering to enable students to apply, adapt, and connect fundamental principles across multiple disciplines[209]. Moreover, educating students about the science and ethics associated with biological technologies will empower them to make informed policy decisions in the future about how such technologies should be used and regulated.

However, due to the cutting-edge nature of these technologies, their translation into educational activities has been limited. In fact, hands-on resources for teaching molecular and synthetic biology have been limited in general, despite evidence that hands-on science activities have been shown to improve student understanding and academic performance[63]. Synthetic biology-based educational efforts such as BioBuilder Educational Foundation, the International Genetically Engineered Machines competition, Amino Labs, and The ODIN have made great strides toward integration of cutting-edge, hands-on biology research into classrooms. However, these resources rely on cell-based experimentation, which requires (i) expensive equipment and specialized expertise to grow and engineer cells[64,65], (ii) extended instructor prep time and in-class time due to the time scales associated with cell growth[66], and (iii) compliance with biosafety regulations that can limit the ability to work with cells outside of a laboratory setting[67,68].

To address these issues, we previously reported adaptation of synthetic biology lab activities into two portable, low-cost, and user-friendly educational kits: BioBits^{TM} Bright[248] and BioBits^{TM} Explorer[66].

These kits are made possible through the use of freeze-dried cell-free (FD-CF) technology, which harnesses an ensemble of catalytic components (e.g., RNA polymerases, ribosomes, aminoacyl-tRNA synthetases, translation initiation and elongation factors, etc.) from cell lysates to synthesize proteins *in vitro*[39,44]. FD-CF reactions circumvent many of the biosafety and biocontainment regulations that exist for living cells because they use cell lysates rather than intact cells to synthesize proteins. Further, FD-CF reactions eliminate the need for specialized equipment or experimental expertise as they are shelf-stable[50] and can be activated simply by adding water and other desired inputs (e.g., DNA, small molecules, and enzymes) to a freeze-dried pellet of reagents. In addition, we showed that FD-CF reactions are robust teaching tools, evidenced by the fact that K-12 students and teachers running FD-CF reactions for the first time were able to obtain the intended experimental results[248]. Together, these features make BioBits™ kits a welcome complement to existing educational kits for classrooms or other nonlaboratory settings. Despite these developments, activities that capture the recent advances and impacts of biological technologies on human health are still limited. If FD-CF technology could be used to develop educational modules about these technologies, it could significantly lower the barrier to entry for teaching emerging health-related topics.

Here, we describe BioBits™ Health, an educational kit that links complex biological experiments to fluorescent readouts in easy-to-use FD-CF reactions (**Figure 5.19**). We and others have demonstrated that fluorescent or chromoproteins are ideal instructional tools because a wide variety of these proteins have been developed which produce colors and/or fluorescence visible to the naked eye. These visual outputs make possible easy qualitative or semiquantitative data collection without the need for expensive analytical equipment[248,280]. To enable BioBits™ Health, we developed two educational modules designed for high school classrooms with wet lab activities that investigate antibiotic resistance and CRISPR-Cas9 gene editing technology. These laboratories use FD-CF reactions with just a few simple inputs (water, DNA, small molecules, and enzymes). For both modules, we show that the fluorescent results from the various lab activities can be analyzed qualitatively by eye using a low-cost blue light imager[248], demonstrating the ability to run and assay reactions without sophisticated laboratory equipment. demonstrating the ability to run and assay reactions without sophisticated laboratory equipment. Each lab

activity can be set up in a single 1 hour class period and produces results that can be analyzed as soon as the following day. The simplified nature of reaction setup and analysis minimizes both the amount of in-class time and out-of-class instructor preparation time required to incorporate hands-on lab activities, which have been cited as limiting factors for high school biology teachers[69,70]. In addition, we show that the laboratories can be run successfully by Chicago high school students and teachers, representing our target audience. Though the primary goal of this work was to develop and validate the lab activities, we also worked with Chicago high school teachers to develop a set of curricula and prelab lecture slides (**Curricula 6-16**) to provide an example framework for teaching using the BioBits™ Health kit. Overall, the BioBits™ Health kit uniquely offers hands-on, cross-cutting educational activities with example supporting curricula that convey some of the most recent health-related advancements in synthetic biology in an interactive way. These resources promise to encourage hands-on education at the high school level about biological technologies for treating and understanding human health and disease.
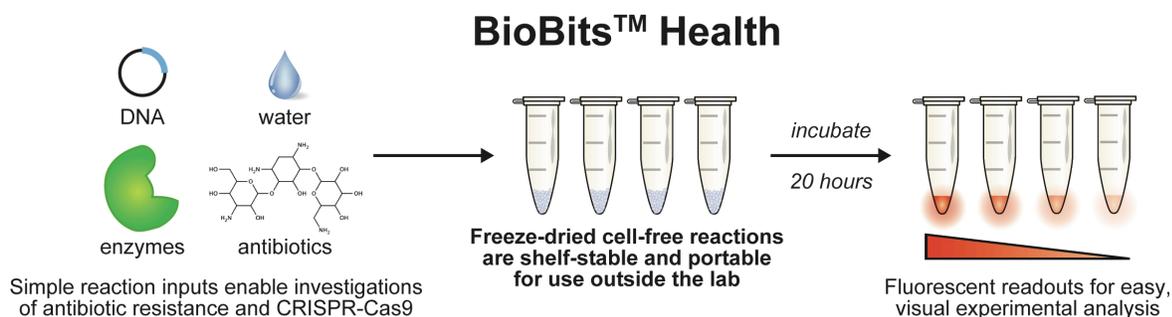
Figure 5.19 BioBits™ Heath is a set of classroom activities and curricula that links cutting-edge, health-related biology experiments to visual, fluorescent readouts. Using freeze-dried cell-free (FD-CF) reactions with simple inputs (DNA, antibiotics, enzymes, and water), BioBits™ Health lab activities enable hands-on, inquiry-guided educational activities focused on antibiotic resistance and CRISPR-Cas9 genome engineering. FD-CF reactions are shelf-stable and can be run and analyzed without expensive equipment, making them well-suited for use in classrooms or other non-laboratory settings.

5.4.2 Methods

5.4.2.1   Bacterial strains and plasmids

E. coli NEB® 5-alpha (NEB) was used in plasmid cloning transformations and for plasmid preparation. Wild-type or *rpsL* R86S *E. coli* BL21 Star™ (DE3) cells (Thermo Fisher Scientific) were used

for preparation of cell-free extracts. Gibson assembly was used for seamless construction of plasmids used in this study (**Table 5.1**). For cloning, the pJL1 vector (Addgene 69496) was digested using restriction enzymes NdeI and SalI-HF® (NEB). Each gene was amplified via PCR using Phusion® High-Fidelity DNA polymerase (NEB) with forward and reverse primers designed with the NEBuilder® Assembly Tool (nebuilder.neb.com) and purchased from IDT. The DNA construct encoding the anti-mRFP1 gRNA was also purchased from IDT. PCR products were gel extracted using an EZNA Gel Extraction Kit (Omega Bio-Tek), mixed with Gibson assembly reagents and incubated at 50°C for 1 hour. Plasmid DNA from the Gibson assembly reactions were transformed into *E. coli* NEB 5-alpha cells and circularized constructs were selected on LB-agar supplemented with kanamycin at 50 µg mL^{-1} (Sigma). Sequence-verified clones were purified using an EZNA Plasmid Midi Kit (Omega Bio-Tek) for use in FD-CF reactions.

5.4.2.2   Construction of *rpsL* R86S mutant strain

The *rpsL* R86S strain of BL21 Star™ (DE3) was generated using a mutagenic oligonucleotide that was designed to introduce the C256A mutation into the *rpsL* gene via a single-cycle of MAGE[281]. Briefly, BL21 Star™ (DE3) cells were transformed with the pKD46 plasmid encoding the lambda Red recombinase system[95]. Transformants were grown to an OD$_{600}$ of 0.5-0.7 in 200 mL LB-Lennox media (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract and 5 g L^{-1} NaCl) with 50 µg mL^{-1} carbenicillin at 30°C, harvested and washed three times with 40 mL ice-cold 10% glycerol to make them electrocompetent, and resuspended in a final volume of 200 µL 10% glycerol. Electrocompetent cells were combined with the *rpsL* MAGE oligonucleotide for a final concentration of 37.5 mM DNA, transformed via electroporation, and plated on LB-agar with 50 µg mL^{-1} streptomycin for selection of resistant colonies. Plates were grown at 37°C to cure cells of the pKD46 plasmid. Colonies that grew on streptomycin were confirmed to have the desired mutation via colony PCR and DNA sequencing and to have lost the pKD46 plasmid via replica plating on LB agar plates with carbenicillin and streptomycin. All primers used for construction and validation of this strain are listed in **Table 5.7**.

5.4.2.3   CFPS extract preparation

CFPS extract was prepared by sonication as previously reported[102]. Briefly, *E. coli* BL21 Star™ (DE3) (Thermo Fisher Scientific) cells were grown in 2xYTPG media (5 g L^{-1} NaCl, 16 g L^{-1} tryptone, 10 g L^{-1} yeast extract, 7 g L^{-1} potassium phosphate dibasic, 3 g L^{-1} potassium phosphate monobasic, 18 g L^{-1} glucose) pH 7.2 at 37°C. BL21 Star™ (DE3) *rpsL* R86S cells were grown in 2xYTPG media supplemented with 50 µg mL^{-1} streptomycin. T7 polymerase expression was induced at an OD_{600} of 0.6-0.8 with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). Cells were grown at 37°C to a final OD_{600} of 3.0, at which point cells were pelleted by centrifugation at 5,000xg for 15 min at 4°C. Cell pellets were then washed three times with cold S30 buffer (10 mM Tris-acetate pH 8.2, 14 mM magnesium acetate, and 60 mM potassium acetate) and pelleted at 5,000xg for 10 min at 4°C. After the final wash, cells were pelleted at 7,000xg for 10 min at 4°C, weighed, flash frozen in liquid nitrogen, and stored at -80°C. For lysis, cell pellets were suspended in 1 mL of S30 buffer per 1 g of wet cell mass. The cell-buffer suspension was transferred into 1.5 mL microcentrifuge tube and placed in an ice-water bath to minimize heat damage during sonication. The cells were lysed using a Q125 Sonicator (Qsonica) with 3.175 mm diameter probe at 20 kHz and 50% amplitude. The input energy was monitored and 640 J was used to lyse 1 mL of suspended cells. The lysate was then centrifuged once at 12,000xg at 4°C for 10 min. Cell extract was aliquoted, flash-frozen on liquid nitrogen, and stored at -80°C. Alternatively, for classroom settings where it is not practical to generate or obtain FD-CF reactions, similar cell-free systems are available commercially from companies such as Promega® (L1130).

5.4.2.4   Cell-free protein synthesis (CFPS)

FD-CF reactions were carried out in PCR tubes (5 µL reactions) or 1.5 mL microcentrifuge tubes (15 µL reactions). The CFPS reaction mixture consists of the following components: 1.2 mM ATP; 0.85 mM each of GTP, UTP, and CTP; 34.0 µg mL^{-1} L-5-formyl-5, 6, 7, 8-tetrahydrofolic acid (folinic acid); 170.0 µg mL^{-1} of *E. coli* tRNA mixture; 130 mM potassium glutamate; 10 mM ammonium glutamate; 8 mM magnesium glutamate; 2 mM each of the 20 amino acids; 0.4 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); 57 mM HEPES; 13.3 µg mL^{-1} plasmid; and 27%

v/v of cell extract[101]. If pre-expressed antibiotic resistance enzymes or Cas9 nuclease were also added, 1 µL of the soluble fraction of CFPS reactions encoding these enzymes was used along with water and plasmid to rehydrate fresh FD-CF reactions. For CRISPR reactions, reactions contained all of the ingredients above, except with 2 µg mL^{-1} target plasmid (mRFP1 or other fluorescent protein constructs). If anti-mRFP1 gRNA plasmid was also added, it was supplied at 6.66 µg mL^{-1}. For quantification of fluorescent protein yields via radioactive leucine incorporation, 10 µM of L-^{14}C-leucine (11.1 GBq mmol^{-1}, PerkinElmer) was added to the CFPS mixture. All reagents required to make CFPS extracts and assemble cell-free reactions are listed in **Table 5.8**.

5.4.2.5   Lyophilization of cell-free reactions

FD-CF reactions were prepared according to the recipe above, but without plasmid added. CFPS reactions and plasmids were separately lyophilized using a VirTis BenchTop Pro lyophilizer (SP Scientific) at 100 mTorr and -80°C overnight or until fully freeze-dried. Following lyophilization, plasmids were rehydrated with nuclease-free water (Ambion) and added to FD-CF reaction pellets at a final concentration of 13.3 µg mL^{-1} unless otherwise noted. CFPS reactions were carried out at 30°C for 20 hours after rehydration unless otherwise noted. In a classroom setting, reactions can be incubated in the BioBits^{TM} portable incubator[248] at 30°C or in a 30°C water bath in an insulated container (Styrofoam, plastic cooler, etc.) for 20 hours or at room temperature for 24-48 hours.

5.4.2.6   Quantification of *in vitro* synthesized protein

Active-full length protein synthesis was measured continuously via fluorescence using a CFX96 Touch Real-Time PCR Detection System (BioRad). If fluorescence saturated the RT-PCR detector, endpoint fluorescence was measured in 96 well half-area black plates (CoStar 3694; Corning Incorporated) using a Synergy2 plate reader (BioTek). Excitation and emission wavelengths used to measure fluorescence of each protein construct were as follows: mCherry, eforRed, mRFP1, dTomato ex 560-590 nm, em 610-650 nm; mOrange ex 515-535 nm, em 560-580 nm; YPet, sfGFP, mTFP1, CyPet, Aquamarine, mTagBFP2, mKalama1, eBFP2 ex 450-490, em 510-530 nm. Fluorescence units were converted to concentrations using a standard curve as previously described[138,248]. For assessing yields of the antibiotic resistance enzymes and the Cas9 nuclease, reaction soluble fractions were analyzed

directly by incorporation of ^{14}C-leucine into trichloroacetic acid (TCA)-precipitable radioactivity using a liquid scintillation counter as described previously[111]. The soluble fractions were also run on a Coomassie-stained SDS-PAGE gel and exposed by autoradiography. Autoradiographs were imaged with a Typhoon 7000 (GE Healthcare Life Sciences).

For quantification without a spectrophotometer, reactions can be semi-quantitatively analyzed via imaging using the BioBits™ 8-well blue light imager[248] and subsequent fluorescence analysis in ImageJ, a free image processing program, using our step-by-step ImageJ tutorial (**Curriculum 6**). Images of FD-CF reactions were taken with a DSLR and arranged in Adobe Illustrator. Protein production can also be qualitatively assessed with the naked eye under white light or blue or black light using our portable blue light imager or others (*e.g.,* Bio-Rad® UV pen lights #1660530EDU, Walmart® black light bulb with fixture #552707607, Home Science Tools® portable UV black light #OP-BLKLITE, miniPCR blueBox™ transilluminator #QP-1700-01).

5.4.2.7   Statistical analysis

Statistical parameters and analytical techniques including the definitions and values of *n*, standard deviations, and/or average errors are reported in the figures and corresponding figure legends.

5.4.3 Results

5.4.3.1   Module I: Investigating mechanisms of antibiotic resistance

Antibiotic resistance is a pressing global issue, projected to threaten up to 10 million lives per year by 2050[124]. To help address this growing worldwide problem, it is important to educate students about how resistance occurs. To meet this educational need, we developed a classroom module that aims to teach (i) antibiotic mechanisms of action, (ii) mechanisms by which pathogenic bacteria can develop resistance, and (iii) how human behaviors can accelerate the development of resistance. As early as the 1960s, cell-free systems have been used to interrogate the mechanisms of action of antibiotics that inhibit the ribosome[282,283]. In this module, we show that it is possible to use synthesis of the orange fluorescent protein dTomato in FD-CF reactions as a reporter of antibiotic efficacy (**Figure 5.20a**). In the first activity, students use a panel of six antibiotic ribosome inhibitors, including large (50S) and small (30S) subunit inhibitors, to inhibit protein expression in FD-CF reactions (**Figure 5.20b**).

Antibiotic-mediated inhibition of protein synthesis is visible by eye when reactions are viewed using

the BioBits™ portable blue light imager [248] (**Figure 5.20c**). This activity is analogous to running a

minimum inhibitory concentration assay, a common technique used to determine the potency of

antibiotics against pathogenic bacteria [284], but without the complications and biosafety concerns of using

live pathogenic cells. This first lab can be used as part of an inquiry-based classroom activity in which

students are tasked with identifying the mechanisms of action for various classes of antibiotics (*e.g.*,

Curricula 7, 8). Students can investigate ribosome inhibitors, which will inhibit protein synthesis, as well

as other antibiotic classes, such as cell wall biosynthesis inhibitors, which have no effect on cell-free

protein synthesis (**Figure 5.20b**). To successfully identify the mechanism of action of the latter class,

students will have to demonstrate understanding of cellular architecture and which cellular components

are present in the cell-free system (**Curriculum 8**). This particular curriculum example further offers the

opportunity to introduce students to image analysis and biological statistics using ImageJ, facilitated by

our step-by-step ImageJ tutorial for quantifying fluorescence in cell-free reactions (**Curriculum 6**). These

represent important analytical skills for students interested in pursuing careers in biology-related fields.
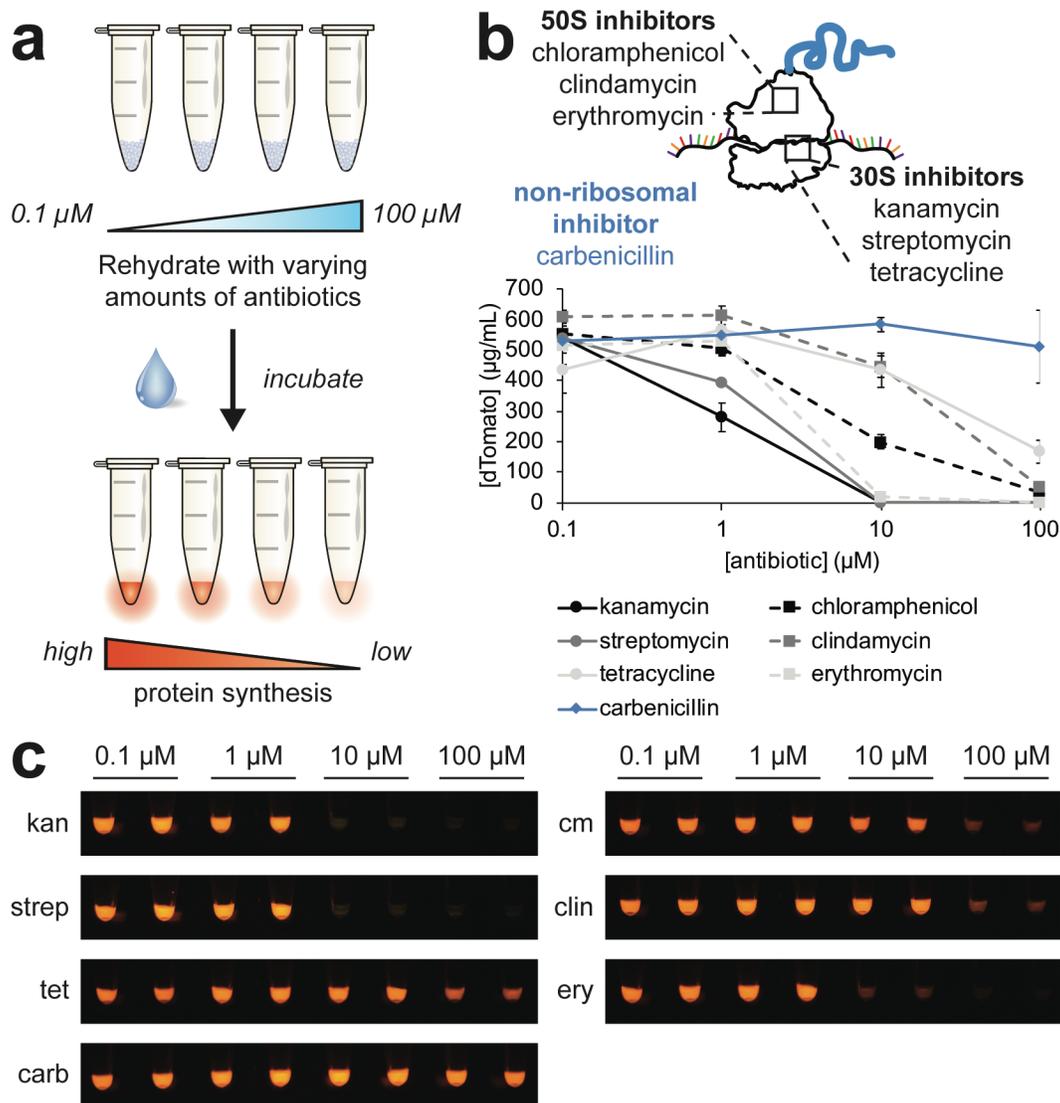
Figure 5.20 Fluorescent protein expression in FD-CF reactions can be used to assay antibiotic potency for a variety of antibiotic ribosome inhibitors. (**a**) Through addition of varying amounts of antibiotic ribosome inhibitors to FD-CF reactions expressing a fluorescent protein, fluorescence can be used as a reporter of antibiotic efficacy. (**b**) Cell-free protein synthesis of dTomato was carried out in FD-CF reactions containing 0.1–100 μM antibiotic. Antibiotics tested included the 50S inhibitors chloramphenicol (cm), clindamycin (clin), and erythromycin (ery), the 30S inhibitors kanamycin (kan), streptomycin (strep), and tetracycline (tet), and the cell wall biosynthesis inhibitor carbenicillin (carb) (**top**). For all antibiotics tested, except for carbenicillin (carb), protein synthesis was suppressed with increasing levels of antibiotics (**bottom**). Values represent averages, and error bars represent standard deviations of $n \geq 3$ biological replicates. (**c**) When representative FD-CF reactions from part (**b**) are imaged using a low-cost blue light imager, inhibition of protein synthesis can be observed by eye.

If class time allows, this activity can be extended with an additional lab that guides students

through an investigation of two potential mechanisms of antibiotic resistance: horizontal gene transfer and

genetic mutation. To mimic horizontal gene transfer (**Figure 5.21a**), we show that FD-CF reactions can be used to express the aminoglycoside O-phosphotransferase (NeoR) and streptomycin 3'-adenylytransferase (AadA) enzymes that confer kanamycin and streptomycin resistance, respectively (**Figure 5.22**). When added to fresh FD-CF reactions expressing dTomato, we observe that these pre-expressed enzymes can "rescue" the FD-CF reactions from the effects of their target antibiotic (**Figure 5.21b**). Further, the resistance conferred through the addition of NeoR or AadA is visible by eye compared to reactions lacking the enzymes using our portable blue light imager (**Figure 5.21c**). Similarly, to explore the idea of genetic mutation (**Figure 5.21d**), we made lysate from BL21 Star™ (DE3) cells in which we introduced an R86S mutation in the *rpsL* gene, which codes for ribosomal protein S12. This mutation has been previously reported to result in resistance to streptomycin, likely by preventing the antibiotic from binding to the ribosome [285]. When the mutant cell extract is used in FD-CF reactions, dTomato expression is observed even in the highest concentrations of streptomycin tested (**Figure 5.21e**), and the level of expression can again be easily distinguished from reactions with wild-type lysate using the BioBits™ imager (**Figure 5.21f**). We also show that similar levels of protein expression are achieved in both wild-type and resistant reactions assembled by high school students and teachers (**Figure 5.23a**). We developed a protocol with pre- and post-lab questions for implementing this lab in high school classrooms (**Curricula 7, 9**) with the option for students to quantify their results with ImageJ (**Curriculum 6**) and carry out a statistical analysis of their results. In sum, this module offers rich, inquiry-guided educational experiences that can be used to meet educational standards for high school biology (**Table 5.9**) by exploring antibiotic mechanisms of action, demonstrating multiple ways in which antibiotic resistance can be acquired, and offering opportunities to expose students to biological statistics and real-world data analysis methods. Further, understanding mechanisms of antibiotic resistance opens the door to classroom discussions or independent research projects that could help students make informed choices about the use of antimicrobial consumer products in their own lives, as well as broader policy decisions about antibiotic use and misuse (**Curricula 7, 9**).
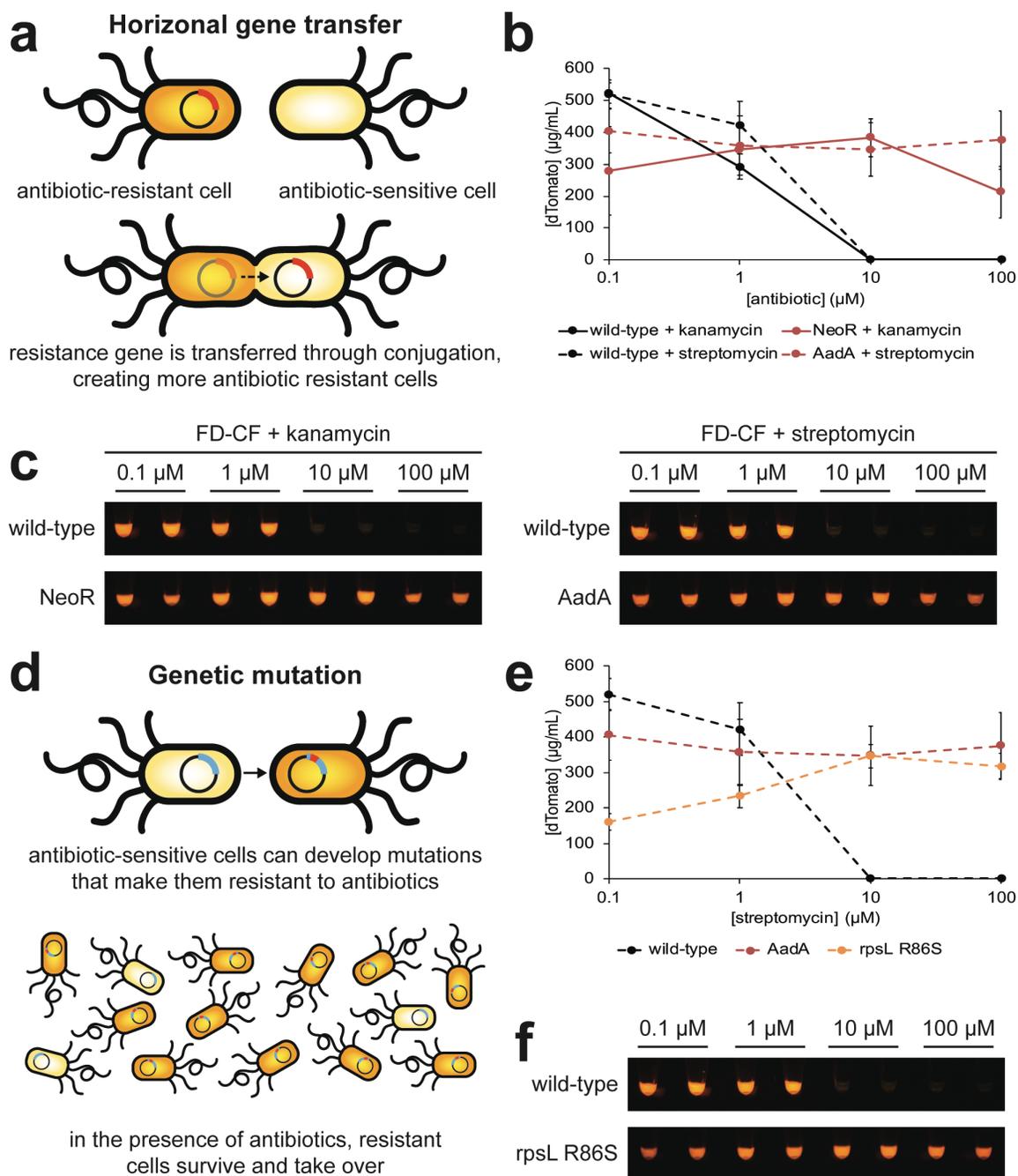
Figure 5.21 Demonstrating mechanisms of antibiotic resistance in FD-CF reactions. (**a**) To demonstrate the concept of horizontal gene transfer, kanamycin (*neoR*) and streptomycin (*aadA*) resistance genes were pre-expressed in FD-CF reactions for 20 hours at 30°C (**Figure 5.22**) and 1 µL of the soluble fraction was added to fresh FD-CF reactions encoding dTomato and containing 0.1-100 µM of either kanamycin or streptomycin. (**b**) Following cell-free protein synthesis for 20 hours at 30°C, reactions containing the resistance enzymes retain the ability to synthesize protein, even in the highest concentrations of antibiotics tested. Values represent averages and error bars represent standard deviations of $n \geq 3$ biological replicates. (**c**) When representative FD-CF reactions from (**b**) are imaged using a low-cost blue-light imager, differences in results using wild-type and resistant reactions can be distinguished by eye. (**d**) To illustrate genetic mutation and selection, we generated lysate from cells with

an R86S mutation in the *rpsL* gene, which are resistant to streptomycin. (**e**) FD-CF reactions with the resistant lysates also retain the ability to synthesize dTomato in the presence of up to 100 µM streptomycin. Values represent averages and error bars represent standard deviations of $n \geq 3$ biological replicates. (**f**) Images of representative FD-CF reactions from (**e**) imaged using a low-cost blue-light imager show that the difference between resistant and wild-type reactions can be observed qualitatively with the naked eye.
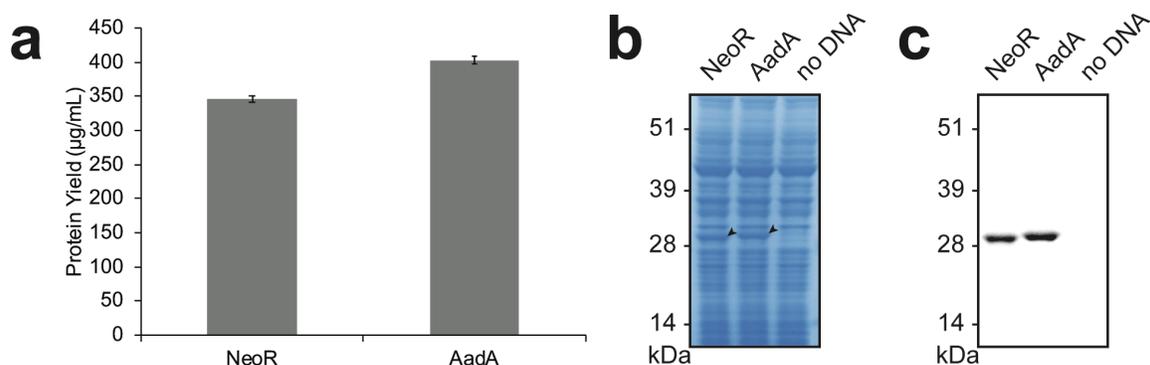
Figure 5.22 Expression of NeoR and AadA enzymes in FD-CF reactions. Following cell-free protein synthesis for 20 hours at 30°C, reactions containing ^{14}C-leucine were centrifuged at 20,000xg for 10 min to remove insoluble or aggregated protein products. (**a**) Soluble yields of the NeoR and AadA enzymes were measured by ^{14}C-leucine incorporation. Values represent averages and error bars represent standard deviations of n = 3 biological replicates. Soluble fractions were then analyzed by (**b**) SDS-PAGE and (**c**) ^{14}C-autoradiogram. Both enzymes expressed with exclusively full-length product observable by both SDS-PAGE and autoradiogram.
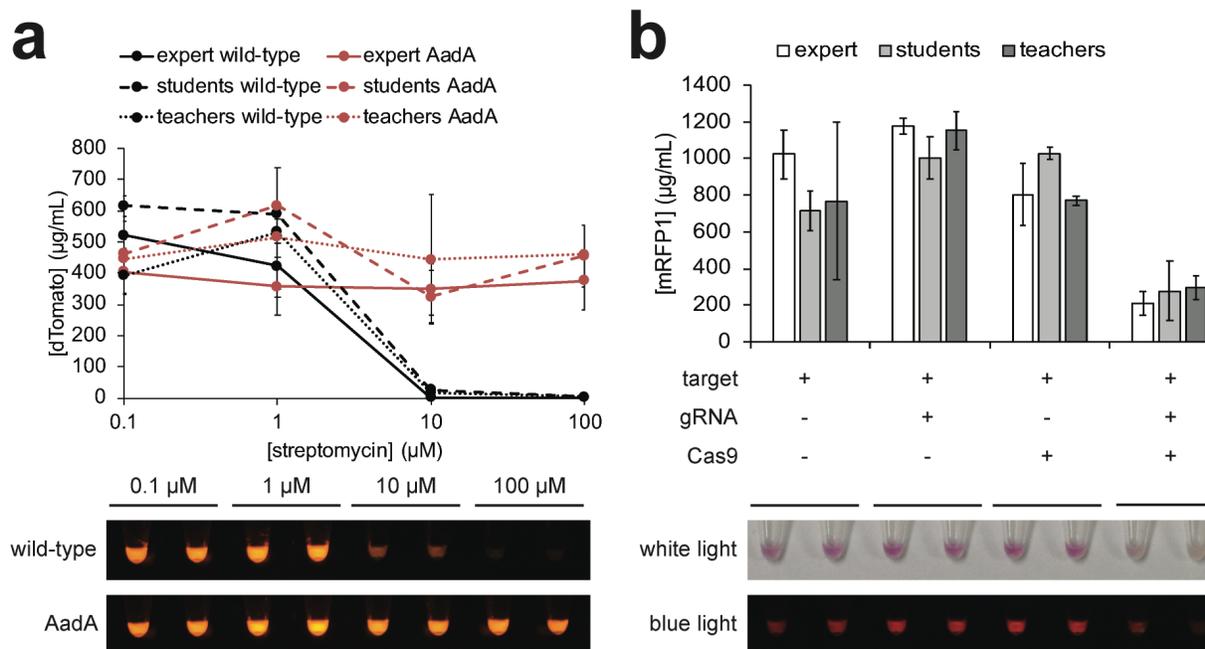
Figure 5.23 High school students and teachers can successfully run BioBits™ Health labs as first-time kit users. Antibiotic resistance and CRISPR gene editing lab activities were run by 5 high school students and 4 high school teachers. (**a**) Both students and teachers observed that FD-CF reactions containing the AadA resistance enzyme retain the ability to synthesize protein in the presence of up to 100 µM streptomycin (**top**) and can be visually distinguished from reactions lacking AadA (**bottom**). (**b**) Students and teachers observed 3.7 and 2.6-fold repression of mRFP1 fluorescence, respectively, when all CRISPR components were present in FD-CF reactions expressing mRFP1 (**top**). This level of repression was visible with the naked eye (**bottom**). Values represent averages and error bars represent average errors of $n \geq 2$ biological replicates. Images are representative examples of reactions assembled by high school students.

5.4.3.2 Module II: Fundamentals of CRISPR-Cas9 gene editing

CRISPR-Cas systems for editing DNA promise to reshape the way we understand and treat human diseases and genetic disorders[66,286]. For example, researchers recently showed that a CRISPR-based therapy could restore function of the protein dystrophin in mouse[287] and canine[288] models of Duchenne muscular dystrophy, and this therapy is now advancing toward human clinical trials. With gene editing therapies starting to make their way into the clinic and garnering broad public interest[289,290], it is important to educate students about the fundamentals of gene editing. Especially in light of the recent controversial report of CRISPR editing of humans in China{Cyranoski, 2019 #975, such educational activities could help students make informed decisions about the politics and ethics surrounding CRISPR and other gene editing technologies. To support this goal, we developed a classroom module that

requires students to (i) demonstrate understanding of the biological mechanism of CRISPR-Cas9 gene editing and (ii) consider the ethical implications of using CRISPR or other gene editing technologies to address societal problems.

In this module, we link activity of the *Streptococcus pyogenes* Cas9 nuclease to a fluorescent readout through the design of a synthetic guide RNA (gRNA) targeting the red fluorescent protein mRFP1, enabling straightforward investigation of the mechanism of action of CRISPR-Cas9. First, we showed that full-length *S. pyogenes* Cas9 nuclease could be expressed in FD-CF reactions (**Figure 5.24**). Next, we designed a gRNA construct to target a 20 base pair sequence in the first 50 nucleotides of the *mrfp1* gene and adjacent to an NGG protospacer adjacent motif (PAM) site that is required for *S. pyogenes* Cas9 activity (**Figure 5.25a**). To test whether our anti-mRFP1 gRNA could effectively repress mRFP1 expression and fluorescence, we rehydrated fresh FD-CF reactions encoding mRFP1 with gRNA plasmid, pre-expressed Cas9 nuclease, or both gRNA and pre-expressed Cas9. When both gRNA and Cas9 are present, we observe an approximately 4-fold reduction in fluorescent signal, indicative of Cas9 cleavage of the mRFP1 DNA template and silencing of fluorescent protein synthesis (**Figure 5.25b**). Similar levels of repression are observed in reactions assembled by high school student and teachers (**Figure 5.23b**). This experiment represents a simple lab activity that could be paired with a sorting and classification exercise like the one we developed (**Curriculum 10**) to help students understand the functions of the necessary components for Cas9 activity (*i.e.*, nuclease, gRNA, target, PAM site).
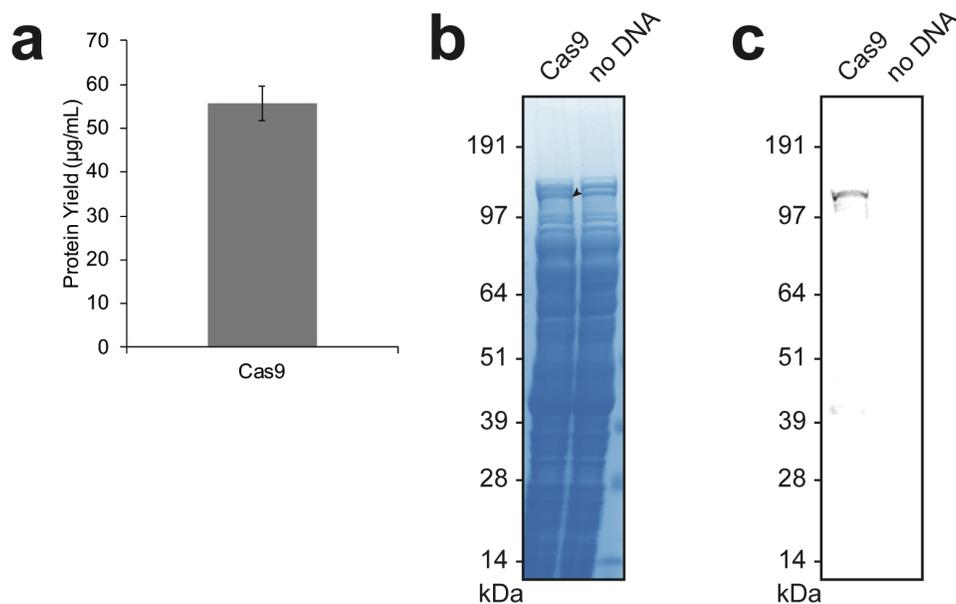
Figure 5.24 Expression of *S. pyogenes* Cas9 in FD-CF reactions. Following cell-free protein synthesis for 20 hours at 30°C, reactions containing ^{14}C-leucine were centrifuged at 20,000xg for 10 min to remove insoluble or aggregated protein products. (**a**) Soluble yield of *S. pyogenes* Cas9 was measured by ^{14}C-leucine incorporation. Value represents the average and error bars represent the standard deviation of *n* = 3 biological replicates. The soluble fraction was also analyzed by (**b**) SDS-PAGE and (**c**) ^{14}C-autoradiogram, with full-length product observed that is consistent with the predicted molecular weight of 158 kDa.
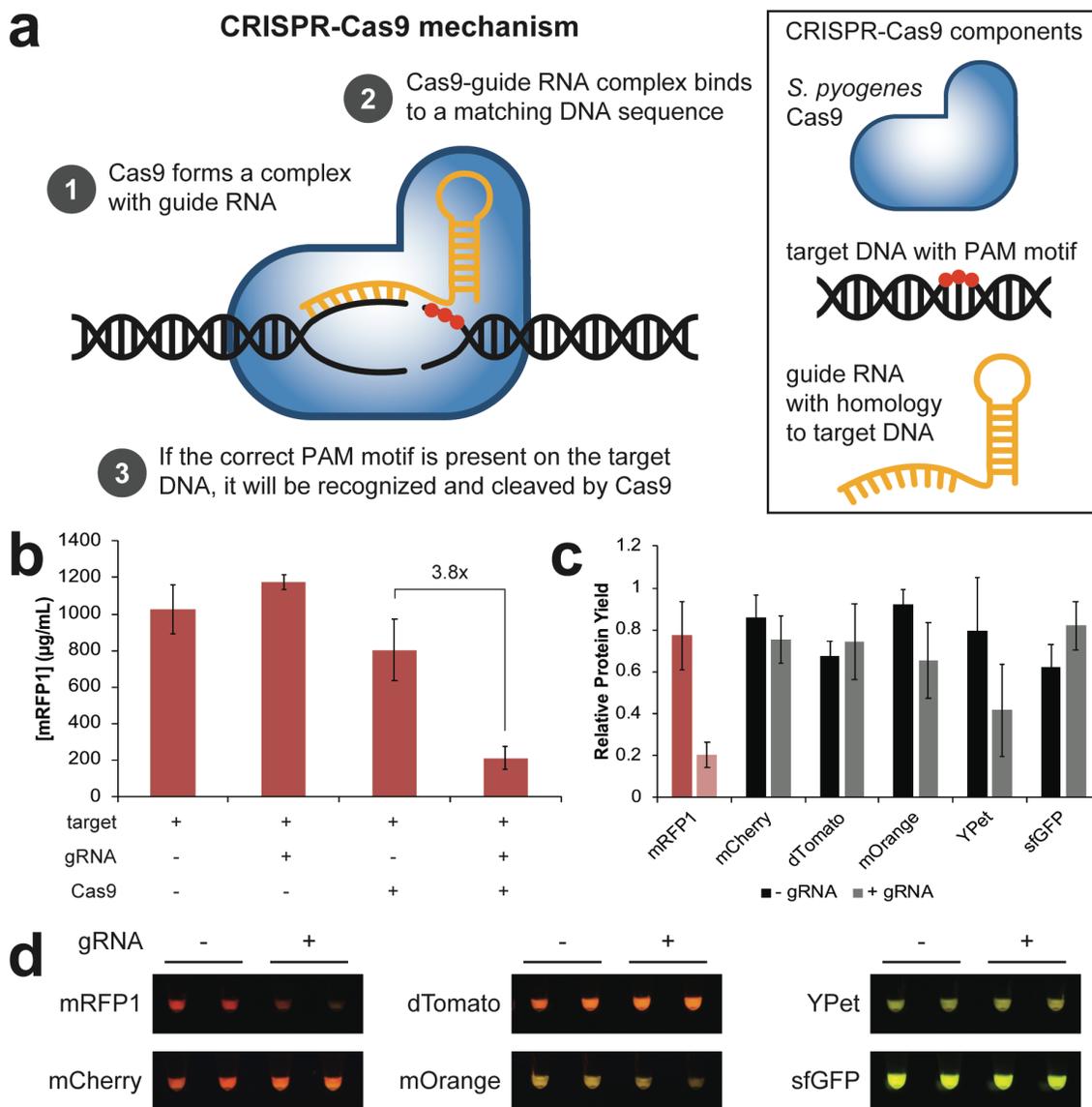
Figure 5.25 Interrogating the mechanism of a CRIPSR-Cas9 nuclease system using fluorescence. (**a**) To illustrate the mechanism of action of a CRISPR-Cas9 system (**left**) and outline the components required for Cas9 nuclease activity (**right**), we designed a synthetic guide RNA construct that targets the gene for the red fluorescent protein mRFP1. This makes it possible to use repression of mRFP1 fluorescence as a reporter of Cas9 activity in FD-CF reactions. (**b**) To test whether our anti-mRFP1 gRNA construct could effectively repress mRFP1 expression, we added gRNA plasmid, Cas9, or both gRNA plasmid and Cas9 to FD-CF reactions expressing mRFP1. The Cas9 nuclease from *S. pyogenes* was pre-expressed in FD-CF reactions for 20 hours at 30°C (**Figure 5.24**) and 1 µL of the soluble fraction was added to the mRFP1 reactions. When both gRNA plasmid and Cas9 are added to the reaction, ~4 fold repression of mRFP1 fluorescence was observed after incubation for 20 hours at 30°C. Values represent averages and error bars represent standard deviations of $n \geq 3$ biological replicates. (**c**) We next tested the orthogonality of our anti-mRFP1 gRNA construct by screening for Cas9 activity against a set of five other fluorescent proteins. Reactions contained template for the fluorescent protein of interest and pre-expressed Cas9, with or without anti-mRFP1 gRNA plasmid. Expression of the other fluorescent proteins tested is not markedly repressed by the anti-mRFP1 gRNA. Values represent averages and error bars represent standard deviations of $n \geq 3$ biological replicates. (**d**) Blue light images of representative FD-CF

reactions from (**c**) show that repression of mRFP1 can be observed with the naked eye, while fluorescence of the other protein targets is retained in the presence of gRNA.

Having demonstrated the activity of our gRNA construct against mRFP1, we designed a second lab activity that challenges students to discover which sequence of DNA this gRNA targets by screening its activity against a set of six fluorescent proteins. We measured production of each of the six fluorescent proteins in reactions containing the target DNA and pre-expressed Cas9, with or without gRNA plasmid. With the exception of its target, mRFP1, our gRNA construct did not greatly impact expression of the other fluorescent proteins, although our results suggest that there may be a low level of off-target activity against mOrange and YPet (**Figure 5.25c**). This is not surprising given that the *morange* and *ypet* gene sequences have homologies of ~79 and ~44% to *mrfp1*, respectively. This off-target activity offers the opportunity to discuss how non-specific targeting of genes is possible and must be thoroughly characterized to ensure the safety of using CRISPR technologies for clinical applications. Despite this low amount of off-target activity, repression of mRFP1 expression could be observed by eye when the reactions were imaged with our portable blue light imager, while fluorescence of the other target proteins was not visibly affected (**Figure 5.25d**). We developed pre-lab slides as well as a protocol with pre- and post-lab questions for running this lab in a high school classroom (**Curricula 11, 12**). Like the laboratories in Module I, this lab activity also offers the option for students to quantify their results with ImageJ (**Curriculum 6**) and carry out a statistical analysis of their results. This activity could open the door to additional exercises that investigate how gene editing technologies can be used to treat human diseases and how they compare to existing treatments. For example, we developed an assignment to help students investigate a variety of gene editing technologies (*i.e.*, CRISPR, zinc finger nucleases, etc.) for treating Huntington's disease (**Curriculum 13**), as well as a process oriented guided inquiry learning (POGIL) activity that explores a variety of possible solutions to treat a patient with sickle cell anemia (**Curriculum 14**). Finally, all of the activities described will prepare students for independent research projects in which they can investigate a problem of their choice that could be solved with CRISPR technology (**Curriculum 15**), and consider the ethics involved in such interventions (**Curriculum 16**). Overall, this module offers diverse, student-guided educational activities that can be used to meet

educational standards for high school biology (**Table 5.9**) and that are centered on an important, cutting-edge topic in synthetic biology. These activities have the potential to empower students with the knowledge to make informed ethical and policy decisions about how CRISPR technology could be used and regulated to yield the most benefit for society.

5.4.4 Discussion

We present here the BioBits™ Health educational kit and an accompanying collection of curricula and data for teaching health-related synthetic biology topics in high school classrooms. To enable BioBits™ Health, we designed wet lab activities with fluorescent readouts that can be run in FD-CF reactions with just a few simple inputs. These activities are organized into two lab modules that investigate antibiotic resistance and CRISPR-Cas9 gene editing technology. For both modules, we show that experimental results can be analyzed qualitatively by eye and successfully reproduced by high school students and teachers using the kit for the first time (**Figure 5.23**), demonstrating the utility of these resources for use by untrained operators without sophisticated laboratory equipment. Each lab activity can be set up in a single 1-hour class period and results can be analyzed as soon as the following day, limiting the class time required to run hands-on activities. Finally, we developed a set of example supporting curricula for facile integration of these activities into high school STEM classrooms. In pre- and post-assessments collected while beta-testing the BioBits™ Health lab activities with high school students, we found that students reported significantly increased confidence in their understanding of the mechanisms of antibiotic resistance and CRISPR-Cas9 genome editing as well as increases in their self-identification as engineers after running BioBits™ Health laboratories (**Figure 5.26**). While these data suggest that positive educational outcomes are possible using the BioBits™ Health resources, longitudinal studies will be needed to more fully assess educational benefits. It should also be noted that recent work has described educational activities for teaching CRISPR-Cas9 genome editing using cell-free systems{Collias, 2019 #974}, highlighting broad interest from educators in integrating this topic into educational curricula. However, these resources rely on expensive commercial cell-free kits and require access to expensive laboratory equipment such as temperature-controlled spectrophotometers and −80

°C freezers. Thus, our work meets a need for economical and accessible hands-on biology activities designed to teach cutting-edge health-related topics in high school classrooms.
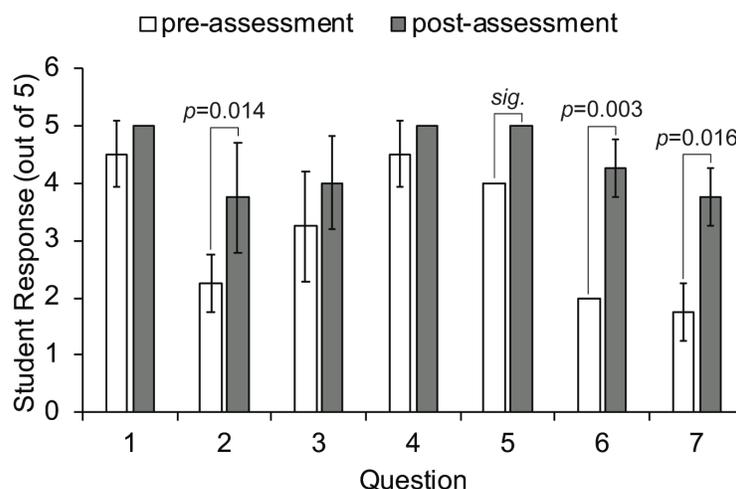
Students were asked to rate the extent to which they agreed or disagreed with the following statements (1 = strongly disagree, 3 = neutral, 5 = strongly agree):

1. Biology can be controlled or engineered
2. I see myself as an engineer
3. I see myself as a biologist or scientist
4. I understand how cells become resistant to antibiotics
5. I can explain to others how antibiotic resistance works
6. I understand CRISPR gene editing
7. I can explain to others how CRISPR works

Figure 5.26 Students' self-reported learning outcomes after running BioBits^{TM} Health labs. Statistically significant results from pre- and post-assessments show that high school students self-report increased self-identification as engineers as well as increased confidence in their understanding of antibiotic resistance and CRISPR-Cas9 genome editing mechanisms after running BioBits^{TM} Health labs. Statistical significance was assessed using paired t-tests comparing pre- and post-assessment responses for each question. If a statistically significant *p* value was obtained, its value is reported. For question 5, the variance of both the pre- and post-assessment responses was zero and a *p*-value could not be calculated, despite the fact that there is no statistical doubt that the responses are significantly different (labeled *sig.*). Differences between pre- and post- assessment responses were not found to be statistically significant for all other questions. Values represent means and error bars represent standard deviations from *n* = 4 student responses (one student did not complete assessments).

Future development of the BioBits^{TM} Health kit will involve scaling its production, further developing and beta-testing supporting curricula, and implementation in high school classrooms where both short- and long-term educational impact can be quantitatively measured. We recently launched a

website (www.mybiobits.org) that we plan to use as the online home for an open-source community centered around the BioBits™ kits. The website currently houses a curriculum database where the BioBits™ curricula can be freely accessed. In the future, we plan to provide information on the Web site about how (i) educators can order kits or participate in pilot programs as these become available, (ii) kit users can share data, and (iii) researchers and educators can upload new lab activities or curriculum pieces that use FD-CF reactions as the community of BioBits™ users and developers grows. We also think there are many exciting future directions for health-related educational modules using FD-CF technology. In particular, RNA toehold switches[51,66] or CRISPR-based RNA[291] or DNA[292] sensing technologies could be adapted to detect infectious agents (for example, viruses and pathogenic bacteria) in environmental samples or cancer-causing mutations in mock patient samples. Implementing such tools in FD-CF reactions could support classroom activities on infectious disease and cancer, respectively.

We designed the BioBits™ Health kits to be economically accessible, priced at less than $150 per 30-person classroom (**Table 5.10**) and within the range of prices teachers reported being willing to pay for such a kit (**Figure 5.27**). This includes a complete set of reagents to run all of the lab activities described here, as well as the custom, low-cost blue light imager and incubator that we developed as part of the BioBits™ Bright kit[248]. We also show that FD-CF reactions are stable for two weeks at room temperature and at least six months in a commercial refrigerator using packaging that could easily and economically be replicated for shipping to classrooms (**Figure 5.28**). This allows for ambient temperature shipping and short-term storage of kits as well as longer term refrigerated storage if necessary. In addition, production of the FD-CF reactions using gas-flushing or modified atmosphere packaging could extend the shelf-life to years, as shown in previous studies[14,50,184]. Due to the highly portable, cost-effective, and user-friendly nature of the reagents and lab activities, the BioBits™ Health kit has utility both inside and outside of a formal classroom or laboratory setting. We anticipate that these resources will increase access to cell-free technologies, enhance basic biology education through integration of cutting-edge health topics, and help the next generation of students make informed decisions about new and transformative biological technologies.
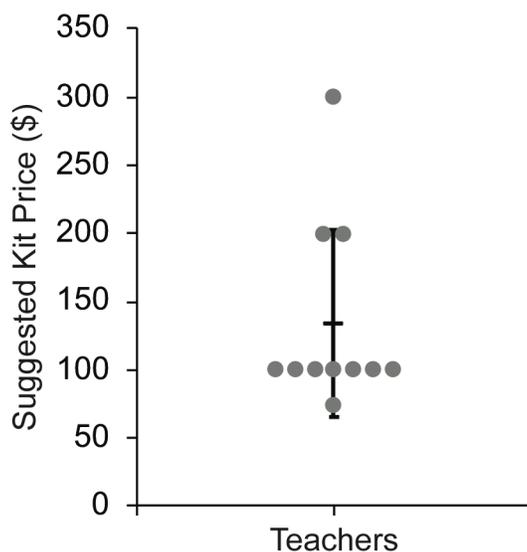
Figure 5.27 Teachers' suggested pricing for BioBits™ Health kit. We asked 11 high school teachers what they thought would be a fair price to pay for the BioBits™ Health kit. Their responses ranged from $75-$300, with an average suggested price of ~$134 and a standard deviation of ~$69. This demonstrates that our kit, which we estimate will cost $150 to produce, is likely to be economically accessible. Additionally, some elements of the kit, such as the portable imager and incubator, will likely become less expensive per unit as economies of scale are achieved. For this reason, we may be able to reduce the cost of the BioBits™ Health kit through the scale up and commercialization process, further increasing its accessibility.
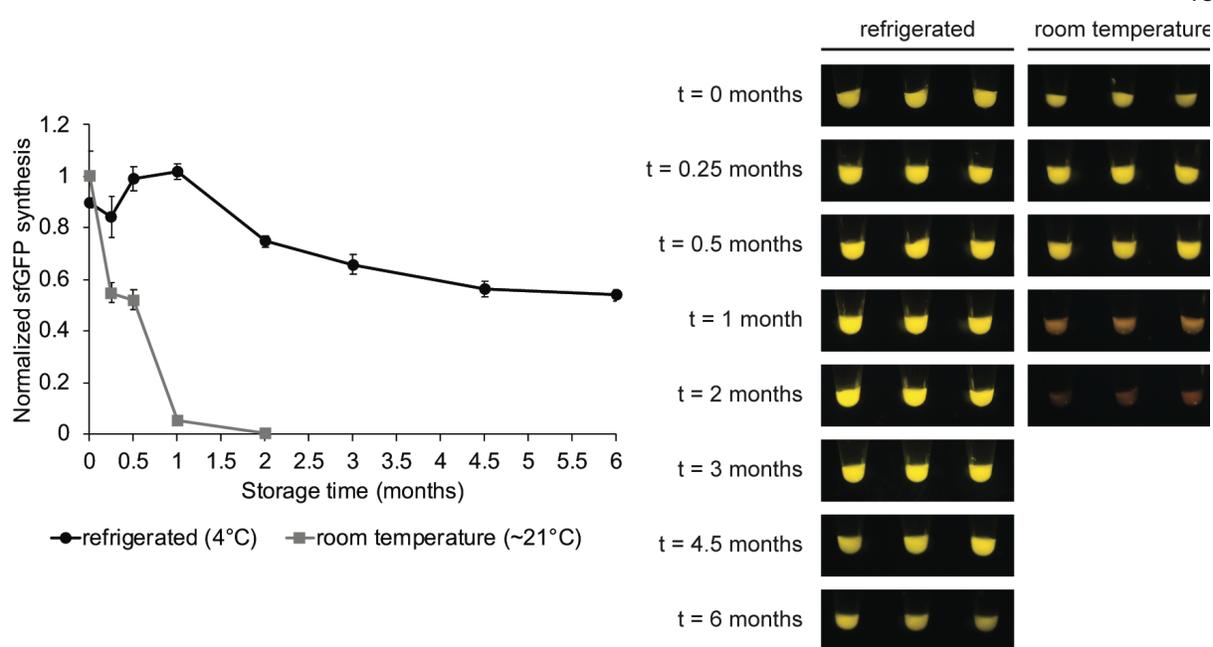
Figure 5.28 Long-term stability of FD-CF reactions under ambient temperature and refrigerated storage conditions. Stability of 5 µL FD-CF reactions was tested over a period of 6 months using packaging that could easily and economically be replicated for shipping to classrooms. Specifically, we vacuum sealed reactions using a commercial FoodSaver® appliance with Dri-Card™ desiccant cards enclosed to prevent rehydration of the FD-CF pellets. At each time point, reactions were rehydrated with 5 ng of pJL1 sfGFP plasmid and incubated at room temperature for 24 hours before quantification and imaging. FD-CF reactions were stable at room temperature for 2 weeks and in a commercial refrigerator for at least 6 months (**left**). Despite the fact that sfGFP synthesis is reduced to ~50% of initial yields after storage for 2 weeks at room temperature or 6 months at 4°C, no visible decrease in protein synthesis is observed (**right**). This allows for ambient temperature shipping and short-term storage of kits, as well as longer term refrigerated storage if necessary. Values represent averages and error bars represent standard deviations of n = 3 biological replicates.

5.5 Summary

Hands-on science activities are known to improve student performance and engagement, but implementation of hands-on biology curriculum has traditionally been challenging due to the expensive equipment and expertise required to grow cells. We developed the BioBits™ Bright, Explorer, and Health kits as user-friendly alternatives that make hands-on molecular and synthetic biology activities possible in classrooms and other non-laboratory settings. Our work has been featured on the covers of *Science Advances* and *ACS Synthetic Biology* and has attracted the attention of scientists and educators worldwide. Since the papers describing the BioBits™ Bright and Explorer kits were published, we have

received about 400 requests for kits from over 20 countries through our website (www.mybiobits.org).

We carried out a pilot program from December 2018 to June 2019, that provided prototype BioBits™ Bright kits to 100 high school classrooms, reaching an estimated 2500 students. We are currently partnering with miniPCR, an established educational company, to launch the first commercial BioBits™ kit in September 2019. Because they contain all of the necessary reagents and equipment for running labs with easy-to-use reactions, we anticipate that the BioBits™ kits will significantly lower the barrier to entry to educational hands-on biology activities. These resources have the potential to broaden participation in molecular biology and synthetic biology education and offer much needed alternatives for students and classrooms that do not have access to the resources required to perform cell-based biology experimentation.

5.6   Additional information

5.6.1 Acknowledgements

BioBits^{TM} Explorer

of Air Force Research Laboratory, Air Force Office of Scientific Research, Defense Threat Reduction

Agency, or the U.S. government.

5.6.2 Author contributions

BioBits^{TM} Bright

J.C.S., A.H., and P.Q.N. designed research, performed research, analyzed data, and wrote the

paper. J.B., T.M., A.M.-B., A.P., K.R., M.S., and L.B. performed research and designed curricula. T.C.F.,

R.S.D., K.J.H., M.A., A.K., Q.M., J.S.P., R.P., P.P., D.Q., T.Z., L.R.H., J.F.C., N.F., S.F., E.G., E.M.G.,

T.G., J.K., B.N., S.O., C.P., A.P., S.S., A.S., and T.W. performed research. N.D. and K.P. aided in research design. M.C.J. and J.J.C. directed research, analyzed data, and wrote the paper.

BioBits^{TM} Explorer

A.H., P.Q.N., and J.C.S. designed the research, performed the research, analyzed the data, and wrote the manuscript. N.D., T.F., M.K.T., and A.J.D. aided in research design and performed the research. R.S.D. and K.J.H. performed the research. K.P. aided in research design. J.J.C. and M.C.J. directed the research.

BioBits^{TM} Health

J.C.S. designed research, performed research, analyzed data, and wrote the paper. A.H. designed research, designed curricula, and edited the paper. J.F., T.M., S.M., F.R., and M.W. performed research and designed curricula. R.S.D., K.J.H., B.H., R.J., R.K., V.K., and W.W. performed research. G.A.R. and P.Q.N. aided in research design, designed curricula, and edited the paper. T.M. aided in curricula design. M.C.J. directed research, analyzed data, and wrote the paper.

5.6.3 Publishing and trademark information

Published manuscripts

Stark, J. C. †, Huang, A. †, Nguyen, P. Q. †, Dubner, R. S., Hsu, K. J., Ferrante, T. C., … , Collins, J. J. *, & Jewett, M. C. * (2018) BioBits^{TM} Bright: A fluorescent synthetic biology education kit. *Science Advances*, doi: 10.1126/sciadv.aat5107.

Huang, A. †, Nguyen, P. Q. †, Stark, J. C. †, Donghia, N., Ferrante, T. C., … , Jewett, M. C. *, & Collins, J. J. * (2018) BioBits^{TM} Explorer: A modular synthetic biology education kit. *Science Advances*, doi: 10.1126/sciadv.aat5105.

Stark, J. C., Huang, A., Hsu, K. J., Dubner, R. S., Forbrook, J., Marshalla, S., Rodriguez, F., Washington, M., … , & Jewett, M. C. * (2018) BioBits^{TM} Health: Classroom activities exploring engineering, biology, and human health with fluorescent readouts. *ACS Synthetic Biology,* doi: 10.1021/acssynbio.8b00381.

† contributed equally; * corresponding author

Trademark application

BioBits^{TM}. 2018. Educational kits sold as a unit in the field of biology consisting primarily of biological components, scientific apparatus for classroom experimentation, and downloadable digital content.

5.7   Tables

Table 5.1 Plasmids used in the BioBits™ studies.

Plasmid	Sequence Source	Plasmid Source	Addgene #
pJL1 mCherry	Clontech 632524	[248]	102629
pJL1 mRFP1	BBa_E1010	[248]	102630
pJL1 eforRed	BBa_K592012	[248]	106320
pJL1 dTomato	Addgene 54856	[248]	102631
pJL1 mOrange	Addgene 54751	[248]	102632
pJL1 YPet	Addgene 54860	[248]	102633
pJL1 sfGFP	Addgene 69496	Jewett lab stock	102634
pJL1 mTFP1	Addgene 54553	[248]	102635
pJL1 CyPet	Addgene 14030	[248]	102636
pJL1 Aquamarine	Addgene 42889	[248]	106285
pJL1 mTagBFP2	Addgene 54572	[248]	102638
pJL1 mKalama1	Addgene 14892	[248]	102639
pJL1 eBFP2	Addgene 14891	[248]	102640
pJL1 ATF1	BBa_J45014	[66]	106286
pJL1 Sortase	Linda Griffith (MIT)	[66]	106288
pNP1 Ecarin	US Patent #US20050164365	[66]	106289
pJL1 Trx-Bx	Trx: Uniprot #P0AA25 Bx: UniProt #P04971.1	[66]	106290
pCOLA banana sensor sfGFP	This work	[66]	107367
pCOLA kiwi sensor sfGFP	This work	[66]	107368
pKD46	—	[95]	—
pJL1 neoR	Addgene 69496	[293]	117048
pJL1 aadA	EXPRESSYS pZS4Int-tetR	[293]	117050
pJL1 *Sp*Cas9	Addgene 62225	[293]	117051
pJL1 mRFP1 gRNA	—	[293]	117052

Table 5.2 Diverse fluorescent protein library enables educational kit development. A 13-member fluorescent protein library was designed to include red, orange, yellow, green, teal, and blue fluorescent protein variants, which were cloned into the in vitro expression vector pJL1. PDB accession numbers are provided if the protein (or a closely related variant) has been crystallized.

Protein	Color	Excitation (nm)	Emission (nm)	PDB entry
mCherry	Red	587	610	2H5Q
mRFP1	Red	584	607	2VAD
eforRed	Red	587	610	2VAD
dTomato	Orange	554	581	-
mOrange	Orange	548	562	2H5O
YPet	Yellow	517	530	1F0B
sfGFP	Green	485	528	2B3P
mTFP1	Cyan	462	492	4Q9W
CyPet	Cyan	435	477	3I19
Aquamarine	Cyan	420	474	2WSN
mTagBFP2	Blue	399	454	3M24
mKalama1	Blue	385	456	4ORN
eBFP2	Blue	383	448	1BFP

Table 5.3 Cost analysis of portable imagers and incubators. The total cost to build working prototypes of our portable 8-well imager, 96-well imager, and incubator (switch and dial versions) are calculated below. Purchasing information for materials used to construct the prototypes are also included.

Component	Cost ($/unit)	Manufacturer	Product No
8-well Imager Total	**14.13**		
Black acrylic	0.95	Inventables	24112-03
Transparent yellow acrylic	0.12	Inventables	24165-03
Transparent orange acrylic	0.12	Inventables	24164-03
SCIGRIP Weld-On 16 Cement	0.12	TAP Plastics	10886
448 nm blue LED	6.97	Luxeon Star LEDs	SP-05-V4
Thermal adhesive tape for LEDs	0.94	Luxeon Star LEDs	LXT-R-10
Aluminum	0.35	McMaster-Carr	9146T38
Battery pack (3 AAA) with switch	1.56	Digi-Key	SBH431-1AS-ND
AAA rechargeable batteries	3.00	Amazon	RFQ420
96-well Imager Total	**32.01**		
Black acrylic	7.05	Inventables	24112-01
Transparent yellow acrylic	0.85	Inventables	24165-03
Transparent orange acrylic	0.85	Inventables	24164-03
SCIGRIP Weld-On 16 Cement	0.29	TAP Plastics	10886
448 nm blue LEDs (2)	13.94	Luxeon Star LEDs	SP-05-V4
Thermal adhesive tape for LEDs	3.77	Luxeon Star LEDs	LXT-R-10
Aluminum	0.70	McMaster-Carr	9146T38
Battery pack (3 AAA) with switch	1.56	Digi-Key	SBH431-1AS-ND
AAA rechargeable batteries	3.00	Amazon	RFQ420
Switch Incubator Total	**19.79**		
Black medium-density fiberboard	5.67	Inventables	30463-04
Screw, 0.25 in	1.20	Inventables	25297-01
Gorilla® wood glue	0.20	Inventables	26032-01
Mini solder-able breadboard	2.95	SparkFun	PRT-12702
N-channel MOSFET	0.95	SparkFun	COM-10213
Heating pad	3.95	SparkFun	COM-11288
MicroB USB breakout	1.95	SparkFun	BOB-12035
10K resistor	0.10	Digi-Key	10.0KXBK-ND
1-20K resistor (2x)	0.20	Digi-Key	variable
110K resistor (2x)	0.20	Digi-Key	110KXBK-ND
Temperature sensor	1.84	Digi-Key	TC622VAT-ND
Switch	0.58	Digi-Key	EG1903-ND
Dial Incubator Total	**19.67**		
Black medium-density fiberboard	5.67	Inventables	30463-04
Screw, 0.25 in	1.20	Inventables	25297-01
Gorilla® wood glue	0.20	Inventables	26032-01
Mini solder-able breadboard	2.95	SparkFun	PRT-12702
N-channel MOSFET	0.95	SparkFun	COM-10213
Heating pad	3.95	SparkFun	COM-11288
MicroB USB breakout	1.95	SparkFun	BOB-12035
10K resistor	0.10	Digi-Key	10.0KXBK-ND
110K resistor	0.10	Digi-Key	110KXBK-ND
Temperature sensor	1.84	Digi-Key	TC622VAT-ND
Potentiometer	0.76	Digi-Key	987-1726-ND

Table 5.4 Cost analysis for BioBits™ Bright. An estimate of the total cost to assemble BioBits™ Bright is calculated below. BioBits™ Bright will include enough reagents to run lab modules I and II for a 30-person classroom with groups of 2 students. The kit also includes our low-cost, portable imagers and incubators to enable use outside of a laboratory setting or in resource-limited classrooms. See **Table 5.5** for a detailed cost analysis of FD-CF reactions.

Component	Cost ($/unit)	Quantity	Cost in Kit ($)	Module
5 µL FD-CF rxns	0.05	150	7.50	I
96-well FD-CF rxns	1.00	15	15.00	II
Incubator	19.79	1	19.79	I & II
8-well imager	14.13	1	14.13	I & II
96-well imager	32.01	1	32.01	I & II
Freeze-dried plasmid (1200+ rxns)	1.60	6	9.60	I & II
Nuclease-free water (mL)	0.04	25	1.07	I & II
		BioBits™ Bright Total	**99.10**	**$/kit**

Table 5.5 Cost analysis of FD-CF reactions. The total cost to assemble FD-CF reactions is ~$0.01 per µL. This comes out to ~$0.05 per 5 µL reaction used in the DNA titration module, or ~$1.00 per 96 well plate used in the design-build-test module. In the table, amino acid cost accounts for 2 mM each of the 20 canonical amino acids purchased individually from Sigma. Extract cost is based on a single employee making 50 mL lysate from a 10 L fermentation, assuming 30 extract batches per year and a 5-year equipment lifetime. Component source is also included in the table if it is available to purchase directly from a supplier. Homemade or user-supplied components cannot be purchased directly and must be prepared by the end user according to procedures described in the Methods section.

Component	Cost ($/µL rxn)	Supplier	Product No
Mg(Glu)$_2$	negligible	Sigma	49605
NH$_4$Glu	negligible	MP	02180595
KGlu	negligible	Sigma	G1501
ATP	negligible	Sigma	A2383
GTP	0.00026	Sigma	G8877
UTP	0.00022	Sigma	U6625
CTP	0.00019	Sigma	C1506
Folinic acid	0.00002	Sigma	47612
tRNA	0.00021	Roche	10109541001
Amino acids	negligible	Sigma	
PEP	0.00178	Roche	10108294001
NAD	negligible	Sigma	N8535-15VL
CoA	0.00033	Sigma	C3144
Oxalic acid	negligible	Sigma	P0963
Putrescine	negligible	Sigma	P5780
Spermidine	negligible	Sigma	S2626
HEPES	negligible	Sigma	H3375
Extract	0.00737	homemade	
Total	**0.01**	**$/µL rxn**	
	0.05	**$/5 µL rxn**	
	1.00	**$/96 well plate**	

Table 5.6 Cost analysis for BioBits™ Explorer.

BioBits™ Explorer – Core Kit

Component	Cost/rxn ($)	Quantity	Cost in Kit ($; w/ 10 repeats)
Activity: Co-expressed fluorescent proteins			
Plasmids	0.04	16	6.40
Cell-free reagents	0.10	16	16.00
		Sub-total	**22.40**
Activity: Odor-generating enzymes			
Plasmids	0.04	3	1.20
Cell-free reagents	0.10	3	3.00
Odor substrates	2.29	3	68.70
		Sub-total	**72.90**
Activity: Sortase-generated hydrogels			
Plasmids	0.07	3	2.10
Cell-free reagents	0.21	3	6.30
Hydrogel substrates	2.34	3	70.20
		Sub-total	**78.60**
Reusable components			
96-well incubator			20.00
8-well incubator			15.00
Nuclease-free water			1.00
		Sub-total	**36.00**
		Total	**209.90**

BioBits™ Explorer – Add-on Kit

Component	Cost/rxn ($)	Quantity	Cost in Kit ($; w/ 10 repeats)
Activity: Fruit sensors			
DNA extraction supplies	0.21	4	8.40
RPA reagents (1 rxn supplies whole class)	4.27	4	17.08
Cell-free reagents	4.41	4	176.40
		Total	**201.88**

Table 5.7 Primers used in the BioBits™ Health study. Primers used to construct and verify the BL21 Star™ (DE3) rpsL R86S mutant strain are listed below. The first four bases (from the 5' end) of the MAGE oligonucleotide were phosphorothioated (*) for improved editing efficiency[281]. Underlined bold text in the MAGE oligonucleotide indicates the location of the C256A mismatch mutation.

Primer Name	DNA Sequence (5' to 3')
rpsL MAGE oligo	G*T*C*A*AGCGCACCACGTACGGTGTGGTAACGAACACCCTGGAGGTCTTTA ACAC**T**ACCGCCACGGATCAGGATCACGGAGTGCTCCTGCAG
rpsL seq for	GGCGTATGTACTCGTGTATATACTACC
rpsL rev	CCATACTTGGAACGAGCCTGC

Table 5.8 Reagents needed to make CFPS extracts and assemble cell-free reactions. Vendors and catalog numbers for necessary reagents are listed below.

Component	Supplier	Product No
Tryptone	Fisher	211699
Yeast extract	Fisher	212720
Sodium chloride	Sigma	S3014
Glucose	Sigma	G8270
K_2HPO_4	Sigma	60353
KH_2PO_4	Sigma	P9791
IPTG	Sigma	I6758
TrisOAc	Sigma	T6066
MgOAc	Sigma	M5661
KOAc	Sigma	P1190
$Mg(Glu)_2$	Sigma	49605
NH_4Glu	MP	2180595
KGlu	Sigma	G1501
ATP	Sigma	A2383
GTP	Sigma	G8877
UTP	Sigma	U6625
CTP	Sigma	C1506
Folinic acid	Sigma	47612
tRNA	Roche	10109541001
Amino acids	Sigma	LAA21-1KT
PEP	Roche	10108294001
NAD	Sigma	N8535-15VL
CoA	Sigma	C3144
Oxalic acid	Sigma	P0963
Putrescine	Sigma	P5780
Spermidine	Sigma	S2626
HEPES	Sigma	H3375
Plasmid prep kit	Omega Bio-Tek	D6904-04

Table 5.9 NGSS alignment of the BioBits™ Health kit. A list of next-generation science standards for high school life science that could be either partially or fully met using the BioBits™ Health kit is provided.

NGSS High School Life Science Standard	Module
HS-LS1-1: Construct an explanation based on evidence for how the structure of DNA determines the structure of proteins which carry out the essential functions of life through systems of specialized cells.	I & II
HS-LS1-6: Construct and revise an explanation based on evidence for how carbon, hydrogen, and oxygen from sugar molecules may combine with other elements to form amino acids and/or other large carbon-based molecules.	I & II
HS-LS3-1: Ask questions to clarify relationships about the role of DNA and chromosomes in coding the instructions for characteristic traits passed from parents to offspring.	II
HS-LS3-2: Make and defend a claim based on evidence that inheritable genetic variations may result from (1) new genetic combinations through meiosis, (2) viable errors occurring during replication, and/or (3) mutations caused by environmental factors.	I & II
HS-LS4-2: Construct an explanation based on evidence that the process of evolution primarily results from four factors: (1) the potential for a species to increase in number, (2) the heritable genetic variation of individuals in a species due to mutation and sexual reproduction, (3) competition for limited resources, and (4) the proliferation of those organisms that are better able to survive and reproduce in the environment.	I
HS-LS4-3: Apply concepts of statistics and probability to support explanations that organisms with an advantageous heritable trait tend to increase in proportion to organisms lacking this trait.	I
HS-LS4-4: Construct an explanation based on evidence for how natural selection leads to adaptation of populations.	I
HS-LS4-5: Evaluate the evidence supporting claims that changes in environmental conditions may result in (1) increases in the number of individuals of some species, (2) the emergence of new species over time, and (3) the extinction of other species.	I & II
HS-LS4-6: Create or revise a simulation to test a solution to mitigate adverse impacts of human activity on biodiversity.	I & II

Table 5.10 Cost analysis of the BioBits™ Health kit. An estimate of the total cost to assemble BioBits™ Health is calculated below. The BioBits™ Health kit will include enough reagents to run lab modules I and II in groups of 2 students for a 30-person classroom. A detailed cost analysis of our custom incubator and 8-well imager is described in **Table 5.3** and a similar cost analysis of our FD-CF reactions is described in **Table 5.5**.

Component	Cost ($/unit)	Quantity	Cost in Kit ($)	Module
5 µL FD-CF rxns	0.05	1920	96.00	I
15 µL FD-CF rxns	0.16	24	3.84	I & II
Antibiotics	varying	6	negligible	I
Incubator	19.79	1	19.79	I & II
8-well imager	14.13	1	14.13	I & II
Freeze-dried plasmids	varying	10	8.99	I & II
Nuclease-free water (mL)	0.04	150	6.41	I & II
	BioBits™ Health Total		**149.16**	**$/kit**

6. Summary of Research Achievements

The goal of my PhD research has been to develop transformative technologies for point-of-care production of glycoprotein medicines and vaccines. The World Health Organization estimates that 30% of the world's population lacks access to essential medicines. This disparity necessitates new technologies for delivering medicines to the developing world. A key challenge is our inability to establish refrigerated supply chains in remote or resource-limited settings. Refrigeration is required to support current manufacturing strategies that produce protein medicines in centralized production facilities. Cell-free protein synthesis (CFPS) is an emerging technology that offers a unique solution to this problem by enabling decentralized, point-of-care production of protein medicines. This is possible because cell-free reactions synthesize proteins in as little as 1 hour and can be freeze-dried for distribution without refrigeration. However, existing CFPS systems are limited in their ability to efficiently co-activate protein synthesis and glycosylation, or the decoration of proteins with sugars, which is required for many classes of protein medicines.

To address this limitation, I developed the first cell-free glycoprotein synthesis (CFGpS) platform with the ability to produce glycosylated protein therapeutics at the point-of-care (Chapter 2)[134]. To do this, I generated *Escherichia coli* cell lysates containing the biological machinery for both protein synthesis and glycosylation, making it possible to produce glycoproteins in simple, one-pot reactions. I demonstrated that CFGpS lysates can produce relevant glycoprotein medicines, including the therapeutic hormone erythropoietin. Further, CFGpS has enabled additional platforms for characterization of glycosylation enzymes[97,100]. I have been selected to give oral presentations on this work at three conferences and was recognized with an oral presentation award at the 2016 American Institute of Chemical Engineers national meeting. In parallel, I reduced cell-free reaction cost in a yeast CFPS system by activating native metabolism to power protein synthesis. This work promises to increase access to yeast CFPS technology and could be combined with advancements in yeast glycoengineering[60-62] to develop an additional platform for cell-free glycoprotein synthesis (Chapter 4)[294]. I was recognized for this work with 2nd place at

the 2015 International Conference on Biomolecular Engineering. My novel CFGpS technology promises to enable point-of-care production of glycoprotein medicines as well as fundamental studies of glycosylation enzymes and glycoprotein properties.

Development of the CFGpS platform coincided with my efforts to generate the first technology for point-of-care production of antibacterial vaccines. Drug-resistant bacteria are predicted to take more lives than cancer by the year 2050, necessitating new strategies for preventative vaccination. In particular, conjugate vaccines are over 90% effective at preventing bacterial infections with extremely rare instances of resistance. However, despite their proven safety and efficacy, global childhood vaccination rates for conjugate vaccines remain as low as ~30%, with lack of access or low immunization coverage accounting for the vast majority of remaining disease burden. *To address this issue, I developed a modular technology for in vitro bioconjugate vaccine expression (iVAX) in portable, freeze-dried lysates from detoxified, nonpathogenic Escherichia coli (Chapter 3)*[295]. iVAX combines cell-free protein synthesis of licensed vaccine carrier proteins and *N*-linked glycosylation of bacterial polysaccharide antigens. I showed that iVAX can be used to synthesize bioconjugate vaccines directed against the highly virulent Schu S4 strain of *Franciscella tularensis*, the causative agent of tularemia and a potential bioterrorism threat, as well as pathogenic *E. coli* strains O78 and O7. Importantly, I showed that anti-*F. tularensis* vaccines can be synthesized in iVAX reactions for ~$6 per human dose and elicited pathogen-specific antibodies in mice at significantly higher levels compared to a control vaccine produced using a state-of-the-art biomanufacturing approach. I have been selected to give oral presentations on this work at four conferences and was recognized with an oral presentation award at the 2018 American Chemical Society national meeting. Overall, my iVAX platform promises to accelerate development of new bioconjugate vaccines with increased access by enabling refrigeration-independent distribution and point-of-care production.

Finally, I merged my thesis research with my passion for outreach by using freeze-dried cell-free technology to develop just-add-water biology education kits. In 2012, the National Science Foundation estimated that 20% of US high school students fail to take a science class by grade 11, highlighting the need to create educational tools that stimulate interest in science. Hands-on science activities are known

to improve student engagement, but implementation of hands-on biology curriculum is challenging due to the expensive equipment and expertise required to grow cells. *I developed the BioBits™ Bright[248], Explorer[66], and Health[293] kits as user-friendly alternatives that make hands-on molecular and synthetic biology activities possible in classrooms and other non-laboratory settings (Chapter 5).* This work has been featured on the covers of *Science Advances* and *ACS Synthetic Biology* and was highlighted in *Science* magazine and on NPR. In addition, I was selected to give an oral presentation on BioBits™ at the 2018 Engineering Biology Research Consortium meeting. Since the Bright and Explorer kits were published, my collaborators and I have received about 400 requests for kits from over 20 countries through our website (www.mybiobits.org). We recently secured sufficient funding to carry out a pilot program from December 2018 to June 2019, that provided prototype BioBits™ Bright kits to 100 high school classrooms, reaching an estimated 2500 students. We are currently partnering with miniPCR, an established educational company, to launch the first commercial BioBits™ kit in September 2019.

My CFGpS and iVAX platforms represent important first steps toward a new manufacturing paradigm for decentralized production of glycosylated protein medicines. These technologies join an emerging set of on-demand biomanufacturing platforms[10-14,168,169] that have the potential to increase global access to costly drugs. In addition, my inexpensive and user-friendly BioBits™ kits alleviate many of the challenges associated with implementing hands-on biology experiments in classrooms. As such, these kits have the potential to increase scientific literacy through the integration of cutting-edge molecular and synthetic biology topics into K-12 STEM education. In recognition of the potential impact of my thesis research, I was recognized with the Northwestern Chemical and Biological Engineering Department's 2018 Distinguished Graduate Researcher Award. Overall, my work lays the foundation for transformative advances in medicine, biotechnology, fundamental biology, and education that will benefit both the developed and the developing world.

7.   Future Directions and Outlook

While this work represents notable advances in our ability to manufacture glycoprotein therapeutics on demand and translate cutting-edge science into hands-on educational activities, there remain many significant research opportunities in each of these areas. For example, while the CFGpS platform described in Chapter 2 enables on-demand production of structurally diverse glycoproteins *in vitro*, this technology remains limited in its ability to produce human protein therapeutics decorated with human glycan structures. There are two reasons for this: (i) our use of the *Cj*PglB enzyme to catalyze glycosylation reactions and (ii) challenges associated with biosynthesis of the human biantennary *N*-glycan structure in *E. coli*. We chose to use *Cj*PglB for the majority of this work because its activity has been extensively characterized[21,96,151]. However, *Cj*PglB is known to require an aspartic or glutamic acid residue in the -2 position relative to the glycosylated asparagine, so that the glycosylation consensus sequence for this enzyme is D/E-x-N-x-S/T, where x is any amino acid except proline[96,151]. In contrast, acidic residues are disfavored in the -2 position of eukaryotic glycosylation sites[296]. As a result, the strict specificity of *Cj*PglB is considered a major challenge for the glycosylation of eukaryotic proteins by bacterial OSTs[297]. Recent studies identified multiple engineered variants and natural homologs of *Cj*PglB that exhibited relaxed amino acid sequence specificities[96,123]. Demonstrating that these more permissive OST variants could be used in the CFGpS platform would represent an important step toward the glycosylation of eukaryotic protein sequences. Importantly, we have already demonstrated that OSTs can be easily prototyped in CFGpS via mixing lysates enriched with LLOs and OSTs, so many candidate OSTs could be tested to determine those with the highest activities on eukaryotic protein sequences.

Additionally, though we showed that it was possible to glycosylate proteins with the eukaryotic trimannose core glycan, a majority of human *N*-linked glycoproteins bear glycans that are further elaborated with *N*-acetylglucosamine, galactose, and sialic acid monomers on each of the two branching mannose residues[298]. Thus, in order to use CFGpS to produce humanized glycoprotein therapeutics, glycosylation with a more complex eukaryotic *N*-glycan must be achieved. However, at present, the ability

to produce humanized *N*-glycans and attach them to proteins in *E. coli* remains an outstanding challenge in the field due to difficulties expressing eukaryotic glycosyltransferases in bacteria. Despite this limitation, recent work reported a synthetic enzyme pathway capable of elaborating the trimannose core glycan structure to produce lipid-linked *N*-glycans resembling human structures *in vitro*[299]. In addition, a separate study described minimal human-like glycan structures that recapitulate beneficial properties of the biantennary human *N*-glycan[80], some of which can be produced in *E. coli* using OST-independent enzymatic approaches capable of modifying eukaryotic glycosylation sites[300]. Integration of either of these advances with CFGpS technology represent promising directions towards enabling the production of proteins with human-like *N*-glycans. Together, the ability to attach human-like *N*-glycan structures to human proteins using the CFGpS platform would make it an attractive technology for on-demand production and prototyping of human glycoprotein therapeutics.

A second research area with exciting future directions is our work on bioconjugate vaccines, described in Chapter 4. While we showed that our iVAX platform produces effective bioconjugates, there is much work to do to continue to make the platform more effective and useful for on-demand production of antibacterial vaccines. For instance, there are a number of additional O-PS biosynthetic pathways from bacterial pathogens that have been previously shown to express functionally in *E. coli*. Demonstrating that these pathways are compatible with the iVAX platform would expand the diversity of bioconjugate vaccines that can be synthesized using our cell-free approach. This would involve demonstrating the ability to synthesize bioconjugates using all-in-one lysates from chassis strains expressing *Cj*PglB along with one of the following O-PS gene clusters from bacterial pathogens: (1) *B. pseudomallei* K96234[35]; (2) *B. mallei* ATCC 23344[301]; (3) *C. jejuni* 81116[302]; (4) ETEC *E. coli* strain O148[170]; (5) *Shigella dysenteriae* 1 strain W30864[303]; (6) *S. flexneri* serotype 2a[304]; and (7) *Yersinia enterocolitica* strain 6471/76[305]. Notably, we have many of these plasmids already in house, making this a tractable next step for this work.

A key finding from the initial iVAX study showed that cell-free-derived vaccine products were significantly more effective at eliciting *Ft*LPS-specific IgG than the state-of-the-art *in vivo*-derived conjugates. We hypothesize that this is a result of either enhanced carbohydrate per carrier protein, or

decoration with a more broad range of higher molecular weight O-PS species. Additionally, recent work has revealed that the site of O-PS attachment plays a key role in determining vaccine immunogenicity for a handful of protein carriers[306-309]. More studies are needed to fully understand the roles that carbohydrate loading and O-PS attachment site play in the immunogenicity of bioconjugates. The ability of the iVAX platform to rapidly synthesize bioconjugates with O-PS attached at defined glycosylation sites enables us to test the relationships between O-PS attachment site, glycan loading (through the production of molecules with single or multiple attachment sites), and immunogenicity at a resolution that was not possible before. Combined with recent developments in organoid technologies capable of mimicking antibody production in the germinal center[310], the iVAX platform could enable high-throughput and systematic interrogation of structure-immunogenicity relationships for licensed vaccine proteins, with the potential to enable the production of next generation bioconjugate vaccines with improved efficacies.

Finally, there are exciting opportunities to expand our work related to the BioBits™ kits. First, as the field of synthetic biology continues to develop new biological technologies, it will be necessary to convert these into teachable hands-on activities using the FD-CF reaction format described in Chapter 5. This is critical to both educate the next generation of scientists as well as to help students make informed policy decisions about how these technologies should be used and regulated. Beyond bench work, there is much work to be done in terms of implementing BioBits™ kits in classrooms and evaluating their educational impact. We began this work through our pilot program, which gathered data on the use of BioBits™ kits in 100 high school classrooms. As we learn how teachers use the kits in their classrooms, our supporting curriculum will need to continue to be developed and refined. Additionally, we plan to study students' educational outcomes from using the BioBits™ kits, which will involve careful design of pre- and post-assessments to quantitatively measure student learning. Ultimately, we will need to perform longitudinal studies to better understand the long-term outcomes that result from implementing hands-on biology curriculum. These types of studies will help us improve the BioBits™ resources and K-12 biology education in general.

I look forward to seeing how our CFGpS, iVAX, and BioBits^{TM} technologies are implemented

by others both inside and outside of the lab. I feel lucky to have been able to lead them this far.

References

1   CDC. *Epidemiology and Prevention of Vaccine-Preventable Diseases*. 13 edn, (2015).

2   FDA okays marketing of human insulin. *Chemical & Engineering News Archive* **60**, 5, doi:10.1021/cen-v060n045.p005 (1982).

3   Mullin, R. Cost to develop new pharmaceutical drug now exceeds $2.5 B. *Scientific American* **24** (2014).

4   Thiel, K. A. Biomanufacturing, from bust to boom...to bubble? *Nat. Biotechnol.* **22**, 1365, doi:10.1038/nbt1104-1365 (2004).

5   Jayapal, K. P., Wlaschin, K. F., Hu, W. & Yap, M. G. Recombinant protein therapeutics from CHO cells-20 years and counting. *Chem. Eng. Prog.* **103**, 40 (2007).

6   Dolsten, M. & Sogaard, M. Precision medicine: an approach to R&D for delivering superior medicines to patients. *Clin. Transl. Med.* **1**, 7, doi:10.1186/2001-1326-1-7 (2012).

7   Ashok, A., Brison, M. & LeTallec, Y. Improving cold chain systems: Challenges and solutions. *Vaccine* **35**, 2217-2223, doi:10.1016/j.vaccine.2016.08.045 (2017).

8   Kumru, O. S. *et al.* Vaccine instability in the cold chain: mechanisms, analysis and formulation strategies. *Biologicals* **42**, 237-259, doi:10.1016/j.biologicals.2014.05.007 (2014).

9   Choi, E. J. & Ling, G. S. Battlefield medicine: paradigm shift for pharmaceuticals manufacturing. *PDA J. Pharm. Sci. Technol.* **68**, 312, doi:10.5731/pdajpst.2014.01002 (2014).

10   Perez-Pinera, P. *et al.* Synthetic biology and microbioreactor platforms for programmable production of biologics at the point-of-care. *Nat. Commun.* **7**, 12211, doi:10.1038/ncomms12211 (2016).

11   Crowell, L. E. *et al.* On-demand manufacturing of clinical-quality biopharmaceuticals. *Nat. Biotechnol.*, doi:10.1038/nbt.4262 (2018).

12   Pardee, K. *et al.* Portable, on-demand biomolecular manufacturing. *Cell* **167**, 248-259.e212, doi:10.1016/j.cell.2016.09.013 (2016).

13   Salehi, A. S. *et al.* 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 (2016).

14   Adiga, R. *et al.* Point-of-care production of therapeutic proteins of good-manufacturing-practice quality. *Nat. Biomed. Eng.*, doi:10.1038/s41551-018-0259-1 (2018).

15   Galarneau, A., Primeau, M., Trudeau, L. E. & Michnick, S. W. Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nature biotechnology* **20**, 619-622, doi:10.1038/nbt0602-619 (2002).

16   Sola, R. J. & Griebenow, K. Glycosylation of therapeutic proteins: an effective strategy to optimize efficacy. *BioDrugs* **24**, 9-21, doi:10.2165/11530550-000000000-00000 (2010).

17   Sola, R. J. & Griebenow, K. Effects of glycosylation on the stability of protein pharmaceuticals. *J. Pharm. Sci.* **98**, 1223-1245, doi:10.1002/jps.21504 (2009).

18   Jefferis, R. Glycosylation as a strategy to improve antibody-based therapeutics. *Nat. Rev. Drug Discov.* **8**, 226-234, doi:10.1038/nrd2804 (2009).

19   Baker, J. L., Celik, E. & DeLisa, M. P. Expanding the glycoengineering toolbox: the rise of bacterial N-linked protein glycosylation. *Trends Biotechnol.* **31**, 313-323, doi:10.1016/j.tibtech.2013.03.003 (2013).

20   Merritt, J. H., Ollis, A. A., Fisher, A. C. & DeLisa, M. P. Glycans-by-design: engineering bacteria for the biosynthesis of complex glycans and glycoconjugates. *Biotechnology and bioengineering* **110**, 1550-1564, doi:10.1002/bit.24885 (2013).

21   Wacker, M. *et al.* N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli. *Science* **298**, 1790-1793, doi:10.1126/science.298.5599.1790 (2002).

22   Feldman, M. F. *et al.* Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 3016-3021, doi:10.1073/pnas.0500044102 (2005).

23   Valderrama-Rincon, J. D. *et al.* An engineered eukaryotic protein glycosylation pathway in Escherichia coli. *Nat. Chem. Biol.* **8**, 434-436, doi:10.1038/nchembio.921 (2012).

24   Rappuoli, R. Glycoconjugate vaccines: Principles and mechanisms. *Sci Transl Med* **10**, doi:10.1126/scitranslmed.aat4615 (2018).

25   Astronomo, R. D. & Burton, D. R. Carbohydrate vaccines: Developing sweet solutions to sticky situations? *Nat. Rev. Drug Discov.* **9**, 308-324, doi:10.1038/nrd3012 (2010).

26   Avci, F. Y. & Kasper, D. L. How bacterial carbohydrates influence the adaptive immune system. *Annu. Rev. Immunol.* **28**, 107-130, doi:10.1146/annurev-immunol-030409-101159 (2010).

27   Lockhart, S. Conjugate vaccines. *Expert Rev. Vaccines* **2**, 633-648 (2003).

28   Mitchison, N. T-cell–B-cell cooperation. *Nat. Rev. Immunol.* **4**, 308 (2004).

29   Weintraub, A. Immunology of bacterial polysaccharide antigens. *Carbohydr. Res.* **338**, 2539-2547 (2003).

30   Jones, C. Vaccines based on the cell surface carbohydrates of pathogenic bacteria. *An. Acad. Bras. Cienc.* **77**, 293-324, doi:/s0001-37652005000200009 (2005).

31   Trotter, C. L. *et al.* Optimising the use of conjugate vaccines to prevent disease caused by Haemophilus influenzae type b, Neisseria meningitidis and Streptococcus pneumoniae. *Vaccine* **26**, 4434-4445, doi:10.1016/j.vaccine.2008.05.073 (2008).

32   Frasch, C. E. Preparation of bacterial polysaccharide-protein conjugates: analytical and manufacturing challenges. *Vaccine* **27**, 6468-6470, doi:10.1016/j.vaccine.2009.06.013 (2009).

33   Wacker, M. *et al.* Prevention of Staphylococcus aureus infections by glycoprotein vaccines synthesized in Escherichia coli. *J. Infect. Dis.* **209**, 1551-1561, doi:10.1093/infdis/jit800 (2014).

34   Cuccui, J. *et al.* Exploitation of bacterial N-linked glycosylation to develop a novel recombinant glycoconjugate vaccine against Francisella tularensis. *Open Biol.* **3**, 130002, doi:10.1098/rsob.130002 (2013).

35   Garcia-Quintanilla, F., Iwashkiw, J. A., Price, N. L., Stratilo, C. & Feldman, M. F. Production of a recombinant vaccine candidate against Burkholderia pseudomallei exploiting the bacterial N-glycosylation machinery. *Front. Microbiol.* **5**, 381, doi:10.3389/fmicb.2014.00381 (2014).

36   Ma, Z. *et al.* Glycoconjugate vaccine containing Escherichia coli O157:H7 O-antigen linked with maltose-binding protein elicits humoral and cellular responses. *PloS one* **9**, e105215, doi:10.1371/journal.pone.0105215 (2014).

37   Wetter, M. *et al.* Engineering, conjugation, and immunogenicity assessment of Escherichia coli O121 O antigen for its potential use as a typhoid vaccine component. *Glycoconj. J.* **30**, 511-522, doi:10.1007/s10719-012-9451-9 (2013).

38   WHO. *Temperature Sensitivity of Vaccines*, <http://www.who.int/immunization/programmes_systems/supply_chain/resources/VaccineStability_EN.pdf> (2014).

39   Carlson, E. D., Gan, R., Hodgman, C. E. & Jewett, M. C. Cell-free protein synthesis: Applications come of age. *Biotechnol. Adv.* **30**, 1185-1194, doi:10.1016/j.biotechadv.2011.09.016 (2012).

40   Hodgman, C. E. & Jewett, M. C. Cell-free synthetic biology: Thinking outside the cell. *Metab. Eng.* **14**, 261-269, doi:10.1016/j.ymben.2011.09.002 (2012).

41   Nirenberg, M. W. & Matthaei, J. H. 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 (1961).

42   Chambers, D. A. & Zubay, G. The stimulatory effect of cyclic adenosine 3'5'-monophosphate on DNA-directed synthesis of beta-galactosidase in a cell-free system. *Proc. Natl. Acad. Sci. U. S. A.* **63**, 118-122 (1969).

43   Zalkin, H., Yanofsky, C. & Squires, C. L. Regulated in vitro synthesis of Escherichia coli tryptophan operon messenger ribonucleic acid and enzymes. *J. Biol. Chem.* **249**, 465-475 (1974).

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

45   Caschera, F. & Noireaux, V. 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 (2014).

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

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

48   Lian, Q., Cao, H. & Wang, F. The cost-efficiency realization in the Escherichia coli-based cell-free protein synthesis systems. *Appl. Biochem. Biotechnol.* **174**, 2351-2367, doi:10.1007/s12010-014-1143-4 (2014).

49   Zawada, J. F. *et al.* Microscale to manufacturing scale-up of cell-free cytokine production--a new approach for shortening protein production development timelines. *Biotechnology and bioengineering* **108**, 1570-1578, doi:10.1002/bit.23103 (2011).

50   Pardee, K. *et al.* Paper-based synthetic gene networks. *Cell* **159**, 940-954, doi:10.1016/j.cell.2014.10.004 (2014).

51   Pardee, K. *et al.* Rapid, low-cost detection of Zika virus using programmable biomolecular components. *Cell* **165**, 1255-1266, doi:10.1016/j.cell.2016.04.059 (2016).

52   Tarui, H., Imanishi, S. & Hara, T. A novel cell-free translation/glycosylation system prepared from insect cells. *J. Biosci. Bioeng.* **90**, 508-514 (2000).

53   Moreno, S. N., Ip, H. S. & Cross, G. A. An mRNA-dependent in vitro translation system from Trypanosoma brucei. *Mol. Biochem. Parasitol.* **46**, 265-274 (1991).

54   Mikami, S., Kobayashi, T., Yokoyama, S. & Imataka, H. A hybridoma-based in vitro translation system that efficiently synthesizes glycoproteins. *J. Biotechnol.* **127**, 65-78, doi:10.1016/j.jbiotec.2006.06.018 (2006).

55   Brodel, A. K. *et al.* IRES-mediated translation of membrane proteins and glycoproteins in eukaryotic cell-free systems. *PloS one* **8**, e82234, doi:10.1371/journal.pone.0082234 (2013).

56   Shibutani, M., Kim, E., Lazarovici, P., Oshima, M. & Guroff, G. Preparation of a cell-free translation system from PC12 cell. *Neurochem. Res.* **21**, 801-807 (1996).

57   Gurramkonda, C. *et al.* Improving the recombinant human erythropoietin glycosylation using microsome supplementation in CHO cell-free system. *Biotechnol. Bioeng.* **115**, 1253-1264, doi:10.1002/bit.26554 (2018).

58   Lingappa, V. R., Lingappa, J. R., Prasad, R., Ebner, K. E. & Blobel, G. Coupled cell-free synthesis, segregation, and core glycosylation of a secretory protein. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 2338-2342 (1978).

59   Rothblatt, J. A. & Meyer, D. I. Secretion in yeast: reconstitution of the translocation and glycosylation of alpha-factor and invertase in a homologous cell-free system. *Cell* **44**, 619-628 (1986).

60   Hamilton, S. R. *et al.* Production of complex human glycoproteins in yeast. *Science* **301**, 1244-1246, doi:10.1126/science.1088166 (2003).

61   Hamilton, S. R. *et al.* Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* **313**, 1441-1443, doi:10.1126/science.1130256 (2006).

62   Li, H. *et al.* Optimization of humanized IgGs in glycoengineered Pichia pastoris. *Nature biotechnology* **24**, 210-215, doi:10.1038/nbt1178 (2006).

63   Kontra, C., Lyons, D. J., Fischer, S. M. & Beilock, S. L. Physical experience enhances science learning. *Psychol. Sci.* **26**, 737-749, doi:10.1177/0956797615569355 (2015).

64   Francis, L. Y. *A high school biology teacher's development through a new teaching assignment coupled with teacher-led professional development* Master of Arts thesis, Brigham Young University, (2016).

65   Moore, A. Breathing new life into the biology classroom. *EMBO Rep.* **4**, 744-746, doi:10.1038/sj.embor.embor907 (2003).

66   Huang, A. *et al.* BioBits™ Explorer: A modular synthetic biology education kit. *Sci. Adv.* **4** (2018).

67   Baumgaertner, E. in *The New York Times*   (2018).

68   ECDC. *Risk related to the use of 'do-it-yourself' CRISPR-associated gene engineering kit contaminated with pathogenic bacteria.*, <https://ecdc.europa.eu/sites/portal/files/documents/2-May-2017-RRA_CRISPR-kit-w-pathogenic-bacteria_2.pdf> (2017).

69   Wood, W. B. Advanced high school biology in an era of rapid change: a summary of the biology panel report from the NRC Committee on Programs for Advanced Study of Mathematics and Science in American High Schools. *Cell Biol. Educ.* **1**, 123-127, doi:10.1187/cbe.02-09-0038 (2002).

70   Tickle, L. in *The Guardian*   (2018).

71   Hebert, D. N., Lamriben, L., Powers, E. T. & Kelly, J. W. The intrinsic and extrinsic effects of N-linked glycans on glycoproteostasis. *Nat. Chem. Biol.* **10**, 902-910, doi:10.1038/nchembio.1651 (2014).

72   Helenius, A. & Aebi, M. Intracellular functions of N-linked glycans. *Science* **291**, 2364-2369 (2001).

73   Imperiali, B. & O'Connor, S. E. Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. *Curr. Opin. Chem. Biol.* **3**, 643-649 (1999).

74   Rudd, P. M., Elliott, T., Cresswell, P., Wilson, I. A. & Dwek, R. A. Glycosylation and the immune system. *Science* **291**, 2370-2376 (2001).

75   Wolfert, M. A. & Boons, G. J. Adaptive immune activation: glycosylation does matter. *Nat. Chem. Biol.* **9**, 776-784, doi:10.1038/nchembio.1403 (2013).

76   Lanctot, P. M., Gage, F. H. & Varki, A. P. The glycans of stem cells. *Curr. Opin. Chem. Biol.* **11**, 373-380, doi:10.1016/j.cbpa.2007.05.032 (2007).

77   Sinclair, A. M. & Elliott, S. Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. *J. Pharm. Sci.* **94**, 1626-1635, doi:10.1002/jps.20319 (2005).

78   Raman, R., Raguram, S., Venkataraman, G., Paulson, J. C. & Sasisekharan, R. Glycomics: an integrated systems approach to structure-function relationships of glycans. *Nat. Methods* **2**, 817-824, doi:10.1038/nmeth807 (2005).

79   Rudd, P. M. & Dwek, R. A. Glycosylation: heterogeneity and the 3D structure of proteins. *Crit. Rev. Biochem. Mol. Biol.* **32**, 1-100, doi:10.3109/10409239709085144 (1997).

80   Meuris, L. *et al.* GlycoDelete engineering of mammalian cells simplifies N-glycosylation of recombinant proteins. *Nat. Biotechnol.* **32**, 485-489, doi:10.1038/nbt.2885 (2014).

81   Kiga, D. *et al.* An engineered Escherichia coli tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 9715-9720, doi:10.1073/pnas.142220099 (2002).

82   Kightlinger, W. *et al.* Design of glycosylation sites by rapid synthesis and analysis of glycosyltransferases. *Nat. Chem. Biol.* **14**, 627-635, doi:10.1038/s41589-018-0051-2 (2018).

83   Martin, R. W. *et al.* 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 (2018).

84   Oza, J. P. *et al.* Robust production of recombinant phosphoproteins using cell-free protein synthesis. *Nat. Commun.* **6**, 8168, doi:10.1038/ncomms9168 (2015).

85   Stapleton, J. A. & Swartz, J. R. 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 (2010).

86   Albayrak, C. & Swartz, J. R. Cell-free co-production of an orthogonal transfer RNA activates efficient site-specific non-natural amino acid incorporation. *Nucleic Acids Res.* **41**, 5949-5963, doi:10.1093/nar/gkt226 (2013).

87   Karim, A. S. & Jewett, M. C. 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 (2016).

88   Kaiser, L. *et al.* Efficient cell-free production of olfactory receptors: detergent optimization, structure, and ligand binding analyses. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15726-15731, doi:10.1073/pnas.0804766105 (2008).

89   Dudley, Q. M., Anderson, K. C. & Jewett, M. C. 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 (2016).

90   Moore, S. J. *et al.* 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 (2018).

91   Goshima, N. *et al.* Human protein factory for converting the transcriptome into an in vitro-expressed proteome. *Nat. Methods* **5**, 1011-1017 (2008).

92   Matsuoka, K., Komori, H., Nose, M., Endo, Y. & Sawasaki, T. Simple screening method for autoantigen proteins using the N-terminal biotinylated protein library produced by wheat cell-free synthesis. *J. Proteome. Res.* **9**, 4264-4273, doi:10.1021/pr9010553 (2010).

93   Guarino, C. & DeLisa, M. P. A prokaryote-based cell-free translation system that efficiently synthesizes glycoproteins. *Glycobiology* **22**, 596-601, doi:10.1093/glycob/cwr151 (2012).

94   Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006 0008 (2006).

95   Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640-6645, doi:10.1073/pnas.120163297 (2000).

96   Ollis, A. A. *et al.* Substitute sweeteners: Diverse bacterial oligosaccharyltransferases with unique N-glycosylation site preferences. *Sci. Rep.* **5**, 15237, doi:10.1038/srep15237 (2015).

97   Schoborg, J. A. *et al.* A cell-free platform for rapid synthesis and testing of active oligosaccharyltransferases. *Biotechnol. Bioeng.*, doi:10.1002/bit.26502 (2017).

98   Bundy, B. C. & Swartz, J. R. Site-specific incorporation of p-propargyloxyphenylalanine in a cell-free environment for direct protein-protein click conjugation. *Bioconjug. Chem.* **21**, 255-263, doi:10.1021/bc9002844 (2010).

99   Kowarik, M. *et al.* N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* **314**, 1148-1150, doi:10.1126/science.1134351 (2006).

100   Jaroentomeechai, T. *et al.* A pipeline for studying and engineering single-subunit oligosaccharyltransferases. *Methods Enzymol.* **597**, 55-81, doi:10.1016/bs.mie.2017.07.011 (2017).

101   Jewett, M. C. & Swartz, J. R. Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.* **86**, 19-26, doi:10.1002/bit.20026 (2004).

102   Kwon, Y. C. & Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* **5**, 8663, doi:10.1038/srep08663 (2015).

103   Weerapana, E. & Imperiali, B. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* **16**, 91R-101R, doi:10.1093/glycob/cwj099 (2006).

104   Liu, D. & Reeves, P. R. Escherichia coli K12 regains its O antigen. *Microbiology* **140 (Pt 1)**, 49-57, doi:10.1099/13500872-140-1-49 (1994).

105   Chen, L. *et al.* Outer membrane vesicles displaying engineered glycotopes elicit protective antibodies. *Proc. Natl. Acad. Sci. U. S. A.*, doi:10.1073/pnas.1518311113 (2016).

106   Schwarz, F. *et al.* Relaxed acceptor site specificity of bacterial oligosaccharyltransferase in vivo. *Glycobiology* **21**, 45-54, doi:10.1093/glycob/cwq130 (2011).

107   Schwarz, F. *et al.* A combined method for producing homogeneous glycoproteins with eukaryotic N-glycosylation. *Nat. Chem. Biol.* **6**, 264-266, doi:10.1038/nchembio.314 (2010).

108   Jervis, A. J. *et al.* Characterization of the structurally diverse N-linked glycans of Campylobacter species. *J. Bacteriol.* **194**, 2355-2362, doi:10.1128/jb.00042-12 (2012).

109   Hagelueken, G. *et al.* A coiled-coil domain acts as a molecular ruler to regulate O-antigen chain length in lipopolysaccharide. *Nat. Struct. Mol. Biol.* **22**, 50-56, doi:10.1038/nsmb.2935 (2015).

110   Srichaisupakit, A., Ohashi, T., Misaki, R. & Fujiyama, K. Production of initial-stage eukaryotic N-glycan and its protein glycosylation in Escherichia coli. *J. Biosci. Bioeng.* **119**, 399-405, doi:10.1016/j.jbiosc.2014.09.016 (2015).

111   Jewett, M. C., Calhoun, K. A., Voloshin, A., Wuu, J. J. & Swartz, J. R. An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol.* **4**, 220, doi:10.1038/msb.2008.57 (2008).

112   Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).

113   Jackson, K., Kanamori, T., Ueda, T. & Fan, Z. H. Protein synthesis yield increased 72 times in the cell-free PURE system. *Integr. Biol.* **6**, 781-788, doi:10.1039/c4ib00088a (2014).

114   Stech, M. *et al.* 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 (2017).

115   Ahn, J. H., Hwang, M. Y., Lee, K. H., Choi, C. Y. & Kim, D. M. Use of signal sequences as an in situ removable sequence element to stimulate protein synthesis in cell-free extracts. *Nucleic Acids Res.* **35**, e21, doi:10.1093/nar/gkl917 (2007).

116   Sun, Z. Z. *et al.* Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. *Journal of visualized experiments : JoVE*, e50762, doi:10.3791/50762 (2013).

117   Hayes, C. *Biomolecular Breadboards:Protocols:cost estimate.*, <http://www.openwetware.org/wiki/Biomolecular_Breadboards:Protocols:cost_estimate> (2012).

118   Glover, K. J., Weerapana, E., Numao, S. & Imperiali, B. Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from Campylobacter jejuni. *Chemistry & biology* **12**, 1311-1315, doi:10.1016/j.chembiol.2005.10.004 (2005).

119   Lizak, C., Gerber, S., Numao, S., Aebi, M. & Locher, K. P. X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474**, 350-355, doi:10.1038/nature10151 (2011).

120   Musumeci, M. A. *et al.* In vitro activity of Neisseria meningitidis PglL O-oligosaccharyltransferase with diverse synthetic lipid donors and a UDP-activated sugar. *The Journal of biological chemistry* **288**, 10578-10587, doi:10.1074/jbc.M112.432815 (2013).

121   Ramirez, A. S. *et al.* Characterization of the single-subunit oligosaccharyltransferase STT3A from Trypanosoma brucei using synthetic peptides andlipid-linked oligosaccharide analogs. *Glycobiology*, doi:10.1093/glycob/cwx017 (2017).

122   Sullivan, C. J. *et al.* A cell-free expression and purification process for rapid production of protein biologics. *Biotechnology journal* **11**, 238-248, doi:10.1002/biot.201500214 (2016).

123   Ollis, A. A., Zhang, S., Fisher, A. C. & DeLisa, M. P. Engineered oligosaccharyltransferases with greatly relaxed acceptor-site specificity. *Nat. Chem. Biol.* **10**, 816-822, doi:10.1038/nchembio.1609 (2014).

124   The Review on Antimicrobial Resistance, C. b. J. O. N. Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. (2014).

125   Jin, C. *et al.* Efficacy and immunogenicity of a Vi-tetanus toxoid conjugate vaccine in the prevention of typhoid fever using a controlled human infection model of Salmonella Typhi: a randomised controlled, phase 2b trial. *Lancet (London, England)* **390**, 2472-2480, doi:10.1016/s0140-6736(17)32149-9 (2017).

126   Novak, R. T. *et al.* Serogroup A meningococcal conjugate vaccination in Burkina Faso: analysis of national surveillance data. *Lancet Infect. Dis.* **12**, 757-764, doi:10.1016/s1473-3099(12)70168-8 (2012).

127   Poehling, K. A. *et al.* Invasive pneumococcal disease among infants before and after introduction of pneumococcal conjugate vaccine. *Jama* **295**, 1668-1674, doi:10.1001/jama.295.14.1668 (2006).

128   Roush, S. W., McIntyre, L. & Baldy, L. M. Manual for the surveillance of vaccine-preventable diseases. *Atlanta: Centers for Disease Control and Prevention*, 4 (2008).

129   Wahl, B. *et al.* Burden of Streptococcus pneumoniae and Haemophilus influenzae type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000-15. *Lancet Glob. Health* **6**, e744-e757, doi:10.1016/s2214-109x(18)30247-x (2018).

130   Bhushan, R., Anthony, B. F. & Frasch, C. E. Estimation of group B streptococcus type III polysaccharide-specific antibody concentrations in human sera is antigen dependent. *Infect. Immun.* **66**, 5848-5853 (1998).

131   Marshall, L. E. *et al.* An O-antigen glycoconjugate vaccine produced using protein glycan coupling technology is protective in an inhalational rat model of tularemia. *J. Immunol. Res.* **2018**, 8087916, doi:10.1155/2018/8087916 (2018).

132   Ihssen, J. *et al.* Production of glycoprotein vaccines in Escherichia coli. *Microb. Cell. Fact.* **9**, 61, doi:10.1186/1475-2859-9-61 (2010).

133   Lydon, P. *et al.* Economic benefits of keeping vaccines at ambient temperature during mass vaccination: the case of meningitis A vaccine in Chad. *Bull. World Health Organ.* **92**, 86-92, doi:10.2471/blt.13.123471 (2014).

134   Jaroentomeechai, T. *et al.* Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nat. Commun.* **9**, 2686, doi:10.1038/s41467-018-05110-x (2018).

135   Fisher, A. C. *et al.* Production of secretory and extracellular N-linked glycoproteins in Escherichia coli. *Applied and environmental microbiology* **77**, 871-881, doi:10.1128/aem.01901-10 (2011).

136   Needham, B. D. *et al.* Modulating the innate immune response by combinatorial engineering of endotoxin. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1464-1469, doi:10.1073/pnas.1218080110 (2013).

137   Knapp, K. G., Goerke, A. R. & Swartz, J. R. Cell-free synthesis of proteins that require disulfide bonds using glucose as an energy source. *Biotechnol. Bioeng.* **97**, 901-908, doi:10.1002/bit.21296 (2007).

138   Hong, S. H. *et al.* 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 (2014).

139   Kim, D. M. & Swartz, J. R. Regeneration of adenosine triphosphate from glycolytic intermediates for cell-free protein synthesis. *Biotechnol. Bioeng.* **74**, 309-316 (2001).

140   Chen, D. J. *et al.* Delivery of foreign antigens by engineered outer membrane vesicle vaccines. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3099-3104, doi:10.1073/pnas.0805532107 (2010).

141   Haghi, F., Peerayeh, S. N., Siadat, S. D. & Montajabiniat, M. Cloning, expression and purification of outer membrane protein PorA of Neisseria meningitidis serogroup B. *J. Infect. Dev. Ctries.* **5**, 856-862 (2011).

142   Stefan, A. *et al.* Overexpression and purification of the recombinant diphtheria toxin variant CRM197 in Escherichia coli. *J. Biotechnol.* **156**, 245-252, doi:10.1016/j.jbiotec.2011.08.024 (2011).

143   Figueiredo, D. *et al.* Characterization of recombinant tetanus toxin derivatives suitable for vaccine development. *Infect. Immun.* **63**, 3218-3221 (1995).

144   Fernandez, S. *et al.* Potential role for Toll-like receptor 4 in mediating Escherichia coli maltose-binding protein activation of dendritic cells. *Infect. Immun.* **75**, 1359-1363, doi:10.1128/iai.00486-06 (2007).

145   Chen, M. M., Glover, K. J. & Imperiali, B. From peptide to protein: comparative analysis of the substrate specificity of N-linked glycosylation in C. jejuni. *Biochemistry* **46**, 5579-5585, doi:10.1021/bi602633n (2007).

146   Bayburt, T. H. & Sligar, S. G. Membrane protein assembly into Nanodiscs. *FEBS Lett.* **584**, 1721-1727, doi:10.1016/j.febslet.2009.10.024 (2010).

147   Oyston, P. C., Sjostedt, A. & Titball, R. W. Tularaemia: bioterrorism defence renews interest in Francisella tularensis. *Nat. Rev. Microbiol.* **2**, 967-978, doi:10.1038/nrmicro1045 (2004).

148   Fulop, M., Mastroeni, P., Green, M. & Titball, R. W. Role of antibody to lipopolysaccharide in protection against low- and high-virulence strains of Francisella tularensis. *Vaccine* **19**, 4465-4472 (2001).

149   Lu, Z. *et al.* Protective B-cell epitopes of Francisella tularensis O-polysaccharide in a mouse model of respiratory tularaemia. *Immunology* **136**, 352-360, doi:10.1111/j.1365-2567.2012.03589.x (2012).

150   Prior, J. L. *et al.* Characterization of the O antigen gene cluster and structural analysis of the O antigen of Francisella tularensis subsp. tularensis. *J. Med. Microbiol.* **52**, 845-851, doi:10.1099/jmm.0.05184-0 (2003).

151   Kowarik, M. *et al.* Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J.* **25**, 1957-1966, doi:10.1038/sj.emboj.7601087 (2006).

152   Stefanetti, G., Okan, N., Fink, A., Gardner, E. & Kasper, D. L. Glycoconjugate vaccine using a genetically modified O antigen induces protective antibodies to Francisella tularensis. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 7062-7070, doi:10.1073/pnas.1900144116 (2019).

153   Hatz, C. F. *et al.* Safety and immunogenicity of a candidate bioconjugate vaccine against Shigella dysenteriae type 1 administered to healthy adults: A single blind, partially randomized Phase I study. *Vaccine* **33**, 4594-4601, doi:10.1016/j.vaccine.2015.06.102 (2015).

154   Huttner, A. *et al.* Safety, immunogenicity, and preliminary clinical efficacy of a vaccine against extraintestinal pathogenic Escherichia coli in women with a history of recurrent urinary tract infection: a randomised, single-blind, placebo-controlled phase 1b trial. *Lancet Infect. Dis.*, doi:10.1016/s1473-3099(17)30108-1 (2017).

155   Riddle, M. S. *et al.* Safety and immunogenicity of a candidate bioconjugate vaccine against Shigella flexneri 2a administered to healthy adults: a single blind, randomized phase I study. *Clin. Vaccine Immunol.*, doi:10.1128/cvi.00224-16 (2016).

156   Humphreys, G.   (World Health Organization, Bulletin of the World Health Organization, 2011).

157   Qadri, F., Svennerholm, A. M., Faruque, A. S. & Sack, R. B. Enterotoxigenic Escherichia coli in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clin. Microbiol. Rev.* **18**, 465-483, doi:10.1128/cmr.18.3.465-483.2005 (2005).

158   Johnson, J. R. Virulence factors in Escherichia coli urinary tract infection. *Clin. Microbiol. Rev.* **4**, 80-128 (1991).

159   Jansson, P. E., Lindberg, B., Widmalm, G. & Leontein, K. Structural studies of the Escherichia coli O78 O-antigen polysaccharide. *Carbohydr. Res.* **165**, 87-92 (1987).

160   L'vov, V. L. *et al.* Structural studies of the O-specific side chain of the lipopolysaccharide from Escherichia coli O:7. *Carbohydrate research* **126**, 249-259 (1984).

161   Russell, J. A. Management of sepsis. *N. Engl. J. Med.* **355**, 1699-1713 (2006).

162   Brito, L. A. & Singh, M. Acceptable levels of endotoxin in vaccine formulations during preclinical research. *J. Pharm. Sci.* **100**, 34-37, doi:10.1002/jps.22267 (2011).

163   Petsch, D. & Anspach, F. B. Endotoxin removal from protein solutions. *J. Biotechnol.* **76**, 97-119 (2000).

164   Casella, C. R. & Mitchell, T. C. Putting endotoxin to work for us: monophosphoryl lipid A as a safe and effective vaccine adjuvant. *Cell. Mol. Life Sci.* **65**, 3231-3240, doi:10.1007/s00018-008-8228-6 (2008).

165   Bogaert, D., Hermans, P. W., Adrian, P. V., Rumke, H. C. & de Groot, R. Pneumococcal vaccines: an update on current strategies. *Vaccine* **22**, 2209-2220, doi:10.1016/j.vaccine.2003.11.038 (2004).

166   Raetz, C. R. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635-700, doi:10.1146/annurev.biochem.71.110601.135414 (2002).

167   CDC. CDC Vaccine Price List. (2019).

168   Murphy, T. W., Sheng, J., Naler, L. B., Feng, X. & Lu, C. On-chip manufacturing of synthetic proteins for point-of-care therapeutics. *Microsyst. Nanoeng.* **5**, 13, doi:10.1038/s41378-019-0051-8 (2019).

169   Boles, K. S. *et al.* Digital-to-biological converter for on-demand production of biologics. *Nat. Biotechnol.* **35**, 672-675, doi:10.1038/nbt.3859 (2017).

170   Celik, E. *et al.* Glycoarrays with engineered phages displaying structurally diverse oligosaccharides enable high-throughput detection of glycan-protein interactions. *Biotechnol. J.* **10**, 199-209, doi:10.1002/biot.201400354 (2015).

171   Valvano, M. A. & Crosa, J. H. Molecular cloning and expression in Escherichia coli K-12 of chromosomal genes determining the O7 lipopolysaccharide antigen of a human invasive strain of E. coli O7:K1. *Infect. Immun.* **57**, 937-943 (1989).

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

173   Ryabova, L. A., Vinokurov, L. M., Shekhovtsova, E. A., Alakhov, Y. B. & Spirin, A. S. Acetyl phosphate as an energy source for bacterial cell-free translation systems. *Anal. Biochem.* **226**, 184-186, doi:10.1006/abio.1995.1208 (1995).

174   Takai, K., Sawasaki, T. & Endo, Y. Practical cell-free protein synthesis system using purified wheat embryos. *Nat. Protoc.* **5**, 227-238, doi:10.1038/nprot.2009.207 (2010).

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

176   Schoborg, J. A., Hodgman, C. E., Anderson, M. J. & Jewett, M. C. Substrate replenishment and byproduct removal improve yeast cell-free protein synthesis. *Biotechnol. J.* **9**, 630-640, doi:10.1002/biot.201300383 (2014).

177   Swartz, J. Developing cell-free biology for industrial applications. *J. Ind. Microbiol. Biotechnol.* **33**, 476-485, doi:10.1007/s10295-006-0127-y (2006).

178   Kim, T. W. *et al.* Prolonged cell-free protein synthesis using dual energy sources: Combined use of creatine phosphate and glucose for the efficient supply of ATP and retarded accumulation of phosphate. *Biotechnol. Bioeng.* **97**, 1510-1515, doi:10.1002/bit.21337 (2007).

179   Kim, H.-C., Kim, T.-W. & Kim, D.-M. Prolonged production of proteins in a cell-free protein synthesis system using polymeric carbohydrates as an energy source. *Process Biochem.* **46**, 1366-1369 (2011).

180   Caschera, F. & Noireaux, V. 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 (2015).

181   Wang, Y. & Zhang, Y. H. Cell-free protein synthesis energized by slowly-metabolized maltodextrin. *BMC Biotechnol.* **9**, 58, doi:10.1186/1472-6750-9-58 (2009).

182   Bujara, M., Schumperli, M., Billerbeck, S., Heinemann, M. & Panke, S. Exploiting cell-free systems: Implementation and debugging of a system of biotransformations. *Biotechnol. Bioeng.* **106**, 376-389, doi:10.1002/bit.22666 (2010).

183   Chappell, J., Takahashi, M. K. & Lucks, J. B. Creating small transcription activating RNAs. *Nat. Chem. Biol.* **11**, 214-220, doi:10.1038/nchembio.1737 (2015).

184   Karig, D. K., Iyer, S., Simpson, M. L. & Doktycz, M. J. Expression optimization and synthetic gene networks in cell-free systems. *Nucleic Acids Res.* **40**, 3763-3774, doi:10.1093/nar/gkr1191 (2012).

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

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

187   Takahashi, M. K. *et al.* Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. *ACS Synth. Biol.*, doi:10.1021/sb400206c (2014).

188   Yin, G. *et al.* 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 (2012).

189   Choudhury, A., Hodgman, C. E., Anderson, M. J. & Jewett, M. C. Evaluating fermentation effects on cell growth and crude extract metabolic activity for improved yeast cell-free protein synthesis. *Biochemical Engineering Journal* **91**, 140-148 (2014).

190   Gan, R. & Jewett, M. C. A combined cell-free transcription-translation system from Saccharomyces cerevisiae for rapid and robust protein synthesis. *Biotechnology journal* **9**, 641-651, doi:10.1002/biot.201300545 (2014).

191   Fritz, B. R., Jamil, O. K. & Jewett, M. C. Implications of macromolecular crowding and reducing conditions for in vitro ribosome construction. *Nucleic acids research*, doi:10.1093/nar/gkv329 (2015).

192   Fritz, B. R. & Jewett, M. C. The impact of transcriptional tuning on in vitro integrated rRNA transcription and ribosome construction. *Nucleic acids research* **42**, 6774-6785, doi:10.1093/nar/gku307 (2014).

193   Jewett, M. C., Fritz, B. R., Timmerman, L. E. & Church, G. M. In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation. *Mol Syst Biol* **9**, 678, doi:10.1038/msb.2013.31 (2013).

194   Jewett, M. C. & Swartz, J. R. Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. *Biotechnol. Bioeng.* **87**, 465-472, doi:10.1002/bit.20139 (2004).

195   Buchner, E. & Rapp, R. Alkoholische gährung ohne hefezellen. *Berichte der deutschen chemischen Gesellschaft* **30**, 2668-2678 (1897).

196   Khattak, W. A. *et al.* Yeast cell-free enzyme system for bio-ethanol production at elevated temperatures. *Process Biochemistry* **49**, 357-364 (2014).

197   Atkinson, D. E. The energy charge of the adenylate pool as a regulatory parameter. Interaction with feedback modifiers. *Biochemistry* **7**, 4030-4034 (1968).

198   Chapman, A. G., Fall, L. & Atkinson, D. E. Adenylate energy charge in Escherichia coli during growth and starvation. *J Bacteriol* **108**, 1072-1086 (1971).

199   Hong, S. H. *et al.* Improving cell-free protein synthesis through genome engineering of Escherichia coli lacking release factor 1. *Chembiochem* **16**, 844-853, doi:10.1002/cbic.201402708 (2015).

200   Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science* **349**, 1095-1100, doi:10.1126/science.aac9373 (2015).

201   Paddon, C. J. *et al.* High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528-532, doi:10.1038/nature12051 (2013).

202   Nielsen, J. & Keasling, J. D. Engineering cellular metabolism. *Cell* **164**, 1185-1197, doi:10.1016/j.cell.2016.02.004 (2016).

203   Yim, H. *et al.* Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. *Nat. Chem. Biol.* **7**, 445-452, doi:10.1038/nchembio.580 (2011).

204   Liu, C., Colon, B. C., Ziesack, M., Silver, P. A. & Nocera, D. G. Water splitting-biosynthetic system with CO(2) reduction efficiencies exceeding photosynthesis. *Science* **352**, 1210-1213, doi:10.1126/science.aaf5039 (2016).

205   Danino, T. *et al.* Programmable probiotics for detection of cancer in urine. *Sci. Transl. Med.* **7**, 289ra284, doi:10.1126/scitranslmed.aaa3519 (2015).

206   Porter, D. L., Levine, B. L., Kalos, M., Bagg, A. & June, C. H. Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. *N. Engl. J. Med.* **365**, 725-733 (2011).

207   Riglar, D. T. *et al.* Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat. Biotechnol.* **35**, 653-658, doi:10.1038/nbt.3879 (2017).

208   Roybal, K. T. *et al.* Engineering T cells with customized therapeutic response programs using synthetic notch receptors. *Cell* **167**, 419-432.e416, doi:10.1016/j.cell.2016.09.011 (2016).

209   NAE. *Educating the engineer of 2020: Adapting engineering education to the new century.* (National Academies Press, 2005).

210   Chandran, D., Bergmann, F. T. & Sauro, H. M. TinkerCell: Modular CAD tool for synthetic biology. *J. Biol. Eng.* **3**, 19, doi:10.1186/1754-1611-3-19 (2009).

211   Dixon, J. & Kuldell, N. BioBuilding using banana-scented bacteria to teach synthetic biology. *Methods Enzymol.* **497**, 255-271, doi:10.1016/b978-0-12-385075-1.00012-3 (2011).

212   Dixon, J. & Kuldell, N. Mendel's modern legacy: How to incorporate engineering in the biology classroom. *Sci. Teach.* **79**, 52-57 (2012).

213   Levskaya, A. *et al.* Synthetic biology: Engineering Escherichia coli to see light. *Nature* **438**, 441-442, doi:10.1038/nature04405 (2005).

214   International Genetically Engineered Machine (iGEM) Foundation. igem.org (2017).

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

216   Campbell, A. M. Meeting report: Synthetic biology jamboree for undergraduates. *Cell Biol. Educ.* **4**, 19-23, doi:10.1187/cbe.04-11-0047 (2005).

217   Mitchell, R., Dori, Y. J., Kuldell, N. H. Experiential engineering through iGEM—An undergraduate summer competition in synthetic biology. *J. Sci. Educ. Technol.* **20**, 156–160, doi:10.1007/s10956-010-9242-7 (2011).

218   Mitchell, R. & Kuldell, N. in *Inquiry-based learning for science, technology, engineering, and math (STEM) programs: A conceptual and practical resource for educators*   75-92 (2015).

219   Kuldell, N. Authentic teaching and learning through synthetic biology. *J. Biol. Eng.* **1**, 8, doi:10.1186/1754-1611-1-8 (2007).

220   Gootenberg, J. S. *et al.* Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* **356**, 438-442, doi:10.1126/science.aam9321 (2017).

221   Smith, M. T., Berkheimer, S. D., Werner, C. J. & Bundy, B. C. Lyophilized Escherichia coli-based cell-free systems for robust, high-density, long-term storage. *BioTechniques* **56**, 186-193, doi:10.2144/000114158 (2014).

222   Ai, H. W., Henderson, J. N., Remington, S. J. & Campbell, R. E. Directed evolution of a monomeric, bright and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochem. J.* **400**, 531-540, doi:10.1042/bj20060874 (2006).

223   Giepmans, B. N., Adams, S. R., Ellisman, M. H. & Tsien, R. Y. The fluorescent toolbox for assessing protein location and function. *Science* **312**, 217-224, doi:10.1126/science.1124618 (2006).

224   Heim, R. & Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* **6**, 178-182 (1996).

225   Nguyen, A. W. & Daugherty, P. S. Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol.* **23**, 355-360, doi:10.1038/nbt1066 (2005).

226   Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.* **22**, 1567-1572, doi:10.1038/nbt1037 (2004).

227   Subach, O. M., Cranfill, P. J., Davidson, M. W. & Verkhusha, V. V. An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. *PloS one* **6**, e28674, doi:10.1371/journal.pone.0028674 (2011).

228   High school life sciences standards. *Next Generation Science Standards*, nextgenscience.org/sites/default/files/HS%20LS%20topics%20combined%206.213.213.pdf (2013).

229   3-LS4-4 Biological Evolution: Unity and Diversity. *Next Generation Science Standards*, nextgenscience.org/pe/3-ls4-4-biological-evolution-unity-and-diversity (2017).

230   Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79-88, doi:10.1038/nbt1172 (2006).

231   Albayrak, C., Jones, K. C. & Swartz, J. R. Broadening horizons and teaching basic biology through cell-free synthesis of green fluorescent protein in a high school laboratory course. *J. Sci. Educ. Technol.* **22**, 963-973 (2013).

232   Peppler, K. STEAM-powered computing education: Using e-textiles to integrate the arts and STEM. *Computer*, 1 (2013).

233   Anderson, M. *et al.* The production of human leptin using cell-free protein synthesis. *BioTreks* **2**, e201705 (2017).

234   Murray, R. M. Build a DNA circuit. txtlbio.org/waitwhat/ (2017).

235   Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).

236   Steiner, H. & Meany, T. bi.xels. designbio.co.uk/blog/bi-xels (2017).

237   Benitti, F. B. V. Exploring the educational potential of robotics in schools: A systematic review. *Comput. Educ.* **58**, 978-988 (2012).

238   Clark, D. B., Tanner-Smith, E. E. & Killingsworth, S. S. Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Rev. Educ. Res.* **86**, 79-122, doi:10.3102/0034654315582065 (2016).

239   Resnick, M. *et al.* Scratch: programming for all. *Commun. ACM* **52**, 60-67 (2009).

240   Council, N. R.   (The National Academies Press Washington, DC, 2013).

241   Freeman, S. *et al.* 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 (2014).

242   Medema, M. H., Breitling, R., Bovenberg, R. & Takano, E. Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms. *Nat. Rev. Microbiol.* **9**, 131-137, doi:10.1038/nrmicro2478 (2011).

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

244   Smanski, M. J. *et al.* Synthetic biology to access and expand nature's chemical diversity. *Nat. Rev. Microbiol.* **14**, 135-149, doi:10.1038/nrmicro.2015.24 (2016).

245   Fesnak, A. D., June, C. H. & Levine, B. L. Engineered T cells: the promise and challenges of cancer immunotherapy. *Nat. Rev. Cancer* **16**, 566-581, doi:10.1038/nrc.2016.97 (2016).

246   Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**, 751-755, doi:10.1038/90802 (2001).

247   Lynch, S. J. *et al.* Understanding inclusive STEM high schools as opportunity structures for underrepresented students: Critical components. *Journal of Research in Science Teaching* **55**, 712-748 (2018).

248   Stark, J. C. *et al.* BioBits™ Bright: A fluorescent synthetic biology education kit. *Sci. Adv.* **4** (2018).

249   Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* **6**, 343-345, doi:10.1038/nmeth.1318 (2009).

250   Kunjapur, A. M., Hyun, J. C. & Prather, K. L. Deregulation of S-adenosylmethionine biosynthesis and regeneration improves methylation in the E. coli de novo vanillin biosynthesis pathway. *Microb. Cell Fact.* **15**, 61, doi:10.1186/s12934-016-0459-x (2016).

251   Bicchi, C., Iori, C., Rubiolo, P. & Sandra, P. Headspace sorptive extraction (HSSE), stir bar sorptive extraction (SBSE), and solid phase microextraction (SPME) applied to the analysis of roasted Arabica coffee and coffee brew. *J. Agric. Food Chem.* **50**, 449-459 (2002).

252   Alieva, N. O. *et al.* Diversity and evolution of coral fluorescent proteins. *PloS one* **3**, e2680, doi:10.1371/journal.pone.0002680 (2008).

253   Bayle, V., Nussaume, L. & Bhat, R. A. Combination of novel green fluorescent protein mutant TSapphire and DsRed variant mOrange to set up a versatile in planta FRET-FLIM assay. *Plant Physiol.* **148**, 51-60, doi:10.1104/pp.108.117358 (2008).

254   Buren, E. B. *et al.* Toolkit for visualization of the cellular structure and organelles in Aspergillus niger. *ACS Synth. Biol.* **3**, 995-998, doi:10.1021/sb500304m (2014).

255   Campbell, R. E. *et al.* A monomeric red fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 7877-7882, doi:10.1073/pnas.082243699 (2002).

256   Erard, M. *et al.* Minimum set of mutations needed to optimize cyan fluorescent proteins for live cell imaging. *Mol. Biosyst.* **9**, 258-267, doi:10.1039/c2mb25303h (2013).

257   Pham, H. L., Ho, C. L., Wong, A., Lee, Y. S. & Chang, M. W. Applying the design-build-test paradigm in microbiome engineering. *Curr. Opin. Biotechnol.* **48**, 85-93, doi:10.1016/j.copbio.2017.03.021 (2017).

258   Association, A. I. H. *Odor thresholds for chemicals with established occupational health standards*. (Aiha, 1989).

259   Yoshioka, K. & Hashimoto, N. Ester formation by alcohol acetyltransferase from brewers' yeast. *Agric. Biol. Chem.* **45**, 2183-2190 (1981).

260   Mandal, S. M., Chakraborty, D. & Dey, S. Phenolic acids act as signaling molecules in plant-microbe symbioses. *Plant Signal. Behav.* **5**, 359-368 (2010).

261   Wackett, L. P. Microbiology for odour production and abatement: An annotated selection of World Wide Web sites relevant to the topics in Microbial Biotechnology. *Microb. Biotechnol.* **6**, 85-86 (2013).

262   Drury, J. L. & Mooney, D. J. Hydrogels for tissue engineering: scaffold design variables and applications. *Biomaterials* **24**, 4337-4351 (2003).

263   Peppas, N. A., Hilt, J. Z., Khademhosseini, A. & Langer, R. Hydrogels in biology and medicine: from molecular principles to bionanotechnology. *Adv. Mater.* **18**, 1345-1360 (2006).

264   Valdez, J. *et al.* On-demand dissolution of modular, synthetic extracellular matrix reveals local epithelial-stromal communication networks. *Biomaterials* **130**, 90-103, doi:10.1016/j.biomaterials.2017.03.030 (2017).

265   Davie, E. W. & Ratnoff, O. D. Waterfall sequence for intrinsic blood clotting. *Science* **145**, 1310-1312 (1964).

266   Dyr, J. E., Blomback, B. & Kornalik, F. The action of prothrombin activated by Ecarin on fibrinogen. *Thromb. Res.* **30**, 225-234 (1983).

267   Wedgwood, J., Freemont, A. J. & Tirelli, N. in *Macromol. Symp.* 117-125 (Wiley Online Library).

268   Partlow, B. P. *et al.* Highly tunable elastomeric silk biomaterials. *Adv. Funct. Mater.* **24**, 4615-4624, doi:10.1002/adfm.201400526 (2014).

269   Green, A. A., Silver, P. A., Collins, J. J. & Yin, P. Toehold switches: de-novo-designed regulators of gene expression. *Cell* **159**, 925-939, doi:10.1016/j.cell.2014.10.002 (2014).

270   Christelova, P., Valarik, M., Hribova, E., De Langhe, E. & Dolezel, J. A multi gene sequence-based phylogeny of the Musaceae (banana) family. *BMC Evol. Biol.* **11**, 103, doi:10.1186/1471-2148-11-103 (2011).

271   Taguchi, H. *et al.* Specific detection of potentially allergenic kiwifruit in foods using polymerase chain reaction. *J. Agric. Food Chem.* **55**, 1649-1655, doi:10.1021/jf0624446 (2007).

272   Piepenburg, O., Williams, C. H., Stemple, D. L. & Armes, N. A. DNA detection using recombination proteins. *PLoS Biology* **4**, e204 (2006).

273   Sweeney, D. & Williamson, B. *Biology: Exploring Life: Laboratory Manual.* (Pearson Education, Incorporated, 2006).

274   Shepherd, T. R. *et al.* De novo design and synthesis of a 30-cistron translation-factor module. *Nucleic Acids Res.* **45**, 10895-10905, doi:10.1093/nar/gkx753 (2017).

275   Villarreal, F. *et al.* Synthetic microbial consortia enable rapid assembly of pure translation machinery. *Nat. Chem. Biol.* **14**, 29-35, doi:10.1038/nchembio.2514 (2018).

276   Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823, doi:10.1126/science.1231143 (2013).

277   Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821, doi:10.1126/science.1225829 (2012).

278   Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826, doi:10.1126/science.1232033 (2013).

279   Gaudelli, N. M. *et al.* Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* **551**, 464-471, doi:10.1038/nature24644 (2017).

280   Liljeruhm, J., Gullberg, E. & Forster, A. C. *Synthetic Biology: A Lab Manual.* (World Scientific, 2014).

281   Wang, H. H. & Church, G. M. Multiplexed genome engineering and genotyping methods applications for synthetic biology and metabolic engineering. *Methods Enzymol.* **498**, 409-426, doi:10.1016/b978-0-12-385120-8.00018-8 (2011).

282   Day, L. E. Tetracycline inhibition of cell-free protein synthesis. II. Effect of the binding of tetracycline to the components of the system. *J. Bacteriol.* **92**, 197-203 (1966).

283   Rendi, R. & Ochoa, S. Effect of chloramphenicol on protein synthesis in cell-free preparations of Escherichia coli. *J. Biol. Chem.* **237**, 3711-3713 (1962).

284   Wiegand, I., Hilpert, K. & Hancock, R. E. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* **3**, 163-175, doi:10.1038/nprot.2007.521 (2008).

285   Edgar, R., Friedman, N., Molshanski-Mor, S. & Qimron, U. Reversing bacterial resistance to antibiotics by phage-mediated delivery of dominant sensitive genes. *Appl. Environ. Microbiol.* **78**, 744-751, doi:10.1128/aem.05741-11 (2012).

286   Xiong, X., Chen, M., Lim, W. A., Zhao, D. & Qi, L. S. CRISPR/Cas9 for Human Genome Engineering and Disease Research. *Annu. Rev. Genomics. Hum. Genet.* **17**, 131-154, doi:10.1146/annurev-genom-083115-022258 (2016).

287   Amoasii, L. *et al.* Single-cut genome editing restores dystrophin expression in a new mouse model of muscular dystrophy. *Sci. Transl. Med.* **9**, doi:10.1126/scitranslmed.aan8081 (2017).

288   Amoasii, L. *et al.* Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. *Science* **362**, 86-91, doi:10.1126/science.aau1549 (2018).

289   Specter, M. in *The New Yorker*   (2017).

290   Kahn, J. in *The New York Times*   (2015).

291   Myhrvold, C. *et al.* Field-deployable viral diagnostics using CRISPR-Cas13. *Science* **360**, 444-448, doi:10.1126/science.aas8836 (2018).

292   Chen, J. S. *et al.* CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* **360**, 436-439, doi:10.1126/science.aar6245 (2018).

293   Stark, J. C. *et al.* BioBits Health: Classroom Activities Exploring Engineering, Biology, and Human Health with Fluorescent Readouts. *ACS. Synth. Biol.* **8**, 1001-1009, doi:10.1021/acssynbio.8b00381 (2019).

294   Anderson, M. J., Stark, J. C., Hodgman, C. E. & Jewett, M. C. Energizing eukaryotic cell-free protein synthesis with glucose metabolism. *FEBS Lett.*, doi:10.1016/j.febslet.2015.05.045 (2015).

295   Stark, J. C. *et al.* On-demand, cell-free biomanufacturing of conjugate vaccines at the point-of-care. *bioRxiv*, 681841, doi:10.1101/681841 (2019).

296   Petrescu, A. J., Milac, A. L., Petrescu, S. M., Dwek, R. A. & Wormald, M. R. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology* **14**, 103-114, doi:10.1093/glycob/cwh008 (2004).

297   Jaffe, S. R., Strutton, B., Levarski, Z., Pandhal, J. & Wright, P. C. Escherichia coli as a glycoprotein production host: recent developments and challenges. *Curr. Opin. Biotechnol.* **30**, 205-210, doi:10.1016/j.copbio.2014.07.006 (2014).

298   Song, T., Aldredge, D. & Lebrilla, C. B. A method for in-depth structural annotation of human serum glycans that yields biological variations. *Analytical chemistry* **87**, 7754-7762, doi:10.1021/acs.analchem.5b01340 (2015).

299   Hamilton, B. S. *et al.* A library of chemically defined human N-glycans synthesized from microbial oligosaccharide precursors. *Sci. Rep.* **7**, 15907, doi:10.1038/s41598-017-15891-8 (2017).

300   Keys, T. G. *et al.* A biosynthetic route for polysialylating proteins in Escherichia coli. *Metab. Eng.* **44**, 293-301, doi:10.1016/j.ymben.2017.10.012 (2017).

301   Moustafa, D. A. *et al.* Recombinant Salmonella Expressing Burkholderia mallei LPS O Antigen Provides Protection in a Murine Model of Melioidosis and Glanders. *PloS one* **10**, e0132032, doi:10.1371/journal.pone.0132032 (2015).

302   Fry, B. N. *et al.* The lipopolysaccharide biosynthesis locus of Campylobacter jejuni 81116. *Microbiology* **144 (Pt 8)**, 2049-2061, doi:10.1099/00221287-144-8-2049 (1998).

303   Sturm, S., Jann, B., Jann, K., Fortnagel, P. & Timmis, K. N. Genetic and biochemical analysis of Shigella dysenteriae 1 O antigen polysaccharide biosynthesis in Escherichia coli K-12: structure and functions of the rfb gene cluster. *Microb. Pathog.* **1**, 307-324 (1986).

304   Macpherson, D. F., Morona, R., Beger, D. W., Cheah, K. C. & Manning, P. A. Genetic analysis of the rfb region of Shigella flexneri encoding the Y serotype O-antigen specificity. *Mol. Microbiol.* **5**, 1491-1499 (1991).

305   al-Hendy, A., Toivanen, P. & Skurnik, M. Expression cloning of Yersinia enterocolitica O:3 rfb gene cluster in Escherichia coli K12. *Microb. Pathog.* **10**, 47-59 (1991).

306   Adamo, R. *et al.* Synthetically defined glycoprotein vaccines: current status and future directions. *Chem. Sci.* **4**, 2995-3008, doi:10.1039/c3sc50862e (2013).

307   Avci, F. Y., Li, X., Tsuji, M. & Kasper, D. L. A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nat. Med.* **17**, 1602-1609, doi:10.1038/nm.2535 (2011).

308   Hu, Q.-Y. *et al.* Synthesis of a well-defined glycoconjugate vaccine by a tyrosine-selective conjugation strategy. *Chem. Sci.* **4**, 3827-3832 (2013).

309   Stefanetti, G. *et al.* Sugar-protein connectivity impacts on the immunogenicity of site-selective salmonella O-antigen glycoconjugate vaccines. *Angew. Chem. Int. Ed. Engl.* **54**, 13198-13203, doi:10.1002/anie.201506112 (2015).

310   Purwada, A. *et al.* Ex vivo engineered immune organoids for controlled germinal center reactions. *Biomaterials* **63**, 24-34, doi:10.1016/j.biomaterials.2015.06.002 (2015).

Appendix: BioBits^{TM} Curriculum

Descriptions

Curriculum 1. Let it glow! This pre-lab assignment and lab protocol is designed to support the BioBits^{TM} Bright tunable protein expression module with an optional inquiry-based extension activity.

Curriculum 2. What factors affect CFPS yields? This curriculum piece outlines an inquiry-based lab activity modeled on the BioBits^{TM} Bright tunable protein expression lab. In this activity, students are guided through an independent investigation of biochemical factor(s) that influence in vitro protein synthesis yields.

Curriculum 3. Synthetic biology: Looking to nature to engineer new designs. This curriculum piece outlines a research project in which students are challenged to design and present a solution to a societal challenge of their choice using synthetic biology.

Curriculum 4. How fast is it really? This assignment, designed for high school math classrooms, asks students to calculate rates of transcription and translation.

Curriculum 5. Super power protein! This curriculum piece outlines a research project in which students have the opportunity to design their own super power after learning about some of the "super powers" (e.g., fluorescence) of proteins in biology. Students will present their super power and design an in vitro program that illustrates the super power via expression of fluorescent proteins in a 96-well plate.

Curriculum 6. BioBits^{TM} Health ImageJ tutorial: Quantifying fluorescence of cell-free reactions. This tutorial provides an illustrated, step-by-step guide on how to quantify fluorescence of cell-free reactions using ImageJ, a free program for image analysis.

Curriculum 7. BioBits^{TM} Health pre-lab slides: Exploring antibiotic resistance with cell-free reactions. This slide deck is provided to facilitate pre-lab lectures on antibiotic mechanisms of action, mechanisms of antibiotic resistance, laboratory procedures, and strategies to prevent antibiotic resistance.

Curriculum 8. Cell-free protein synthesis and antibiotics. This curriculum piece guides students through an investigation of the mechanisms of action of a variety of different antibiotics through a combination of online research and hands-on lab activities.

Curriculum 9. Exploring antibiotic resistance in cell-free systems. This lab activity asks students to investigate mechanisms of antibiotic resistance (genetic mutation, horizontal gene transfer) using cell-free reactions and research strategies to prevent antibiotic resistance.

Curriculum 10. CRISPR sort. A sorting and classification exercise designed to help students understand the functions of the necessary components for Cas9 activity (*i.e.*, nuclease, gRNA, target, PAM site).

Curriculum 11. BioBits™ Health pre-lab slides: Exploring CRISPR-Cas9 genome editing with cell-free reactions. This slide deck is provided to facilitate pre-lab lectures on CRISPR-Cas9 genome editing, laboratory procedures, and ethical considerations associated with genome editing.

Curriculum 12. Using fluorescent proteins as reporters of CRISPR-Cas9 activity. This lab protocol is designed to support the CRISPR fluorescent protein screening lab activity.

Curriculum 13. CRISPR Huntington's disease activity. This assignment asks students to investigate a variety of gene editing technologies (*i.e.*, CRISPR, zinc finger nucleases, etc.) for treating Huntington's disease.

Curriculum 14. Exploring solutions to genetic problems. A process oriented guided inquiry learning (POGIL) exercise that guides students through an exploration of possible solutions to treat a patient with sickle cell anemia.

Curriculum 15. Applications of CRISPR. An independent or group-based research project in which students investigate a societal problem of their choice that could be solved with CRISPR technology.

Curriculum 16. Ethics of CRISPR. Students are asked to consider the question: "If you can solve a problem should you?". Through this exercise, students carry out independent research to construct an argument on the ethics of using CRISPR technology to address human disease and other societal problems. Students will also research and consider the ethics surrounding the 2018 report of CRISPR editing of humans in China.

Curriculum 1

Name: _____ Date: _____

Cell-Free Protein Synthesis Using Fluorescent Proteins

Introduction:

How can we utilize natural processes to create products for humans, such as medicines or materials needed for life? Synthetic biology aims to do just this. Scientists in this field use living systems to create engineered products, combining engineering and biology into one area of science. In particular, synthetic biologists work to hijack living systems and processes to create a specific product. This has a multitude (many) of applications in the real world. We can make better medications (treatment), safer and more affordable vaccines (prevention), materials such as adhesives, or other products to improve the quality of human life.

Background:

One of the systems that scientists have been able to utilize is protein synthesis. Today, most insulin is mostly made biosynthetically (artificially) by recombinant DNA technology or 'genetic engineering'. Rather than being extracted from animal sources, commercially available human insulin is now manufactured through recombinant DNA technology, in which the gene for making human insulin is transferred into simple cells such as bacteria or baker's yeast. In 1978, Genentech in South San Francisco used this process to create recombinant insulin known as rInsulin, or Humulin©.   Because insulin is a protein coded for by a gene, Genentech developed a process of inserting the gene for human insulin into a bacterial plasmid, tricking bacteria into growing large volumes of insulin for diabetic patients worldwide. This discovery helped Genentech become one of the first biotech companies in the world, and it continues to be one of the largest and most influential even to this day. This is a great example of utilizing a naturally occurring process to create a product to improve the quality of life for humans.

When scientists want to produce a gene or protein of interest they engineer DNA, insert it into the organism that will produce the protein, and hope that the cell will actually

make their protein. However, this doesn't always work. We don't fully understand how cells make proteins, so sometimes the cell may choose not to produce much or any of your protein. Additionally, sometimes the protein you want to make can kill the cell, which also means little or no protein is made. For these reasons, sometimes it can be difficult to produce proteins in cells, although as we know it can be done.

Carlson, et.al. *Biotech Adv* (2011)
Perez, Stark, & Jewett. *Cold Spring Harb Perspect Biol.* (2016)
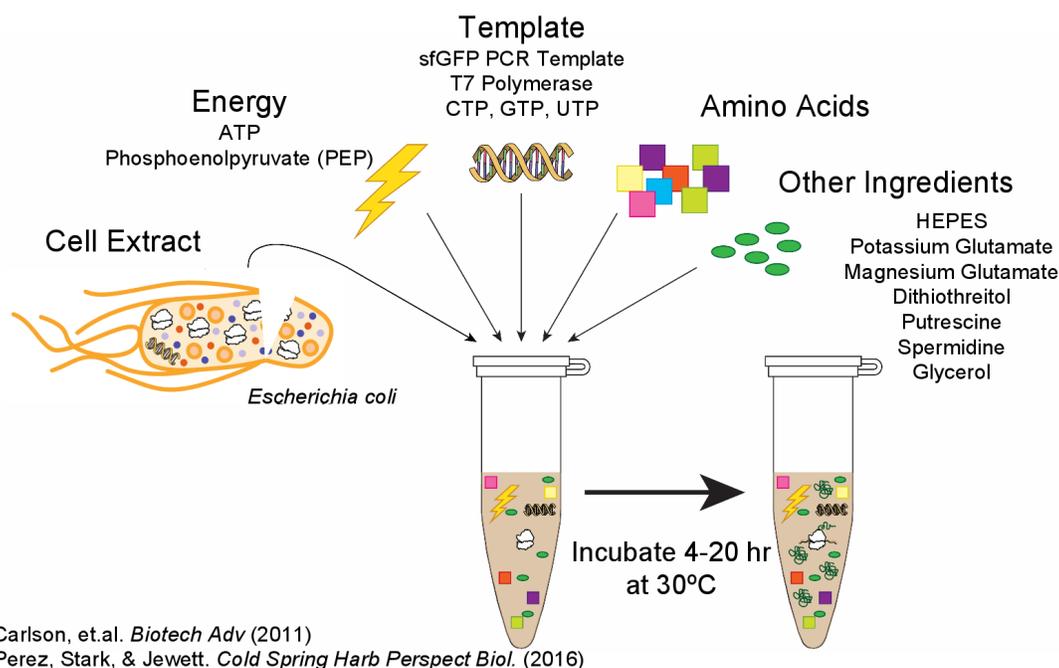
To combat this exact problem, scientists have come up with a process called **Cell-Free Protein Synthesis** (CFPS). CFPS was first developed in the 1950's, when it was used to help crack the code that converts mRNA to proteins, but recently has been adopted for synthetic biology. CFPS is created by lysing the cell (breaking the cellular membrane of a cell) and extracting the all of the parts from inside of the cell. We can use this extract to carry out many of the cellular functions that would occur within the cell, such as protein synthesis, in nonliving or *in vitro* reactions (as opposed to *in vivo* which is in a living cell). Since this is a nonliving extract, this method allows synthetic biologists to bypass normal cellular decision making and allows us to create our protein product. The process is also faster and creates a higher yield of your product as the reaction doesn't have to manufacture any other proteins as would have to be done in a living system.

We can use cell-free protein synthesis to then manufacture proteins of interest, such as insulin or antibiotics. When we create a protein of interest, scientists can sometimes take the product and run it through a gel electrophoresis to make sure that the protein we wanted was actually created. They do this by confirming expected molecular weight or even send the protein off to be sequenced. This can be costly and take up too much valuable time.

Other times, scientists attach their protein of interest to a reporter protein. Reporter proteins often have color or fluorescence that can be seen by eye to indicate whether a

certain protein had been expressed. These proteins indicate that your reaction worked. Some examples of reporter proteins are GFP (green fluorescent protein) and RFP (red fluorescent protein), which are very easy to use because they glow either green (GFP) or red (RFP) when expressed. This method of reporting can be used to indicate successful protein synthesis in CFPS or *in vivo*.
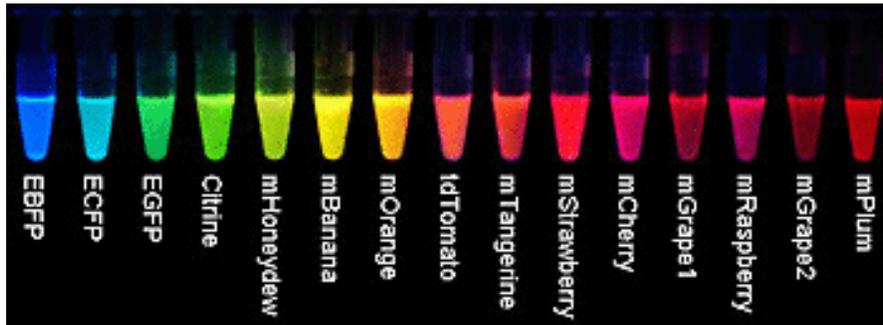
Typically an engineered plasmid looks like this:
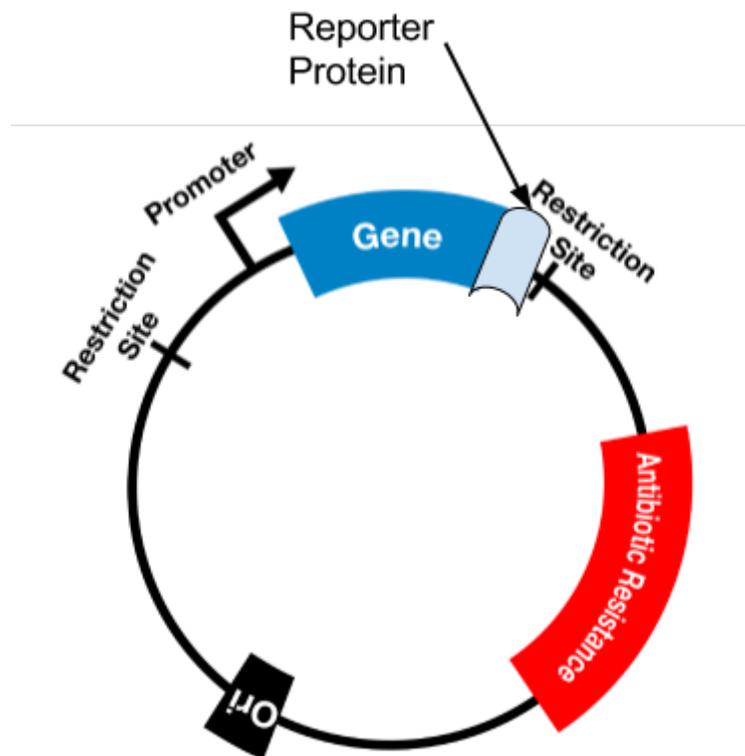
1. **Why might a scientist use this method of reporting? What could you predict that this method can tell us about concentrations of inputs (DNA) and products (protein)? Explain your reasoning.**

Since we are going to be carrying out the process of Cell-Free Protein Synthesis (CFPS) in our lab, we need to review the process. Protein Synthesis is broken into two distinct processes, transcription and translation. **Below, diagram what these two processes are:**
Include the following in your diagram: DNA, mRNA, tRNA, ribosomes, RNA polymerase, amino acid chain, amino acids, free nucleotides.

2. **What are three example proteins that can be made inside of a cell?**

3. **In the following lab, we will be carrying out cell-free protein synthesis. We are going to be adding a plasmid (DNA) into a cell-free environment, so what must already be in that environment for protein synthesis to occur? List at least three things.**

In this lab we will be carrying out cell free protein synthesis to create fluorescent proteins. In the first part of this lab, we will be looking at different concentrations of plasmid DNA and the amount of protein that is synthesized. In the second part of the lab we will be investigating the optimal conditions for protein synthesis to occur in a cell-free system.

How is cell-free protein synthesis used in the real world? What practical applications does this have for other sciences like chemistry or medicine? We talked about a couple of ways that we can use recombinant DNA (inserting DNA from one organism into another organism's DNA) to produce a product. How can we further use this recombinant DNA to make something that does not need to be produced in a cell using cell-free protein synthesis? Already being produced are low-cost diagnostic indicators for things such as the Zika virus, vaccines, and other medicine.

Cell-Free Synthesis of Fluorescent Proteins Lab
<u>Let it Glow!</u>

Objectives
Students will be able to...
- Visually observe and collect data to compare protein concentration using relative fluorescence as a reporter.
- Analyze data to draw conclusions about the effects of (a) changing variable(s) on the rate of protein synthesis.
- Design and implement an experiment to test the effects of a variable on the rate of protein synthesis in a cell-free (*in vitro*) system.
- Use data to determine the optimal conditions for the cell-free transcription and translation reactions.
- Propose an explanation for the observed relationships between variables.

Purpose of the Experiment
To test the effect of DNA concentration on the amount of fluorescent protein that is produced in a cell-free protein synthesis system.

Materials
- Freeze-dried cell-free (FD-CF) reaction in PCR tube strips
- Plasmid DNA
- Nuclease free water
- Sterile pipet tips
- Microcentrifuge tubes for DNA dilution

Important Notes
Be careful to avoid contamination by using a new pipet tip for each solution! Do not touch your pipet tips to any surface other than the inside of the tubes while you are working.

Procedures
DAY 1
1. Locate the lyophilized stock for each plasmid construct. Rehydrate the plasmid DNA by adding 40 µL of nuclease-free water to the tube. Pipet the solution up and down to mix.
2. Label the empty tubes "25 ng," "10 ng," "5 ng," and "2.5 ng." These will be your plasmid dilutions.
3. Add the appropriate amount of water (refer to the table below) to each of the dilution tubes.
   a. You can use the same pipet tip for each water addition as long as you do not touch any surface other than the inside of the tubes! When you start pipetting the DNA later you'll need to switch pipet tips, but not yet.
   b. Remember to keep all of your tubes on ice.

Tube	Water (µL)	Plasmid DNA (µL)
25 ng	1.25	10
10 ng	7.25	4
5 ng	9.25	2
2.5 ng	10.25	1

4. Label your PCR tubes containing the CFPS reactions. The first two tubes in the strip should be labeled "25," the next two "10," the next two "5," and the last two "2.5." Label these on the side of the PCR tubes. Put the initials of your group members or your group number on the SIDE of the strip.
5. Refer to the table above to add the appropriate amount of DNA to your 25 ng tube, ONLY. Vortex or pipet up and down several times to mix.
6. Being careful to *avoid touching the white pellet* at the bottom, add 4.5 µL of the DNA/water mixture from your 25 ng dilution to the first CFPS reaction tube labeled "25." Add the same amount to the second "25" tube.
7. Repeat steps 5 and 6 for your 10, 5, and 2.5 tubes.
8. INSTRUCTOR: add 4.5 µL of nuclease free water to each of 2 additional freeze-dried CFPS reactions. These will serve as a no DNA control (negative control) for your experiment.
9. Seal your CFPS reactions using a 8-strip PCR tube lid. Gently flick the side of the PCR tubes to dissolve the freeze-dried CFPS pellet in the DNA/water mixture that you added.
10. Allow the reactions to incubate until the next time you come to class. Reactions can be incubated in a PCR thermocycler for 20 hours at 30°C or on the benchtop at room temperature for 48 hours. Get excited!

DAY 2
11. Obtain your group's samples from your teacher. Place your strip of reactions in front of a piece of white paper, and record your observations.
12. Obtain a black light pen or a blue light imager and shine it on the tubes. Record your observations.
13. Take a picture of your reactions under blue light and quantify fluorescence produced in each reaction with ImageJ. Record your results in an *appropriately labeled data table* that you should create.

Analysis
1. Generate a graph of your results. Include appropriate scale, labels, units, and a descriptive title.
2. Use your results to make a **claim** about the relationship between plasmid concentration and protein concentration. Use the **evidence** you collected to support your claim. Explain your **reasoning**. (This should be approximately one paragraph).
3. Using your understanding of transcription and translation, propose an explanation for *why* this pattern exists. What is happening in the cell-free system that causes this to occur?
4. If there are any data points that seem to be outliers that do not fit the pattern, propose an explanation for why this may have happened. Remember, simply writing "human error" is *never* an acceptable response! Be specific!
5. Consider the lab procedures. Do you see any sources of error? Propose one change that could be made to improve these procedures.
6. Based on the observations you made in this activity, brainstorm a list of *testable* questions about other factors that might affect the rate of protein synthesis.

Inquiry Lab
As a class, we need to answer the question "What are the optimal conditions for protein synthesis in this cell-free system?" Your group needs to decide on a specific condition that you would like to test, then design a procedure for testing that condition.
1. Your teacher has a list of available supplies that you can use. Before writing your procedure, check with your teacher to make sure that your question is testable and that you have the needed equipment/supplies available.

2. Write out your plan, which should include:
   a. The question you are trying to answer
   b. Identification of independent, dependent, and controlled/constant variables
   c. An experimental control
   d. A hypothesis (prediction *and* reason for that prediction)
   e. Materials
   f. Detailed procedures, including amounts of each of the supplies
   g. Data tables and a plan for collecting/analyzing your results
3. Show your teacher your completed plan for approval.
4. Implement your plan, collect your data, and analyze your results!
   a. Remember to be careful to avoid contamination when pipetting! Suggestion: if water needs to be transferred to multiple tubes, do all of these FIRST so that you can use the same pipet tip for all of your aliquots of water!
   b. Keep everything on ice until you are finished!
   c. All of your data must be organized in data tables and graphed.
5. Submit your data to your teacher.
6. You will be given a summary of data as collected from several groups of students. Use all of the data sets to make a **claim** about the optimal conditions for the transcription and translation process in this cell-free system. Support your claim using **evidence**, and provide **reasoning** for *every piece of evidence*. (This should be approximately 2-3 paragraphs)
7. After thoroughly discussing and supporting your claim, use your understanding of biochemical reactions to propose an explanation for *why* this is the optimal set of conditions.
8. If you were to do this lab again, what changes or improvements would you make?
   a. Discuss ways to improve your procedure for testing the same question.
   b. Additionally, describe another experiment you could do to further test the claim that you made about the optimal conditions for these reactions. The focus for this experiment should be on learning more about the *same* conditions we've already tested, not on proposing a new condition to test.

Teacher Notes
Adapted from an activity developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University), Rachel Dubner (B.S. Candidate, Biology and Psychology, Northwestern University), Karen Hsu (B.S. Candidate, Mechanical Engineering, Northwestern University), and Dr. Michael Jewett (Charles Deering McCormick Professor of Teaching Excellence, Chemical Engineering, Northwestern University). Written by Kara Reichert, M.Ed. (Jones College Prep) and Ashlee Miller-Berry, M.A.T. (Evanston Township High School).

For the inquiry portion of the lab, we suggest that students consider varying the following parameters: concentration of the overall mixture (amount of water added); amount of specific ions (magnesium is a good one to consider, but this can only be done by increasing the amount as we cannot remove it from the original cell-free reaction pellet); temperature; and pH. You might also consider an extension activity to ask students to determine if different plasmids require different sets of "optimal conditions." Another possible extension of this activity would be in titrate in various ribosome inhibitor antibiotics to guide students through an investigation of antibiotic mechanism of action.

Curriculum 2

Lab: "What factors affect cell-free protein synthesis yields?"

Introduction

The purpose of this lab is to propose an answer to the question "what factors affect cell-free protein synthesis yields?" You will answer this question by identifying several factors that affect protein synthesis. Then, you will devise an experiment to test the effect of one of these factors on the cell-free production of a fluorescent protein.

Step 1: Background Research

Protein Factories

    1.   First, let's confirm that we are using the same vocabulary. Define the following terms:

molecule	purpose
DNA	
mRNA	
tRNA	

2. What determines the shape and function of a protein?

3. In this lab, you will use cellular machinery to build and optimize your own protein factories in a test tube. With your group, develop a list of the cellular "ingredients" (organelles, enzymes, and materials) are needed to make a protein.

4. As with all biological processes, the process of protein synthesis can be simplified into a series of inputs, sensors, actuators, and outputs. Fill in the block diagram below, using the list of protein synthesis ingredients you created as a guide.
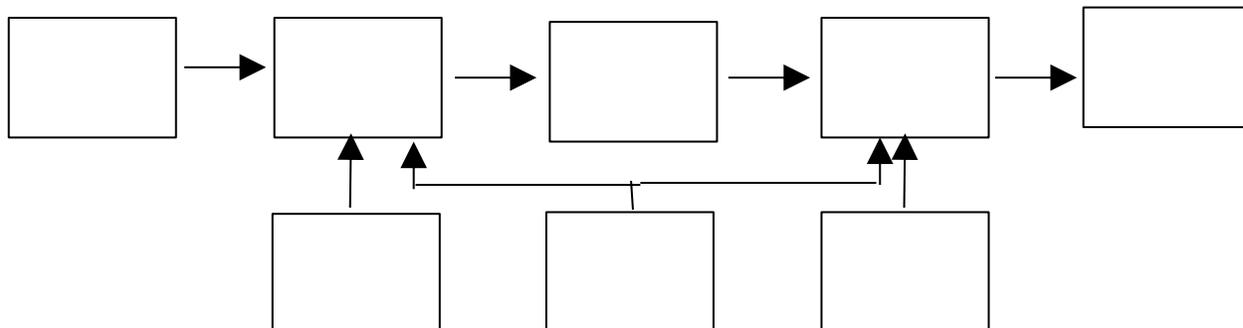
Efficiency and Output

5. The two processors you have listed above are both enzymes. Below, list three environmental factors that can change the rate of an enzyme-catalyzed reaction and explain how a change in each affects reaction rate:

In addition to the environmental factors listed above, scientists can alter the rate of protein synthesis by adding inhibitors, such as the following biological compounds:

> ***Antibiotics***
> All living things need a mechanism for protein synthesis. Though ribosomes are very similar across organisms, there are clear differences between these in eukaryotic and prokaryotic systems. Scientists have developed antibiotics that directly target bacterial ribosomes. The following are some examples:
> - Chloramphenicol: prevents protein chain elongation
> - Kanamycin: causes misalignment of mRNA in ribosome
> - Erythromycin: prevents movement of tRNA from the A- to P-site in the ribosome
>
> ***Protease Inhibitors***
> Proteases are enzymes that digest larger proteins into smaller peptides. These enzymes are incredibly diverse and serve many biochemical roles. Protease inhibitors are chemical compounds that block the action of proteases.

Step 2: Brainstorming investigation strategies

Recall that the purpose of this lab is to answer the question, "what factors affect cell-free protein synthesis yields?" Now, let's design an experiment to test this general question.

5. In this experiment, the output of interest will be a functional, fluorescent protein. How can you measure the output of this reaction? What materials could you use to do this?

6. Review the information you collected in your background research to identify measurable factors can be manipulated in your protein factory:
   ● What environmental factors could be changed? How?

   ● What inputs can be changed? How?

   ● What could you do to change the function of one of the processors?

6. Examine your list above. Choose one variable to manipulate. Briefly describe a method to determine the effect of a change in this factor on the amount of functional protein produced.

7. Why do you think your method of testing will help answer the general question, "what factors affect cell-free protein synthesis yields?" What do you expect will happen?

Step 3: Refine your question and plan your investigation
Your general question is: "how can cell-free protein synthesis be optimized?" Based on the research above, write a question that clearly explains the variables you intend to test and measure. Then, consider what variables must be controlled in your experiment. Be sure to choose variables that can be measured qualitatively or quantitatively!

Question	
Independent Variable (IV, treatment)	
Dependent variable (DV)	
Hypothesis (if, then, because)	
How to measure IV	

How to measure DV	
Control treatment	
Controlled variables	
How will you collect data, and how often will you do it?	
How will you organize your data?	
How will you analyze your data?	

Step 4: Perform your study

1. Develop and refine a procedure
    a. Determine how you will conduct your experiment, using your answers to the table in step 3 as a guide. Create a procedure and check it, as well as the information in step 3, with your teacher.

    b. Test your procedure. Determine whether your setup will generate reliable data that will accurately test your hypothesis and yield meaningful, measurable results. Refine and record your procedure.

    2. Conduct your experiment:
- a. Create a method of collecting data
- b. Record any deviations from the procedure or sources of error
- c. Ensure that your workspace and materials are cleaned up.

Step 5: Report your findings

You will report your results in two ways.
1. Prepare a whiteboard that shows: your experimental question, your data, and your interpretation of the data. Use both a graph and a particle picture to support your claim.
2. Complete a lab report that includes an introduction, materials and methods, results, discussion, and conclusion.

Lab report guidelines:
1. Introduction:
   - a. Provide background on your subject.
   - b. Explain why investigation into this subject is important in a larger scientific context.
   - c. Explain what your experiment sought to test.
   - d. State your hypothesis and justify why your experiment tests this hypothesis.
2. Materials and Methods:
   - a. Explain what you did in your experiment in paragraph form. Include volumes or masses used. Include a diagram of your setup. Do not list every step of your procedure!
3. Results: Clearly summarize your data without making inferences or conclusions. Include quantitative and qualitative observations. Include averages or ranges, when appropriate. Show your results in a table or graph. Be sure to label your tables and graphs, as well as any axes, with units. Title your tables and graphs as "Figure #."
4. Discussion: Describe what you observed.
   - a. Interpret your data and state your findings. Refer to data by the "figure #"
   - b. If some data is aberrant from other trials, explain those data and what might account for the differences.
   - c. State the importance of your findings. Provide alternate explanations for your findings.
5. Conclusion
   - a. Summarize how your findings relate to your hypothesis.
   - b. Provide recommendations for future research, applications of your findings, and additional questions that have arisen during your research.

Teacher notes:
Adapted from an activity developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University), Rachel Dubner (B.S. Candidate, Biology and Psychology, Northwestern University), Karen Hsu (B.S. Candidate, Mechanical Engineering, Northwestern University), and Dr. Michael Jewett (Charles Deering McCormick Professor of Teaching Excellence, Chemical Engineering, Northwestern University). Written by Aparna Puppala (Glenbrook South High School).

Curriculum 3

Names:_____

Synthetic Biology: Looking to Nature to Engineer new Designs

The United States is a leading nation in the development of synthetic biology, an emerging engineering discipline to create, control and reprogram biological systems. Synthetic biology promises to provide sustainable solutions to many grand challenges of the modern society via innovations in agriculture like bioremediation or the production of chemicals, pharmaceuticals, and energy. For example, many pharmaceuticals today are produced by reprogramming cells to produce medicines like insulin or human growth hormone, and the possibilities are almost endless. However, we are just getting started and synthetic biology will allow us to solve problems in completely new ways such as creating living home structures that repair themselves, naturally insect resistant building lumber straight from the tree, or trees that produce light living along road sides. Mother nature has had millions of years to create such characteristics through the process of evolution by natural selection, so we should first look to natural occurring genetic solutions that we can move from one organism to another.

ASK: What is your design supposed to do? What problem do you want to solve?	
IMAGINE: Apply knowledge and creativity to brainstorm ideas together. List 3+ ideas. Agree on one to try!	Idea 1: Idea 2: Idea 3: Final idea:

243

PLAN: Consider what organism you'll be modifying and how you'll be modifying its function. What changes will you make? Sketch/draw the details of your idea	Device/organism: Design sketch:
CREATE: What cell parts will be necessary to change your organism or a product it produces? Is there any naturally occurring organism that possesses some or all the design characteristics you're interested in? What DNA/gene will you need to insert into your organism?	Cellular parts necessary: Potential naturally occurring sources of trait or characteristic: Specific DNA/gene found in that organism:

Teacher notes
Adapted from an activity developed by BioBuilder Educational Foundation and meant to complement freeze-dried cell-free (FD-CF) lab activities developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University), Rachel Dubner (B.S. Candidate, Biology and Psychology, Northwestern University), Karen Hsu (B.S. Candidate, Mechanical Engineering, Northwestern University), and Dr. Michael Jewett (Charles Deering McCormick Professor of Teaching Excellence, Chemical Engineering, Northwestern University). Written by Lance Brand, Ph.D. (Delta High School).

Curriculum 4

Name:_____

How Fast is it Really?

Background on Transcription/ Translation and *E. coli* in Recombinant Systems

In living organisms, genetic information for synthesizing proteins is encoded in DNA molecules made of nucleotides that pair together with a complementary base to create a base pair (bp). In order to use the information, a temporary messenger RNA (mRNA) copy is made in the process of transcription (see below) by an enzyme called DNA polymerase. That mRNA molecule is translated by a ribosome into the protein made of amino acids (aa) that your cell needs!

In this activity, we are going to focus on the genetic information in a bacterium, *E. coli*. Remember that a gene is a piece of DNA which provides the instructions for making a protein. This protein gives an organism a particular trait. *E. coli* is commonly used in a process known as genetic transformation and involves the insertion of a gene into an organism in order to change the organism's trait. Genetic transformation is used in many areas of biotechnology. In agriculture, genes coding for traits such as frost, pest, or spoilage resistance can be genetically transformed into plants. In bioremediation, bacteria can be genetically transformed with genes enabling them to digest oil spills. In medicine, diseases caused by defective genes are beginning to be treated by gene therapy; that is, by genetically transforming a sick person's cells with healthy copies of the defective gene that causes the disease.
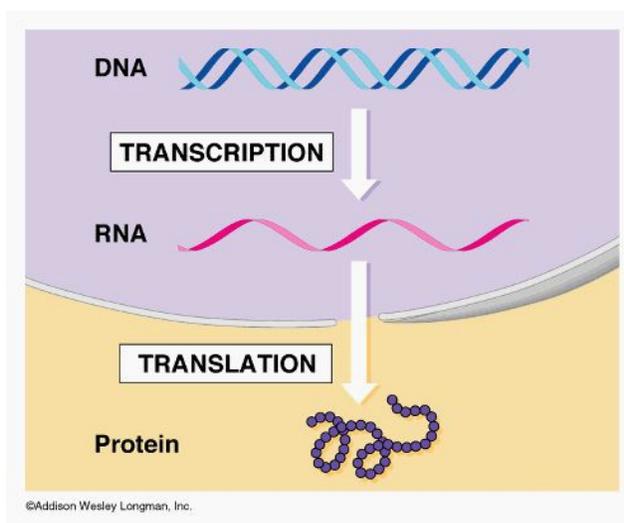
In this activity, we are going to find out how fast this process happens within *E. coli.* Scientists and engineers work to optimize this process in order to produce the highest levels of protein possible. For example, scientists use *E. coli* to produce insulin to treat diabetes. In order to both treat as many people as possible and reduce costs of making insulin, they would like to make protein synthesis happen as quickly as possible.

(Adapted from Bio-Rad pGlo Bacterial Transformation Lab manual)

Part 1: *E. coli* Transcription and Translation Calculations:

The genome (an organism's complete set of DNA) of *E. coli* consists of 4.6 million base pairs (bp) with an average gene length of 834 bp. If the *E. coli* polymerase can transcribe 20 bp/s, how many seconds would it for the polymerase to transcribe the average gene?

Your polymerase has now created an mRNA molecule with 834 bp. Recall that the genetic code is a triplet code (3 nucleotides/ base pair code for 1 a.a). How many amino acids would this molecule encode?

If the *E. coli* ribosome can translate 20 a.a. per second, how long would the translation process take?

How long do the transcription and translation processes take to get from DNA to an unfolded protein?

Part 2: Polymerase Comparison:

In order to optimize protein synthesis, chemical and biological engineers have used several different polymerases for transcription. The Taq polymerase can transcribe 1 kb/ min. How does this transcription rate compare to the native *E. coli* you used in part 1?

A new polymerase named Q5 is being sold by biotech companies that has a transcription rate of 4kb/min. How much time would you save if you used Q5 instead of Taq?

The average human gene is 53.6 kB long, how long would it take for Q5 to transcribe the gene?

A human can read approximately 250 words per min. If each base is considered to be a word, how long would it take the average human to read the average gene?

Part 3: Challenge Questions Genome Comparisons

In 1990, the Human Genome Project was officially launched to study and sequence the entire human genome. Completed in April of 2003, the project found that the human genome consists of 3 billion base pairs. How long would it take Taq to transcribe the entire genome?

The title for the longest known genome is currently held by the rare Japanese flower with a whopping 149 billion base pairs. If Taq, Q5, and native E. Coli polymerases started transcription at the exact same time on identical copies of the *Paris japonica* genome, how long would it take the fastest polymerase?

What percentages would the other two polymerases have completed in the same amount of time?

If a human "read" the genome of *Paris japonica* at 250 bp/min how long would it take?

Teacher Notes:

Use:

This activity is designed to complement a protein synthesis unit and reinforce unit rate calculations typically introduced in middle school math (7.RP). Students can better grasp the scale and rate of the protein synthesis and increase their mathematical reasoning skills to have a more well-rounded picture of how this process takes place in all living cells on a continual basis. Students can work independently or in small groups through the calculations. Choose students based on their methods (proportions, conversion factors, etc.) and have a variety of students present and explain their approach to finding an answer. Focus on different approaches and interpreting results in terms of context.

Standards:

CCSS.MATH.CONTENT.7.RP.A.1
Compute unit rates associated with ratios of fractions, including ratios of lengths, areas and other quantities measured in like or different units. For example, if a person walks 1/2 mile in each 1/4 hour, compute the unit rate as the complex fraction 1/2/1/4 miles per hour, equivalently 2 miles per hour.

NGSS CC.3: Scale, Proportion, and Quantity
In considering phenomena, it is critical to recognize what is relevant at different size, time, and energy scales, and to recognize proportional relationships between different quantities as scales change.

Suggestions:

- First page can be used as a separate activity without Parts 2 and 3 if desired.
- Students may need access to a metric conversion chart specifically kilo to base units.
- Can be used as an out-of-class assignment to start a conversation on scale.

Data:

Taq (1kb/min)/ Q5 (up to 4kb/min) [1] / E.coli native polymerase (10-100 nt/s (b/s))[4] Longest Genome (A rare Japanese flower named *Paris japonica*):149 billion base pairs [2] Human Genome Length 3 billion base pairs [3]

Translation Rate 20 a.a per sec [4]

Longest Human Gene CASPR2 (CNTNAP2) 2.30 Mb [5]

E. coli genome 4.6 million bp [6]

Genome Project Info [7]

Average human gene length 53.6 kB [8]
Average human reading speed 250wpm [9] Average E. coli protein length 278 a.a. [10]

Assumptions:

Translation and transcription do not happen concurrently, which can be true in prokaryotes [4]. Also, time between transcription ending and translation beginning is not included.

Citations:

1. https://www.neb.com/applications/dna-amplification-pcr-and-qpcr/specialty-pcr/fast-pcr
2. http://www.sciencemag.org/news/2010/10/scienceshot-biggest-genome-ever
3. https://www.genome.gov/11006943/human-genome-project-completion-frequently-asked-questions/
4. http://book.bionumbers.org/what-is-faster-transcription-or-translation/
5. http://www.cshlp.org/ghg5_all/section/gene.shtml
6. http://book.bionumbers.org/how-many-genes-are-in-a-genome/
7. https://www.genome.gov/10001772/all-about-the--human-genome-project-hgp/
8. http://kirschner.med.harvard.edu/files/bionumbers/Human%20genome%20and%20human%20gene%20statistics.pdf
9. http://www.healthguidance.org/entry/13263/1/What-Is-the-Average-Reading-Speed-and-the-Best-Rate-of-Reading.html
10. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1150220/

Answer Key:
Part 1: *E. coli* Transcription and Translation Calculations:
The genome (an organism's complete set of DNA) of *E. coli* consists of 4.6 million base pairs (bp) with an average gene length of 834 bp. If the *E. coli* polymerase can transcribe 20 bp/s, how many seconds would it for the polymerase to transcribe the average gene?
41.7s

Your polymerase has now created an mRNA molecule with 834 bp. Recall that the genetic code is a triplet code (3 nucleotides/ base pair code for 1 a.a). How many amino acids would this molecule encode?
278 a.a.

If the *E. coli* ribosome can translate 20 a.a. per second, how long would the translation process take? 13.9s

How long do the transcription and translation processes take to get from DNA to an unfolded protein? 55.6s

Part 2: Polymerase Comparison:
In order to optimize protein synthesis, chemical and biological engineers have used several different polymerases for transcription. The Taq polymerase can transcribe 1 kb/ min. How does this transcription rate compare to the native *E. coli* you used in part 1?
1.2kb/min (*E. coli*) vs.
1kb/min (Taq) OR
20bp/s (*E. coli*) vs.
16.67bp/s (Taq)

A new polymerase named Q5 is being sold by biotech companies that has a transcription rate of 4kb/min. How much time would you save if you used Q5 instead of Taq?
You would save 75 % of your time! It's 4 times faster.

The average human gene is 53.6 kB long, how long would it take for Q5 to transcribe the gene?
13.4 min

A human can read approximately 250 words per min. If each base is considered to be a word, how long would it take the average human to read the average gene?
214 min, 3.6 hours

Part 3: Challenge Questions Genome Comparisons
In 1990, the Human Genome Project was officially launched to study and sequence the entire human genome. Completed in April of 2003, the project found that the human genome consists of 3 billion base pairs. How long would it take Taq to transcribe the entire genome?
5.7 years

The title for the longest known genome is currently held by the rare Japanese flower with a whopping 149 billion base pairs. If Taq, Q5, and native *E. coli* polymerases started transcription at the exact same time on identical copies of the *Paris japonica* genome, how long would it take the fastest polymerase?

70.87 years

What percentages would the other two polymerases have completed in the same amount of time? Taq (25%)
E. coli (30%)

If a human "read" the genome of *Paris japonica* at 250 bp/min how long would it take? 1,133.9 years

Curriculum 5

Super Power Protein Project

You have been selected to create your very own super power. Some superheroes are Superman, Batman, Spiderman, and Wonder Woman. There are also Groot, Iron man, Green Lantern and more. What are their powers? Where do their powers come from? Well now you have an opportunity to gain a superpower. You have learned about cells, DNA, and protein synthesis. You now need to choose what superpower you want, where in nature can you obtain it, and how through protein synthesis will you get it. Ahh, research!!! You need to delve into research to find the superpower that best meets your needs. Think about the way other characters, and then you, have gained the ability to generate an effect. You will present your super power to the class. You will also use cell-free protein synthesis to design a logo or image in a 96 well plate that illustrates your super power. You will have DNA encoding red, pink, orange, yellow, and green proteins available to use in your logo design.

Ways in which superpower characters have gained the ability to generate an effect

Type	Explanation	Examples	Superpowers
Inhuman Nature	Class of partial non-human beings but their superhuman capabilities are typical	➢ Aliens ➢ Zeus & Hippolyta	➢ Groot (Guardians of the Galaxy) alien from a planet of sentient plants that appear like trees ➢ Wonder Women
Object-based powers	Powers derived from objects,	➢ Iron man ➢ Green Lantern ➢ Harry Potter ➢ Wonder women	➢ Metallic suit of armor ➢ Power ring ➢ Wand, invisible cloak ➢ Lasso of Truth
Mutation	These powers come from evolution or natural selection	➢ Angel X-men ➢ Captain Comet ➢ Spider-man	➢ White feather wings began to grow ➢ Comet emitted low-level waves of radiation which triggered multi-genetic cell growth ➢ Bitten by radioactive spider
Induced (Accidental)	These powers are a result of an accident or bizarre event that altered the genetics of the individual	➢ Fantastic Four ➢ The Flash	➢ Exposure to cosmic rays ➢ Lightning strike at his lab and doused with chemicals
Induced (Intentional)	Powers are purposely given as a result of an event, through experimental science.	➢ Captain America	➢ Through the Super-Soldier Serum and "Vita-Ray" treatment his physical abilities, such as strength, endurance, agility, speed, reflexes, are just short of being rated as superhuman.

Superpowers found in organic manipulation and animal mimicry

Type	Explanation	Examples	
Plant Manipulation	The ability to create, shape, and manipulate plants. To cause plants to possibly grow, move/attack, "walk", mutate plants by rearranging DNA structure	➢ Grass-type Pokemon ➢ Plantman	

Animal Imitation	The ability to use or imitate animal powers. They can run fast, swing, move stealthily, track, or have strength.
	
	Animal Man
	
		➢ Fastest animal on
		  ○ Land: Cheetah
		  ○ Air: Peregrine falcon
		  ○ Fish: Black Marlin	

Electricity Manipulation	The ability to manipulate electricity. **Bioelectricity**- the generation of electricity by living organisms. Allows plants, insects, fish, and other organisms to glow or produce electrical charge.	➢ Electric eel ○ Uses electric shocks for hunting and self defense	

Use the 96 well plate template below to design a logo for your superpower
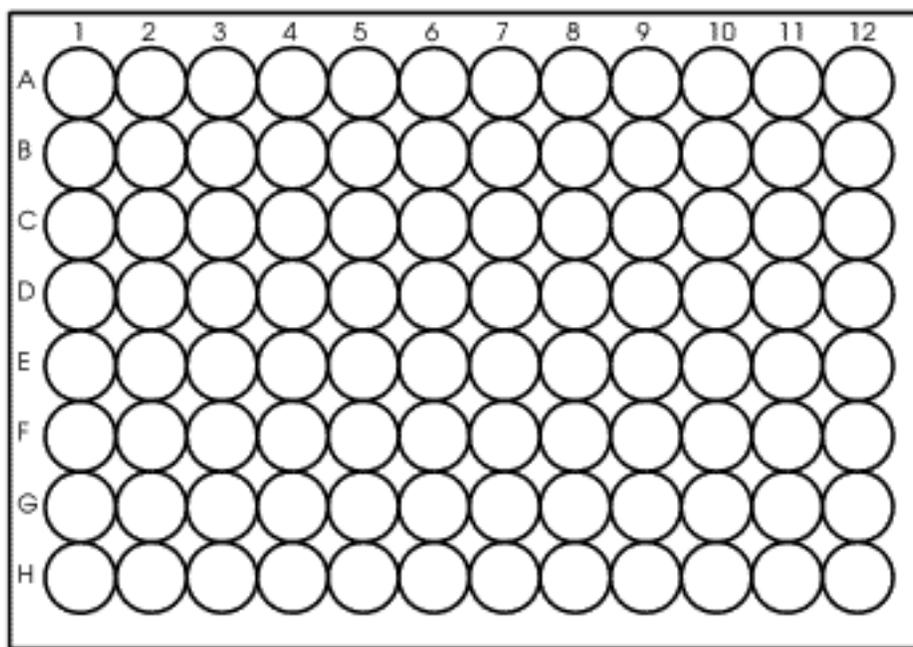
Teacher notes

Adapted from an activity developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University), Rachel Dubner (B.S. Candidate, Biology and Psychology, Northwestern University), Karen Hsu (B.S. Candidate, Mechanical Engineering, Northwestern University), and Dr. Michael Jewett (Charles Deering McCormick Professor of Teaching Excellence, Chemical Engineering, Northwestern University). Written by Julie Burke, M.Ed. (Cleveland School).

Amoeba Sisters Video
https://www.youtube.com/playlist?list=PLwL0Myd7Dk1F0iQPGrjehze3eDpco1eVz

Amoeba Sisters Protein Synthesis
https://www.youtube.com/watch?v=h5mJbP23Buo&index=26&list=PLwL0Myd7Dk1F0iQPGrjehze3eDpco1eVz

What is the relationship between DNA and proteins in a cell?
- Protein synthesis is the process in which a cell makes protein.
- Transcription is the first phase where DNA is converted into mRNA.
- Translation is the second phase where the mRNA is converted to the tRNA and a protein molecule is assembled.

Why are proteins important?
- Protein molecules are important to us in a variety of ways. Our body is constructed from protein molecules, such as muscle, ligaments, skin, hair. There are smaller protein molecules that keep our body working, such as haemoglobin hormones (insulin), antibodies, and enzymes.

<u>Resources</u>
Khan Academy: Video on Protein synthesis
https://www.khanacademy.org/partner-content/nova/rnawondermolecule/v/proteinsynthesis

Fandom- Superpower Wiki, website that lists the different superpower and superhero or villain
http://powerlisting.wikia.com/wiki/Superpower_Wiki

Marvel- Different characters- discusses where or how they received their powers
https://marvel.com/characters
- Example of Incredible Hulk: https://marvel.com/universe/Hulk_(Bruce_Banner)

Plant Manipulation
http://powerlisting.wikia.com/wiki/Organic_Manipulation

Animal Imitation
http://powerlisting.wikia.com/wiki/Category:Organic_Mimicry

Electricity Manipulation
http://powerlisting.wikia.com/wiki/Electricity_Manipulation

Curriculum 6

BioBits™ Health ImageJ Tutorial: Quantifying the Fluorescence of Cell-Free Reactions

Now that you've completed your cell-free expression of fluorescent proteins, you can take a picture of your samples to document your results. You will likely notice that some of your reactions are more brightly fluorescent than others, which is a result of more of the fluorescent protein, dTomato being produced in the cell-free reactions.

Although you can see that some tubes look brighter than others, it can be difficult to tell the differences between the reactions. Was more or less dTomato produced in one reaction versus another? We can tell this from the brightness of the tubes: a tube with more dTomato will be brighter than a tube with less. But even this can be challenging—is one reaction twice as bright as another? Perhaps 3 or 5 times as bright? How do we tell?

In practice, scientists **quantify** (measure) their samples to accurately analyze the results of their experiments. Here, we will use the image taken of your samples to measure and express the brightness of each tube as a number, allowing us to easily and accurately compare the brightness of one tube to another. To do so, we will use ImageJ, a program developed by the National Institutes of Health that is freely available online. This is used worldwide by scientists every day to analyze images of their experiments. Today, we will be using ImageJ to measure **relative fluorescence units** (brightness) values for your reactions so that you can plot them on a graph.

1. Download ImageJ at the link below:
   https://imagej.nih.gov/ij/download.html

2. Load ImageJ by clicking on the ImageJ icon. The main program window looks like this:
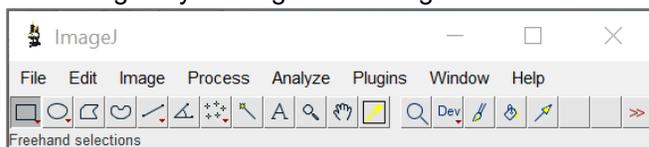
3. <u>Open an image</u>. Open your image by selecting from the drop-down menu **File > Open** and selecting your image. In this tutorial, we are using an example image:

4. Navigate in ImageJ. Some quick tips on navigating – to zoom in on your image, press **Ctrl and the + key**. To zoom out, press **Ctrl and the - key**. To pan the image, hold down the **Spacebar** while dragging with your mouse. Go ahead and zoom out, pan the image, and resize the window to center it on the fluorescent cell-free reactions in the tubes.
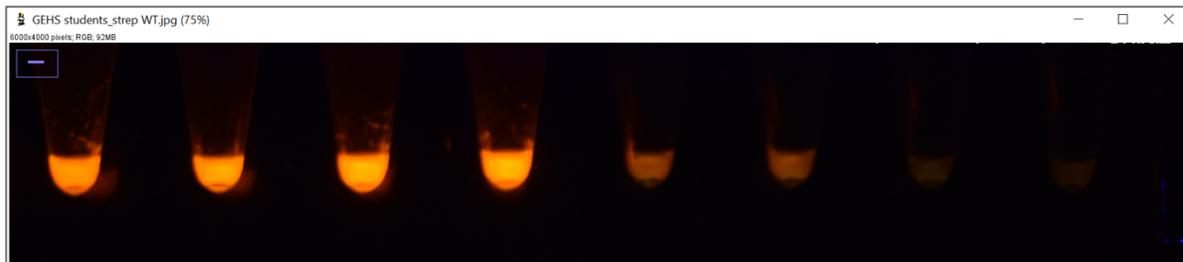
5. Convert to grayscale. Since we have a color image, we will simplify it for our analysis by converting it to a grayscale image. From the dropdown menu, select **Image > Type** and select **8-bit**. The image will now be converted to a black and white grayscale image, where each pixel has a numerical value from 0-255.
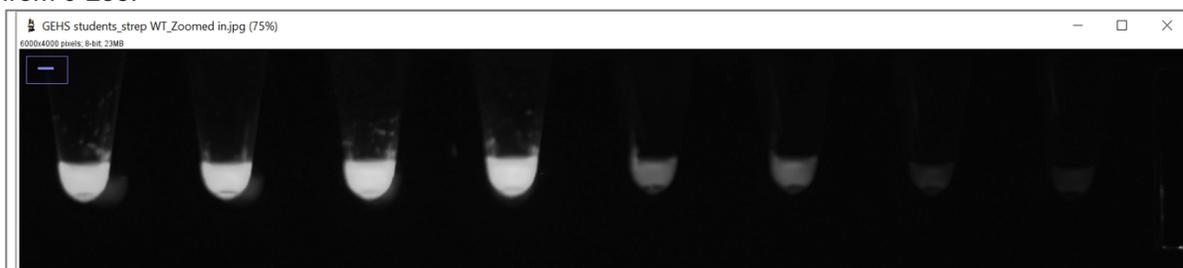
6. Select Region of Interest (ROI). We are going to select the region of the image for each tube that we want to quantify, starting from the first tube on the left. On the main menu, there are icons with shapes. These allow you to select regions of the image using a particular shape selection (a rectangle, circle, etc.). Click on the one shaped like an oval.
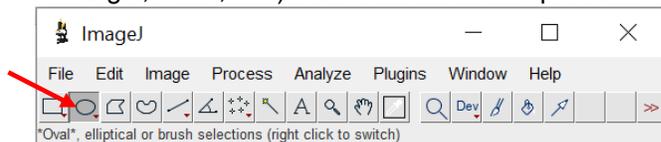
7. <u>Make and move a selection circle</u>. With the oval shape icon selected, if you hold down your left mouse key and drag on the image, you'll see an oval selection appear. We recommend holding down the **shift key** while doing this, which tells ImageJ to make the selection a perfect circle. Once you make a circle selection, you can move it around the image by moving your mouse cursor into the circle until the cross turns into an arrow:

Moving a selection circle.

If your place your mouse cursor right on top of the dotted line, it will turn into a hand, and this will allow you to resize the selection area. Again, hold down the **shift key** to keep the selection a perfect circle:

Resizing a selection circle.

8. <u>Make a selection circle small enough and move it so that it completely fits inside the cell-free reaction part of each tube</u>. We will be acquiring an average of every pixel within the selection area, so it's important that the selection does not include areas outside of the tube. If this happens, the black pixels will also be included in the average, which we do not want.
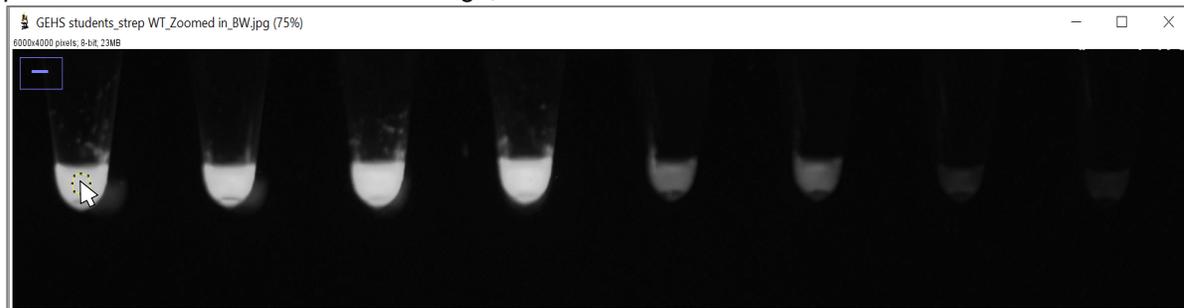
Selection region is moved to the inside of the first cell-free reaction tube.

9. <u>Save the selected region</u>. In the ImageJ dropdown menu, select **Analyze > Tools > ROI Manager**. A new window should pop up:
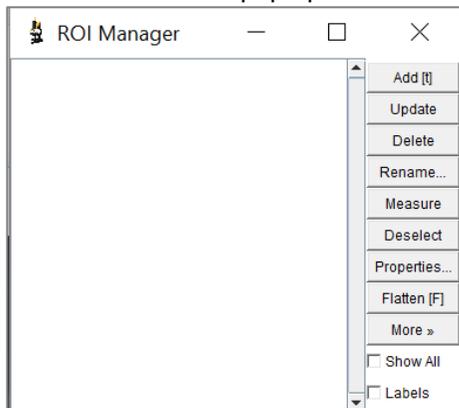
The ROI Manager window.

Click on the **Add [t]** button on the right menu of the ROI Manager window. An entry will appear; this is a record of the current selection circle area.
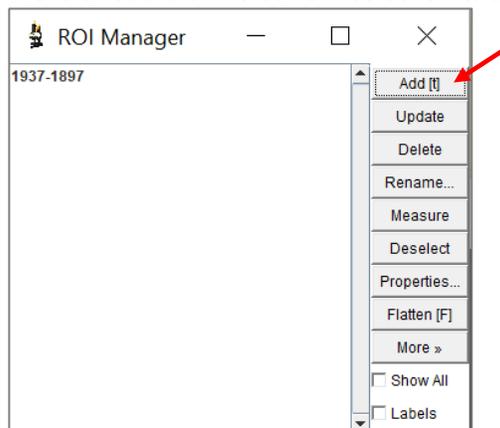
Adding the current selection circle to the ROI Manager.

The title of the entry is a series of numbers that represent the pixel coordinates of your selection. Let's rename that to something that makes more sense to us. Click on the ROI entry to select it and then click the **Rename** button on the right menu of the ROI Manager window to enter a new name. Here we have named it "Tube 1." You could also use the letter that is on the side of the reaction tube ("A", "B", "C", or "D"). It is important to keep track of which tubes are which so that you can plot them later!
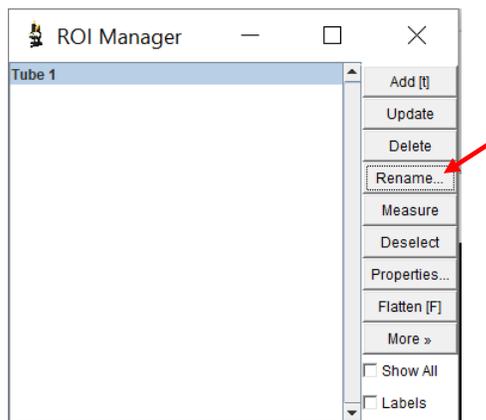
First ROI renamed to 'Tube 1'.

10. Go back to the main image window. <u>Move the selection circle to the inside of the next tube and repeat step 9</u>. Repeat this step for all of the tubes in your image.

11. After the last tube, you should have eight ROI entries. You'll notice if you click on any of the entries in the ROI Manager window, the corresponding selection circle will show up in the main image window.
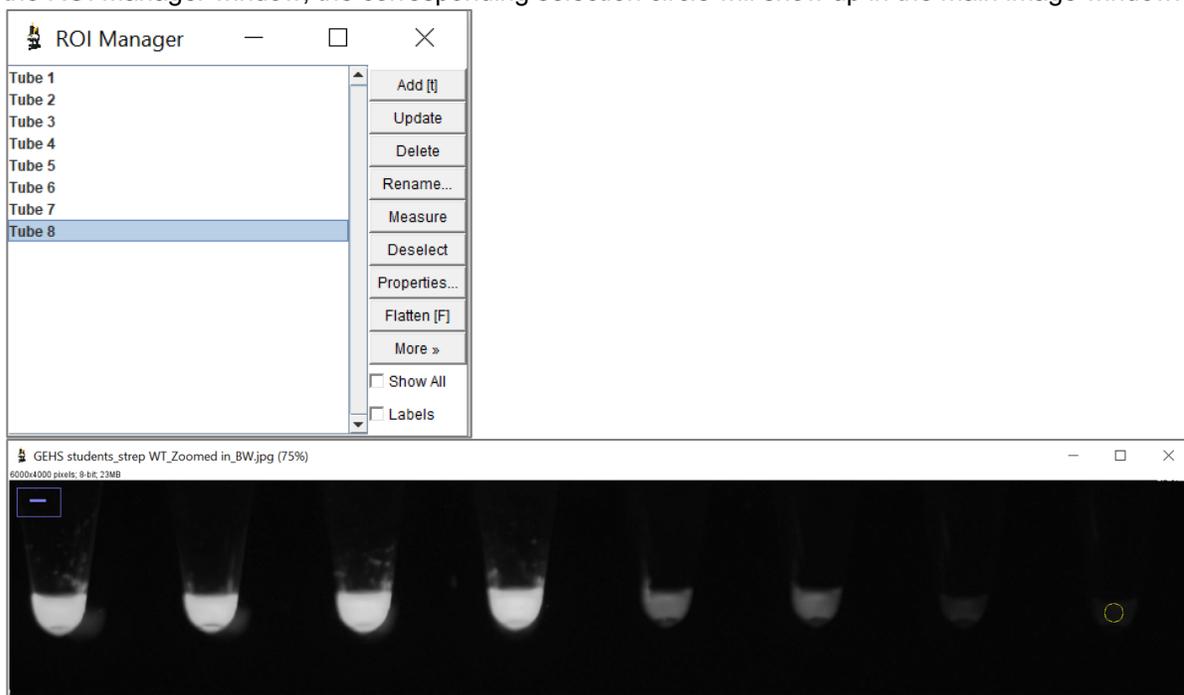
All 8 tube selection regions added to the ROI Manager.

12. Measure average intensities for each ROI. In the ROI Manager window, press **Ctrl-A** to select all eight entries. This should highlight all ROI selections. Now, on the right menu of the ROI Manager window, click the **Measure** button.
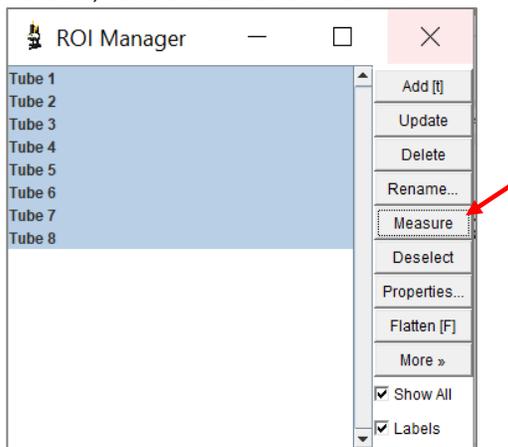
All ROI entries selected, click Measure.

A new window, the Results window, should pop up. It displays measured values for each ROI, such as the calculated Area, Maximum Pixel Intensity, Minimum Pixel Intensity, and many other values. Currently, it is displaying default settings:
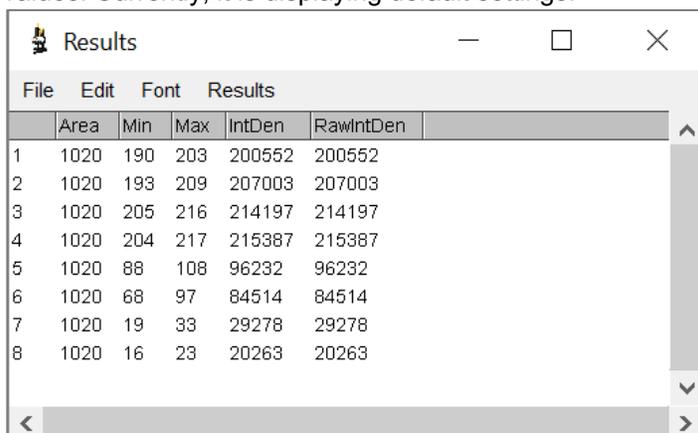
The Results window with default settings.

13. <u>Tell ImageJ which values you want to see in the Results window</u>. Let's change the settings of the Results window to view the measurements we are interested in. In the Results window, select from the drop-down menu **Results > Clear Results** to get rid of the currently displayed results. Then select **Results > Set Measurements** to bring up a new window that allows you to control what is being calculated. Unclick all the boxes, then click on the **Mean gray value** box (this calculates the average pixel brightness value for each ROI) and the **Display value** box (this names the results using the ROI Window entry labels) and press the OK button.
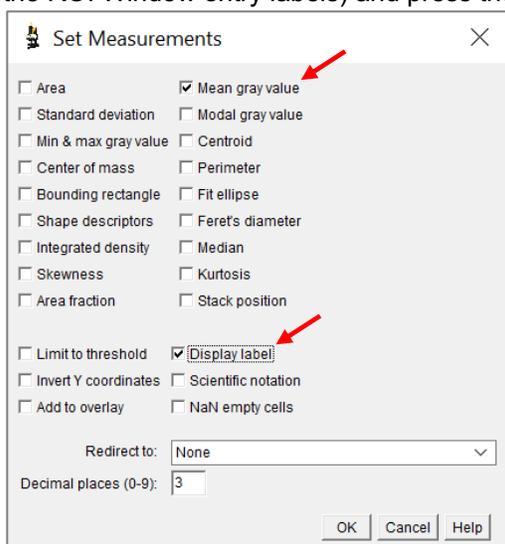
14. <u>Reupdate measurements</u>. Go back to the ROI Manager window, and with all the entries still selected, press the **Measure** button again. The Results window will update, and this time, only the values we selected in the Set Measurement options window are displayed. The 'Label' column contains the name of the file and the name we gave each ROI entry. The 'Mean' column contains the mean gray value for each ROI, a quantitative measure of the fluorescence intensity for each cell-free reaction.

15. <u>Save your measurements</u>. To save your measurements, in the Results window, select from the drop-down menu **File > Save as**, and choose a directory location and name for your results. This will create an Excel-compatible file (comma delimited) of your results. Alternatively, can select all of the data in the Results window (press the keys **Control-A** or simply select all using your mouse), copy (**Control-C**) and paste the values (**Control- V**) directly into Excel or another graphing program to create a quantitative graph of your experiment.
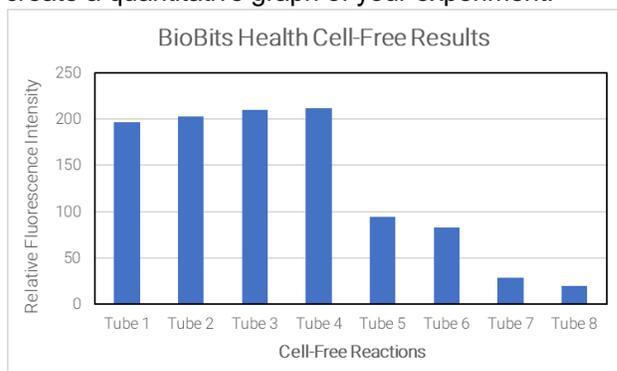
Data from ImageJ measurements plotted in Excel.

Teacher Notes:

Designed to support a lab activity developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University), Rachel Dubner (B.S. Candidate, Biology and Psychology, Northwestern University), and Karen Hsu (B.S. Candidate, Mechanical Engineering, Northwestern University). Written by Dr. Peter Q. Nguyen (Technology Development Postdoctoral Fellow, Wyss Institute and Harvard University) and Jessica Stark.

Curriculum 7

biobits

Exploring antibiotic resistance with cell-free reactions

BioBits™ Health

Outline

• **What is cell-free protein synthesis?**

• Investigate how antibiotics work

• Explore mechanisms of antibiotic resistance

biobits   BioBits™ Health   2

Synthetic biology seeks to engineer biology to benefit society

Using DNA…
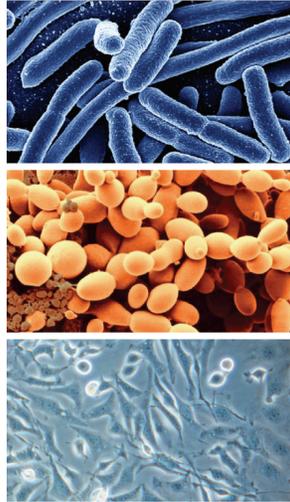
…to sustainably make medicines, chemicals, food, or fuel!

…to program biological systems…

biobits   BioBits™ Health   3

"Tug of war" between cell's goals and engineer's goals

Cell's goal:
- Grow
- Divide
- Adapt/evolve

Synthetic biologist's goal:
- Make non-native protein products
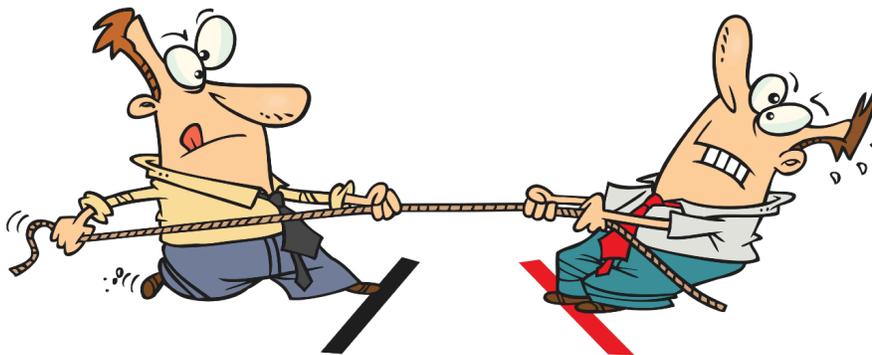
biobits   BioBits™ Health   4

Cell-free protein synthesis: Making proteins without cells

Advantages of CFPS:
- Direct access to biological system
- Add or remove components easily
- DNA → protein overnight!

cell growth

cell lysis

CFPS Substrates
NTPs
template DNA
amino acids
energy source

Legend
- genomic DNA
- ribosomes
- membrane vesicles
- soluble proteins

extract preparation

crude cell lysate

biobits   BioBits™ Health   5
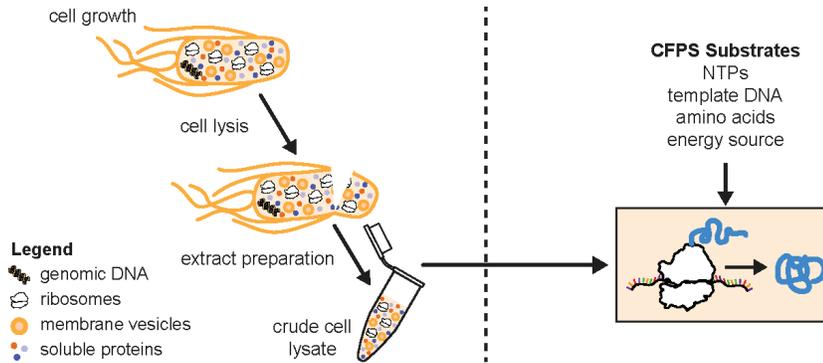
Freeze-dried cell-free (FD-CF) reactions contain machinery for protein synthesis

Transcription

DNA + RNA polymerase → mRNA

Translation

mRNA + Ribosome → Protein

Metabolism

Substrate + Enzyme → Enzyme + Product

biobits   BioBits™ Health   6
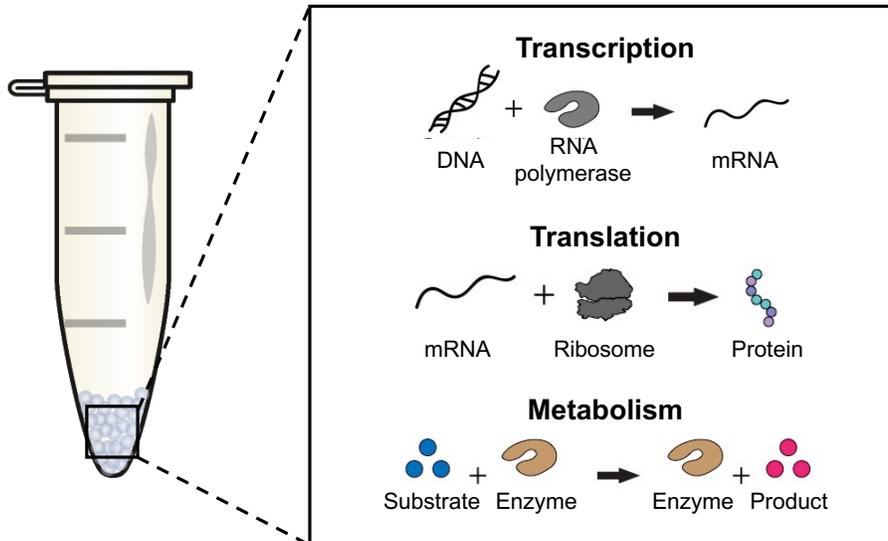
Fluorescent proteins help scientists measure protein synthesis
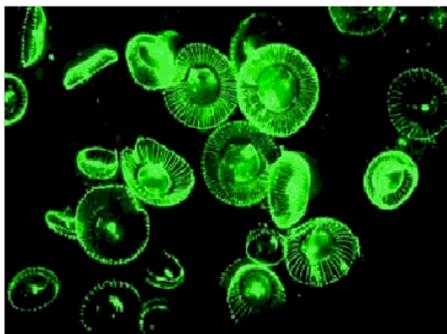
Green fluorescent protein (GFP) from the jellyfish *Aequorea victoria*

Fluorescent proteins allow us to quantify protein production by measuring **relative fluorescence units** (RFUs)

more protein = more fluorescence
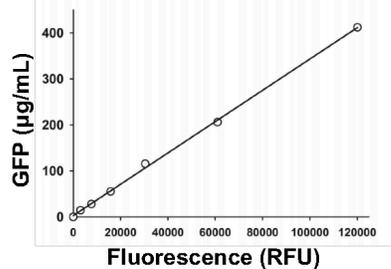
Hong, et.al. *ACS Synth. Biol.* (2014)

biobits   BioBits™ Health   7

You will make fluorescent proteins in FD-CF reactions

We will use expression of a fluorescent protein, dTomato, to measure the effects of antibiotics on protein synthesis in FD-CF reactions.

biobits   BioBits™ Health   8

Outline

- What is cell-free protein synthesis?

- **Investigate how antibiotics work**

- Explore mechanisms of antibiotic resistance

biobits   BioBits™ Health   9

Antibiotic resistance is a growing global issue

Resistance of *Escherichia coli* to fluoroquinolones

% Resistant
100
75
50
25
0
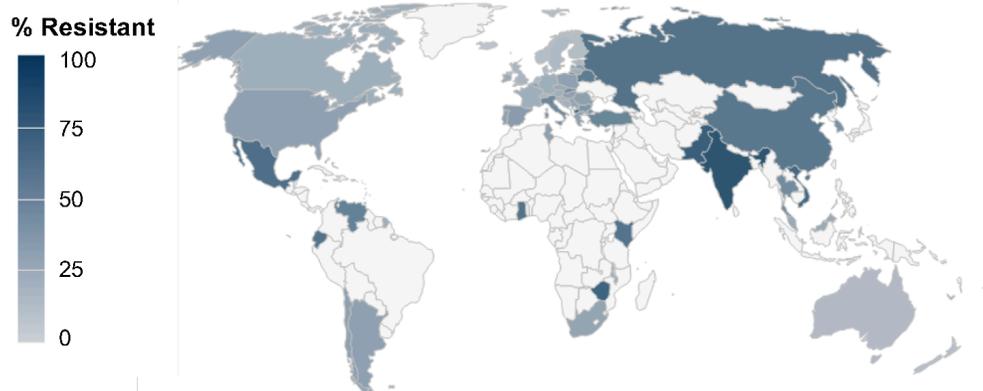
The Center for Disease Dynamics, Economics, and Policy; O'Neill, et. al. *The Review on Antimicrobial Resistance* (2014)

To solve this problem, we need to understand how antibiotics work and how resistance occurs

biobits   BioBits™ Health   10

You will research how antibiotics disrupt essential cell functions
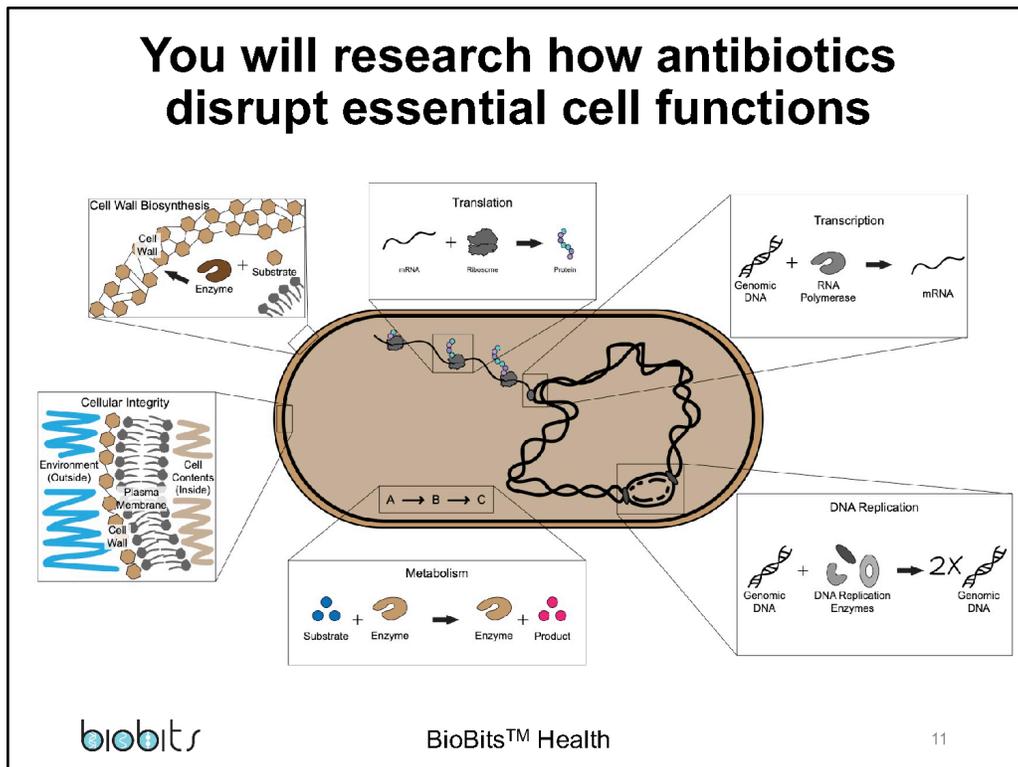
One example: Some antibiotics work by inhibiting the ribosome

Major classes of protein synthesis–inhibiting antibacterials

Chloramphenicol, macrolides, and lincosamides
- Bind to the 50S ribosomal subunit
- Prevent peptide bond formation
- Stop protein synthesis

Aminoglycosides
- Bind to the 30S ribosomal subunit
- Impair proofreading, resulting in production of faulty proteins

Tetracyclines
- Bind to the 30S ribosomal subunit
- Block the binding of tRNAs, thereby inhibiting protein synthesis

https://courses.lumenlearning.com/microbiology/chapter/mechanisms-of-antibacterial-drugs/
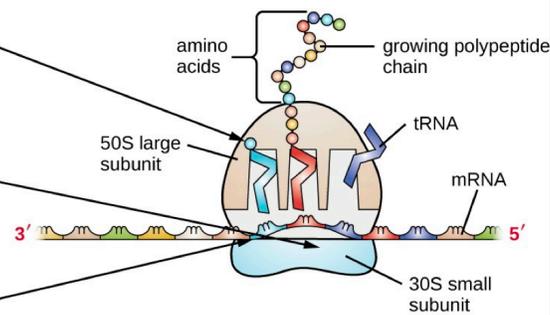
What do you think will happen if we add these antibiotics to FD-CF reactions?

You will test varying amounts of antibiotics in FD-CF reactions

Dilute the antibiotics to get varying amounts

Add to FD-CF reactions

Measure the effects on protein synthesis
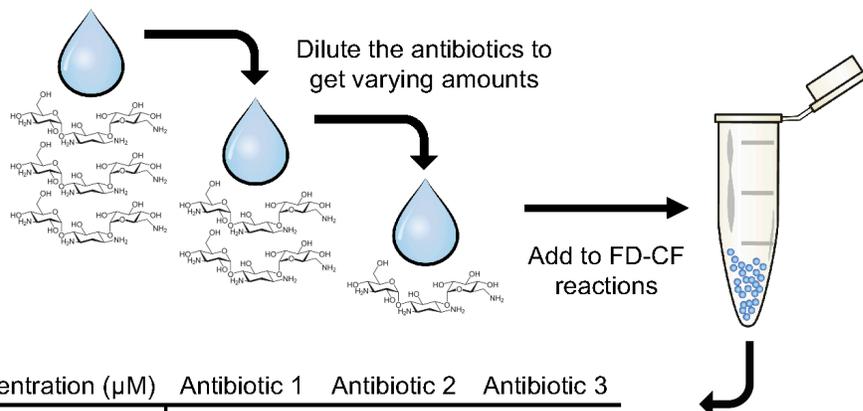
Concentration (μM)		Antibiotic 1	Antibiotic 2	Antibiotic 3
low	0.1			
↓	1			
	10			
high	100			

biobits

BioBits™ Health

13

Outline

• What is cell-free protein synthesis?

• Investigate how antibiotics work

• **Explore mechanisms of antibiotic resistance**

biobits

BioBits™ Health

14

269

Test the effect of antibiotics on protein synthesis in 3 types of FD-CF reactions

Dilute the antibiotics to get varying amounts

Add to FD-CF reactions

Measure the effects on protein synthesis

Concentration (µM)	Scenarios		
	Wild-type	Resistance gene	Mutation
low 0.1			
↓ 1			
10			
high 100			

biobits   BioBits™ Health   17

You will research ways that you can help prevent antibiotic resistance!

Prescribe antibiotics only for bacterial infections and always finish a course of antibiotics when prescribed

Limit the use of antibiotics in agriculture by using them only on sick animals

Limit use of antibacterial soaps and other consumer products that could cause resistance

https://www.cdc.gov/antibiotic-use/community/about/antibiotic-resistance-faqs.html
https://www.cdc.gov/narms/faq.html
https://www.scientificamerican.com/article/strange-but-true-antibacterial-products-may-do-more-harm-than-good/

How else could we address this societal problem?

biobits   BioBits™ Health   18

Curriculum 8

Name _____

Cell-Free Protein Synthesis and Antibiotics

Purpose:
- Students will explore how antibiotics interact with structure and function of cellular components.
- Students will conduct cell-free protein synthesis in the presence of antibiotics and measure the expression of the fluorescent protein dTomato.
- Students will process the data collected using ImageJ and Excel.
- Students will demonstrate understanding of antibiotics by hypothesizing a new antibiotic that targets a cellular component not involved in the cell-free protein synthesis system.

Background:
Antibiotics are a group of medicines that fight bacterial infections. They work by interfering with various structures of bacterial cells which can inhibit growth of the bacteria or even kill the bacteria completely. When used properly, antibiotics save lives. Unfortunately, some bacteria have developed ways to destroy or avoid the antibiotic, so the antibiotic cannot kill the bacterial cell. This is called antibiotic resistance, and it is a growing problem world-wide. Experts predict that by the year 2050, more people will die from antibiotic-resistant bacterial infections than cancer [O'Neill, et al. *The Review on Antimicrobial Resistance* (2014)].

But how do antibiotics work? For this activity, imagine that you are a scientist at a pharmaceutical company. You have been assigned to investigate the mechanisms of action of antibiotics that the company is developing to combat antibiotic-resistant bacteria. Once you have determined how these antibiotics work, you will use what you've learned to propose a new antibiotic that could be used to treat bacteria that have become resistant to the antibiotics you have already characterized.

Resistance of *Escherichia coli* to fluoroquinolones

Source: The Center for Disease Dynamics, Economics, and Policy Resistance Map
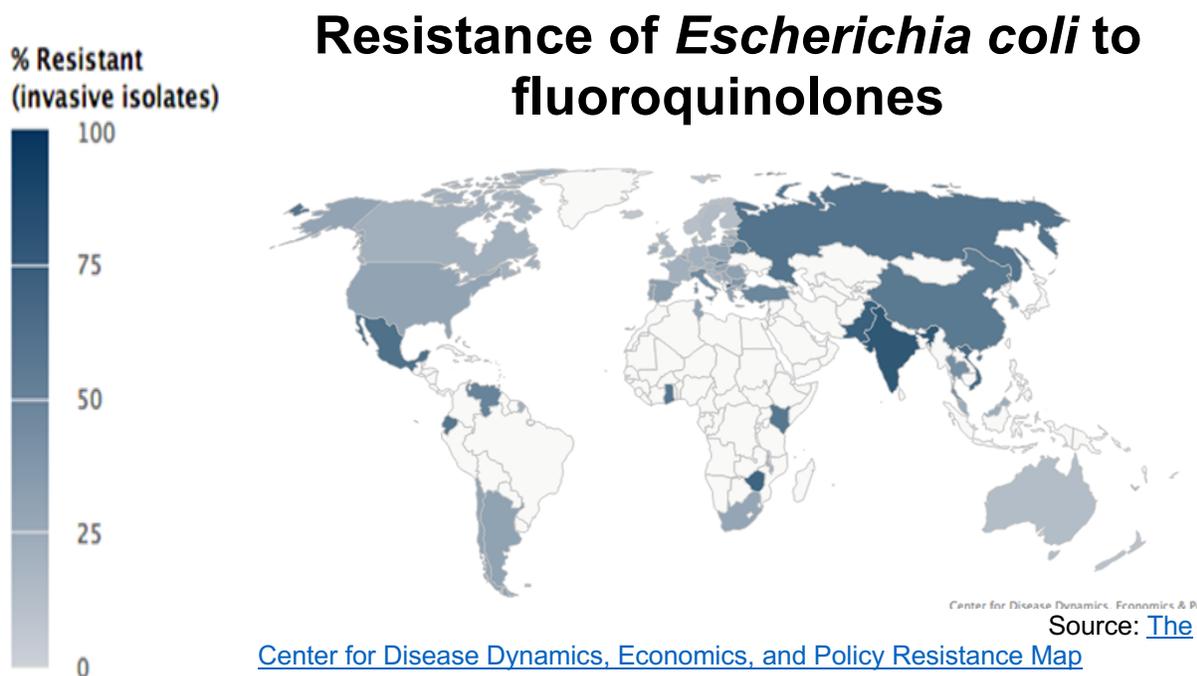
Pre-lab Activity:

First, to familiarize yourself with different types of antibiotics, research the following antibiotics: carbenicillin, streptomycin, rifamycin, polymyxin B, & trimethoprim. Fill in the table provided with how the antibiotic works, the processes it disrupts, and the specific target within that process. Make sure to provide the sources that you used.

Antibiotic	How does the antibiotic work? Describe in a few sentences.	What process does the antibiotic disrupt?	What cellular component does the antibiotic target?
Carbenicillin			
Streptomycin			
Rifampin			
Polymyxin B			

Trimethoprim			

Now that you have learned a variety of different ways that antibiotics can kill bacteria, choose one antibiotic to test in the lab. Ask your teacher which antibiotics are available for you to test. Write down your choice of antibiotic below. How does your antibiotic work? What cellular process does the antibiotic disrupt? What cellular component does the antibiotic target?

Your antibiotic	How does the antibiotic work? Describe in a few sentences.	What process does the antibiotic disrupt?	What cellular component does the antibiotic target?

Procedure:

Exploring antibiotic mechanisms of action with cell-free protein synthesis lab activity

Materials needed:
- Antibiotics of choice (use provided powder stocks or others supplied by the user)
- FD-CF reactions in PCR tube strips (one 8-well strip per group/antibiotic; dTomato plasmid is already included in the FD-CF reactions for this module)
- Nuclease-free water (provided)
- Sterile pipet tips and micropipettes
- Microcentrifuge tubes for antibiotic dilution
- PCR thermocycler OR BioBits™ portable incubator
- Blue light imager OR BioBits™ portable imager

DAY 1
1. Prepare a 500 µM stock solution of your antibiotic by mixing the provided antibiotic powder from the BioBits™ Health kit with 500 µL of nuclease-free water. Vortex or pipet the solution up and down to mix.
2. *Optional: These stocks can be prepared ahead of time and refrigerated.*
3. Label four empty microcentrifuge tubes "A," "B," "C," and "D." These will be your antibiotic dilutions.
4. Add the appropriate amount of water (refer to the table below) to each of the dilution tubes.
   a. You can use the same pipet tip for each water addition as long as you do not touch any surface other than the inside of the tubes! When you start pipetting the antibiotic later you'll need to switch pipet tips, but not yet.

Tube	Antibiotic Concentration	Water (µL)	Antibiotic Source	Antibiotic (µL)
A	100 µM	16	500 µM stock	4
B	10 µM	18	Dilution A	2
C	1 µM	18	Dilution B	2
D	0.1 µM	18	Dilution C	2

5. Refer to the table above to add the appropriate amount of antibiotic to each of the dilution tubes. Remember to switch pipette tips in between tubes. Vortex or pipet up and down several times to mix each dilution tube.
6. Label your PCR tubes containing the CFPS reactions. The first two tubes in the strip should be labeled "A," the next two "B," the next two "C," and the last two "D." Label these on the side of the PCR tubes. Put the initials of your group members or your group number on the SIDE of the strip.
7. Being careful to *avoid touching the white pellet* at the bottom, add 5 µL of the antibiotic/water mixture from dilution tube A to the first CFPS reaction tube labeled "A." Add the same amount to the second "A" tube.
8. Repeat step 6 for your "B", "C", and "D" PCR tubes with dilution tubes B, C, and D.
9. <u>INSTRUCTOR</u>: add 5 µL of nuclease free water to each of 2 additional freeze-dried CFPS reactions. These will serve as the no antibiotic controls (negative controls) for your experiment.
10. Seal your CFPS reactions using an 8-strip PCR tube lid. Gently flick the side of the PCR tubes to dissolve the freeze-dried CFPS pellet in the antibiotic/water mixture that you added.

11. Allow the reactions to incubate until the next time you come to class. Reactions can be incubated in the BioBits™ portable incubator or a PCR thermocycler for 20-24 hours at 30°C or on the benchtop at room temperature for 24-48 hours. Get excited!

DAY 2

12. Obtain your group's samples from your teacher. Place your strip of reactions in front of a piece of white paper and record your observations.
13. Take a picture of your reactions under blue light (using the BioBits™ blue light imager or other blue light source) and record your observations.
14. Quantify fluorescence produced in each reaction with ImageJ, using the accompanying ImageJ tutorial. Record the relative fluorescence unit (RFU) values you measure in ImageJ in the table below.
15. Calculate the average and standard error of the RFU values you measured using Excel. Record these values in the table below.
16. Analyze images from other groups' reactions or share your data with other groups to compile data for all antibiotics tested. Also quantify RFU values for the negative control reactions (no antibiotic added).
17. Plot the average RFU values you calculated as a function of antibiotic concentration for each antibiotic. Choose an appropriate graph type (bar, scatter, etc.) to display your results. Add error bars showing the standard error of your measurements.
18. *Optional: Use the "t-test: Paired Two Sample for Means" function in Excel to determine if the addition of antibiotic creates statistically significant differences in fluorescence measured. The following tutorial can help you get started:*
    https://www.rwu.edu/sites/default/files/downloads/fcas/mns/running_a_t-test_in_excel.pdf

Data:

Relative Fluorescence Units (RFU) vs Antibiotic Concentration

Negative controls (no antibiotic)				
Concentration	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
0 µM				

Group 1 Antibiotic:				
Concentration	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
100 µM				
10 µM				
1 µM				
0.1 µM				

Group 2 Antibiotic:				
Concentration	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
100 µM				
10 µM				
1 µM				
0.1 µM				

Group 3 Antibiotic:				
Concentration	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
100 µM				
10 µM				
1 µM				
0.1 µM				

Post-lab Analysis:

1. Which antibiotics prevented protein synthesis of dTomato? Which did not?

2. Draw a model of the cell-free system without antibiotic interference. Make sure to include transcription, translation, ribosomes, plasmid DNA, mRNA, T7 RNA polymerase & dTomato in your model. *Hint: think back to our pre-lab lecture!*

3. Using the figure on the next page, draw a model of each cellular process indicated. Make sure to include Genomic DNA, mRNA, ribosomes, protein, plasma membrane, cell wall, enzymes, substrates, products, RNA polymerase, & DNA replication enzymes in your answer.

DNA Replication

Translation

Metabolism

A → B → C

Cell Wall Biosynthesis

Cellular Integrity

4. Looking at your answers to questions 2 & 3, what cellular components and processes are missing in a cell-free system compared to a living bacterial cell?

5. Why do you think that not all antibiotics used in this lab affected protein synthesis?

6. Research the other antibiotics used in the lab. Explain the results you and your classmates collected for each antibiotic, using question #4, the pre-lab activity, and new research as evidence. Why did protein synthesis take place or not take place in the presence of each antibiotic in this lab?

7. Reflect on what you have learned in this activity. What makes a good antibiotic? How do they work to slow bacterial growth or kill bacteria? Propose a new antibiotic that targets a part of the <u>bacterial cell</u> (doesn't have to be a part of the cell that is included in the cell-free system) that is different from the targets of the antibiotics you tested in the lab. What process would it disrupt? What does it target? Draw a model of the process and how your antibiotic would interact with the bacterial cell to kill the cell or inhibit its growth. Write an explanation of your model in words.

Answer Key

Pre-lab Activity:

Antibiotic	How does it work? Describe in a few sentences.	What process does it disrupt?	What does the antibiotic target?
Carbenicillin	Bacteria have a cell wall to protect the cell from the environment. Carbenicillin inhibits cell wall synthesis. This makes the cell more fragile and likely to die (lyse).	Cell wall biosynthesis	Proteins that build the cell wall (peptidoglycan binding proteins)
Streptomycin	Bacteria need to produce proteins in order to grow and survive. Streptomycin binds to the ribosome, preventing it from making proteins that function correctly. This makes the cell unable to grow or maintain itself, so the cell dies.	Protein synthesis (translation)	Ribosome (30S subunit)
Rifampin	Bacteria must produce mRNA in order to make protein. Rifampin inhibits RNA polymerase. This makes the cell unable to make RNA or new proteins and kills the cell.	RNA synthesis (transcription)	RNA Polymerase
Polymyxin B	Bacteria have a plasma membrane to separate the cell from the environment. Polymyxin B interacts with the plasma membrane and destabilizes it. This makes the cell unable to maintain the cell-environment boundary, killing the cell.	Cellular integrity	Plasma membrane
Trimethoprim	Bacteria need to consume molecules form the environment to build themselves. Trimethoprim inhibits folic acid synthesis, a key part of nucleic acid biosynthesis. This makes the cell unable to make new DNA & RNA, killing the cell.	Folic acid synthesis (metabolism)	Proteins that produce folic acid (enzymes)

Post-lab Analysis:

1. Will depend on which antibiotics were chosen and based on the data the students collected. Antibiotics that cause decreases in CTCF with increasing antibiotic concentration will have demonstrated inhibition of protein synthesis. Antibiotics that did not cause changes in CTCF with increasing antibiotic concentration did not demonstrate inhibition of protein synthesis.

2.
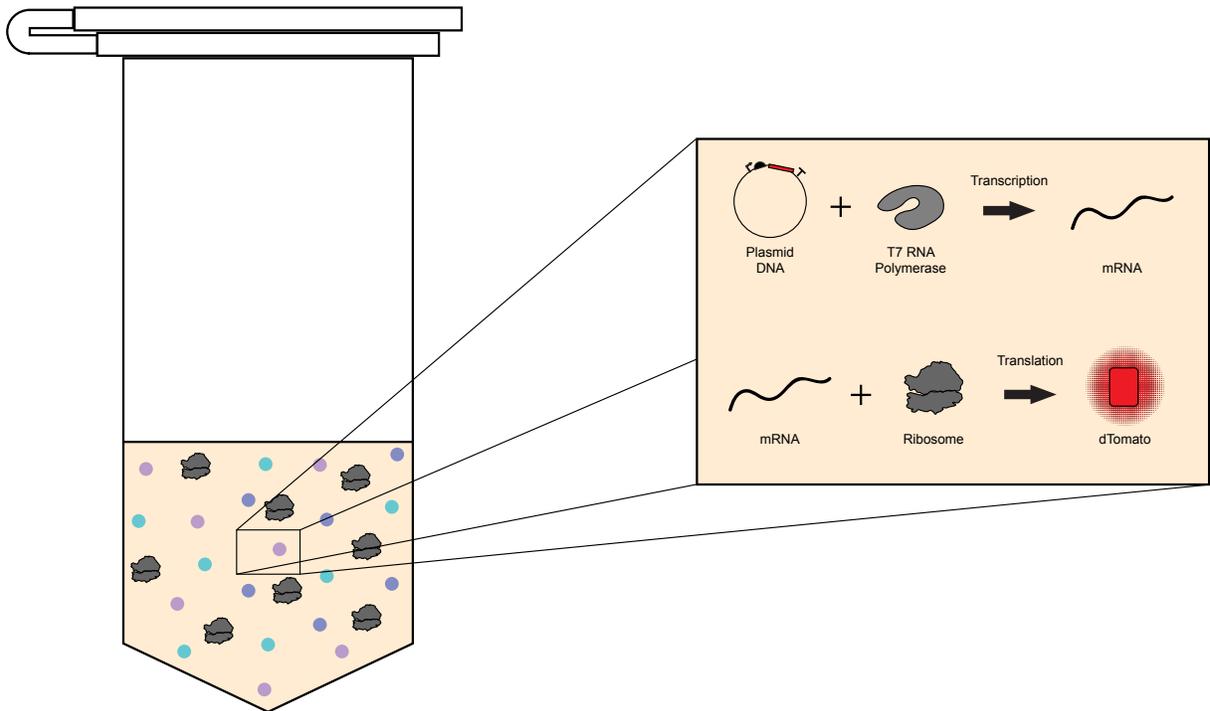
3.
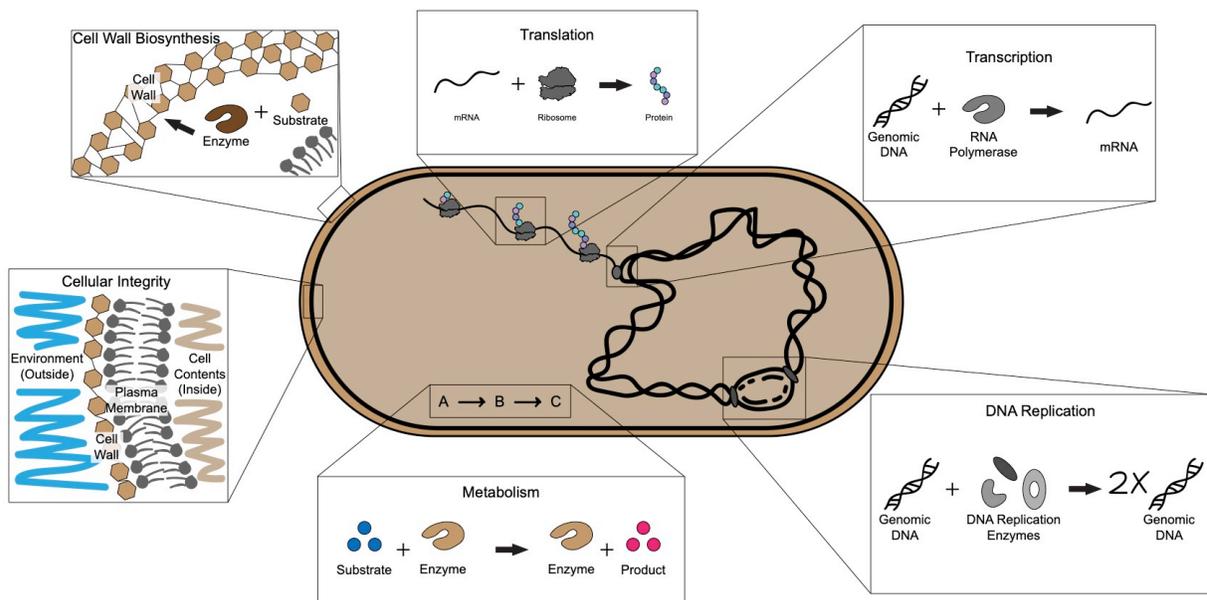
4. The cell-free system is missing a plasma membrane, cell wall, and DNA synthesis machinery. The cell-free system does not need to maintain cellular integrity (the plastic tube does that), be protected from the environment (again, the plastic tube does that), and does not need to make DNA (we add it in).

5. Not all of the antibiotics affected protein synthesis because not all of the antibiotics target pieces of the protein synthesis machinery. The mode of action of these antibiotics are likely inhibiting processes not modeled by the cell-free system.

6. Will depend on which antibiotics were chosen and based on the data the students collected. Decreases in CTCF with increases in antibiotic concentration are due to the antibiotic inhibiting protein synthesis, which results in less dTomato being produced. dTomato produces fluorescence. Constant CTCF with increases in antibiotic concentration are due to the antibiotic inhibiting other cellular processes than protein synthesis, which does not affect the amount of dTomato being produced.

7. A good antibiotic is able to kill or slow the growth of bacteria. Antibiotics work by inhibiting cellular processes within the bacteria and target specific components of the process. Proposals for new antibiotics can vary, but all should include 1) a bacterial process other than protein synthesis, 2) a drawn model of that process, 3) identification of a specific component in that process that the antibiotic will affect, & 4) explanation of how inhibiting that cellular process would slow bacteria growth or kill bacteria.

Antibiotic Resistance in Cell-Free Systems - Teacher Overview

Purpose:
- Students will explore how antibiotics interact with structure and function of cellular components.
- Students will conduct cell-free protein synthesis in the presence of antibiotics and measure the expression of the fluorescent protein dTomato.
- Students will process the data collected using ImageJ and Excel.
- Students will demonstrate understanding of antibiotics by hypothesizing a new antibiotic that targets a cellular component not involved in the cell-free protein synthesis system.

Background:
- How antibiotics work
- Development of antibiotic resistance and global health problem

Pre-lab activity:
- Introduce students to how antibiotics work
- Students research antibiotics including those they will use in the lab and identify their targets in the bacterial cell
- Antibiotics act on specific components and processes in the cell
- A useful resource for teachers to prepare for student responses to the pre-lab activity is: https://courses.lumenlearning.com/microbiology/chapter/mechanisms-of-antibacterial-drugs/

Antibiotic lab:
- Cell-free protein synthesis using dTomato
- Antibiotics with different cellular targets (ribosome, cell wall; for example). In the BioBits™ Health kit, we provide some antibiotics that are ribosome inhibitors (e.g., streptomycin, kanamycin) and will impact cell-free protein synthesis, as well as some that inhibit cell wall synthesis (e.g., carbenicillin) and will not impact cell-free protein synthesis. You can also choose to purchase and test other antibiotics beyond those provided. You can have each group use a different antibiotic or have a few groups use the same antibiotics to get more data points for calculating means and standard deviations or running statistical analyses in Excel.
- Vary concentration of antibiotics
- CFPS lab and measure relative color change (ImageJ)
- Optional use of biological statistics (t-test) to analyze fluorescence data

Post-lab analysis:
- Students discover that protein synthesis is not stopped by all antibiotics
- Why is that?
    - Antibiotics target different parts of the bacterial cell other than protein synthesis
    - The cell-free protein synthesis system does not use all of the bacteria's cellular components/processes
- Draw a model of the cell-free protein synthesis system
- Compare model of cell-free protein synthesis system to model of bacterial cell
- Revise a model of a bacterial cell to include an interaction of a new antibiotic that you designed to target a different part of the cell (not cell-free system)

Curriculum 9

Name _____

Exploring Antibiotic Resistance in Cell-Free Systems

Purpose:

- Students will explore how bacteria develop resistance to antibiotics using cell-free protein synthesis of the fluorescent protein dTomato.
- Students will process the data collected using ImageJ and Excel.
- Students will demonstrate understanding of how bacteria become resistant to antibiotics and will research behaviors and technologies that can help combat this problem.

Background:

Have you ever been to the doctor and been prescribed an antibiotic? You likely took the medicine as prescribed and felt better in a few days. However, that is not always the case. With increased antibiotic use in everything from agriculture to hygiene products, we are experiencing a global crisis: antibiotic resistance. A variety of bacteria that used to be treatable with antibiotics have developed resistance mechanisms that allow them to survive and even thrive in the presence of antibiotics. But how did this happen?

The two primary ways that bacteria develop resistance to antibiotics are through **horizontal gene transfer** or random **genetic mutation**. Horizontal gene transfer describes the process by which bacteria can acquire resistance genes from each other through a process called conjugation. Mutations in a bacteria's own genome (DNA) makes any bacteria with the mutation resistant to an antibiotic. Once bacteria have either acquired resistance genes or mutated their DNA to become resistant, they can continue to grow in the presence of antibiotics while wild-type bacteria are killed.

In this activity, you will use cell-free protein synthesis to simulate both horizontal gene transfer and genetic mutation, as well as research ways **you** can help prevent antibiotic resistance!

Antibiotics are used in many unexpected products, even pencils!
Image source: dixonticonderoga.com

Pre-lab Activity:

1. Recall that antibiotics kill or slow the growth of bacteria. How do antibiotics do this? Describe at least 3 ways.

2. Define antibiotic resistance in your own words.

3. Do some research to find an example of how bacteria can become resistant to antibiotics. List the relevant antibiotic and describe, in your own words, the change that took place in the bacteria that caused it to become resistant to that antibiotic. Include a link to your source.

4. From what you've learned in the pre-lab lecture, do you think the example you found is an instance of horizontal gene transfer or genetic mutation? Explain your reasoning.

Procedure:

Exploring mechanisms of antibiotic resistance with cell-free protein synthesis lab activity

Materials needed:
- Streptomycin (use provided powder stock)
- Large-scale (15 µL) FD-CF reactions in microcentrifuge tubes (10 µL of reaction mixture needed per group)
- Wild-type FD-CF reactions in PCR tube strips (two 8-well strips per group)
- Streptomycin-resistant FD-CF reactions in PCR tube strips (one 8-well strip per group)
- 13.33 ng/µL AadA plasmid DNA stock solution (provided)
- Nuclease-free water (provided)
- Sterile pipet tips and micropipettes
- Microcentrifuge tubes for antibiotic dilution
- PCR thermocycler OR BioBits™ portable incubator
- Blue light imager OR BioBits™ portable imager

DAY 0
1. Rehydrate large-scale FD-CF reactions by adding 15 µL of AadA plasmid DNA stock solution to the FD-CF reactions in microcentrifuge tubes. Close the lid and gently flick the side of the tube to dissolve the freeze-dried CFPS pellet. Reactions can be incubated in the BioBits™ portable incubator or a PCR thermocycler for 20-24 hours at 30°C. These reactions will pre-express the streptomycin resistance protein AadA that you will use in your experiment. <u>You will need to prepare 10 µL of AadA reaction mix per group.</u>
   This step can be done by the instructor or students.

DAY 1
2. Prepare a 500 µM stock solution of streptomycin by mixing the streptomycin powder from the BioBits™ Health kit with 500 µL of nuclease-free water. Vortex or pipet the solution up and down to mix.
   Optional: This stock can be prepared ahead of time and refrigerated.
3. Label four empty microcentrifuge tubes "A," "B," "C," and "D." These will be your antibiotic dilutions.
4. Add the appropriate amount of water (refer to the table below) to each of the dilution tubes.
   a. You can use the same pipet tip for each water addition as long as you do not touch any surface other than the inside of the tubes! When you start pipetting the antibiotic later you'll need to switch pipet tips, but not yet.

Tube	Antibiotic Concentration	Water (µL)	Antibiotic Source	Antibiotic (µL)
A	100 µM	0	500 µM stock	20
B	10 µM	18	500 µM stock	2
C	1 µM	18	Dilution B	2
D	0.1 µM	18	Dilution C	2

5. Refer to the table above to add the appropriate amount of antibiotic to each of the dilution tubes. Remember to switch pipette tips in between tubes. Vortex or pipet up and down several times to mix.
6. Label your PCR tubes containing the CFPS reactions with which extract it contains and if it will have AadA added. The first two tubes in the strip should be labeled "A," the next two "B," the next two "C," and the last two "D." Label these on the side of the PCR tubes. Put the initials of your group members or your group number on the SIDE of the strip.
7. The table below shows the volumes of antibiotic, AadA reaction mixture, and water you will add to each types of FD-CF reaction. Start with your two "A" tubes of wild-type FD-CF reactions. Being careful to *avoid touching the white pellet* at the bottom, add 1 µL from your dilution tube A to the first CFPS reaction tube labeled "A" as indicated in the table below. Add the same amount to the second "A" tube.

Resistance Mechanism	FD-CF reactions	Antibiotic dilution (µL)	AadA FD-CF reaction (µL)	Water (µL)
None	Wild-type	1	0	4
Genetic mutation	Resistant	1	0	4
Horizontal gene transfer	Wild-type	1	1	3

8. Again, being careful to *avoid touching the white pellet* at the bottom of the PCR tubes, add 4 µL of water to both "A" tubes.
9. Repeat steps 7 and 8 for your B, C, and D wild-type *and* A, B, C, and D resistant FD-CF reactions using your A, B, C, and D antibiotic dilution tubes.
10. For the reactions that will contain AadA, repeat step 7, adding 1 µL of the appropriate antibiotic stock to your A, B, C, and D PCR tubes.
11. Next, add 1 µL of AadA FD-CF reaction mixture to all 8 of the AadA tubes. The FD-CF mixture is viscous and can be hard to pipet, but do your best to transfer all of the 1 µL volume to the PCR tubes. As always, *avoid touching the white pellet* at the bottom of the tube.
12. Add 3 µL of water to all AadA tubes.
13. <u>INSTRUCTOR</u>: add 5 µL of nuclease free water to each of 2 additional freeze-dried CFPS reactions. These will serve as the no antibiotic controls (negative controls) for your experiment.
14. Seal your CFPS reactions using an 8-strip PCR tube lid. Gently flick the side of the PCR tubes to dissolve the freeze-dried CFPS pellet in the antibiotic/water/AadA mixture that you added.
15. Allow the reactions to incubate until the next time you come to class. Reactions can be incubated in the BioBits™ portable incubator or a PCR thermocycler for 20-24 hours at 30°C or on the benchtop at room temperature for 24-48 hours. Get excited!

DAY 2

16. Obtain your group's samples from your teacher. Place your strip of reactions in front of a piece of white paper and record your observations.
17. Take a picture of your reactions under blue light (using the BioBits^{TM} blue light imager or other blue light source) and record your observations.
18. Quantify fluorescence produced in each reaction with ImageJ, using the BioBits^{TM} Health ImageJ tutorial. Record the relative fluorescence unit (RFU) values you measure in ImageJ in the table below.
19. Calculate the average and standard error of the RFU values you measured using Excel. Record these values in the table below.
20. Plot the average RFU values you calculated as a function of antibiotic concentration for each antibiotic. Choose an appropriate graph type (bar, scatter, etc.) to display your results. Add error bars showing the standard error of your measurements.
21. *Optional: Use the "t-test: Paired Two Sample for Means" function in Excel to determine if the resistant reactions have statistically significant differences in fluorescence compared to wild-type reactions for various antibiotic concentrations. The following tutorial can help you get started:* https://www.rwu.edu/sites/default/files/downloads/fcas/mns/running_a_t-test_in_excel.pdf

Data:

Relative Fluorescence Units (RFU) vs Antibiotic Concentration

Wild-Type FD-CF				
Concentration	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
100 µM				
10 µM				
1 µM				
0.1 µM				

Wild-Type FD-CF + AadA Protein				
Concentration	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
100 µM				
10 µM				
1 µM				
0.1 µM				

Streptomycin-Resistant FD-CF				
Concentration	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
100 µM				
10 µM				
1 µM				
0.1 µM				

Post-lab Analysis:

1. Draw a model of how a cell could become resistant to an antibiotic through horizontal gene transfer. Make sure to include an antibiotic resistance gene, a donor bacterium, a recipient bacterium, an antibiotic resistant bacteria population, an antibiotic sensitive bacteria population, a timeline and antibiotic in your answer. *Hint: think back to our pre-lab lecture!*

2. Draw a model of how a cell could become resistant to an antibiotic through random mutation. Make sure to include an antibiotic target gene, mutant antibiotic target gene, an antibiotic resistant bacteria population, an antibiotic sensitive bacteria population, a timeline and antibiotic in your answer. *Hint: think back to our pre-lab lecture!*

3. In the lab, you investigated 3 types of FD-CF reactions. Which of these simulated horizontal gene transfer? Which simulated genetic mutation? Which (if any) were experimental controls? Explain how each experimental setup simulates each mechanism of resistance.

4. Research solutions to the antibiotic resistance crisis. List and explain two solutions, other than reducing the amount of antibiotics society uses. Give an example of one behavior and one technology that could help prevent antibiotic resistance. Explain how these could be implemented in your life.

Answer Key

Pre-lab Activity:
1. Antibiotics kill or slow the growth of bacteria by inhibiting cellular processes needed for growth or replication. Antibiotics specifically target a cellular component within the process and cause it not to work, breaking the process. Without the intact cellular process, the bacteria cannot replicate or live.
2. Antibiotic resistance is the ability of a bacterium to no longer be affected by an antibiotic that affected it at one point in time. The bacterium can now survive and multiply in the presence of the antibiotic.
3. This answer will vary depending on the resource students found. They may find an example of a specific genetic mutation or a description of an efflux pump or enzyme that can degrade antibiotics. This article provides some good examples: https://www.reactgroup.org/toolbox/understand/antibiotic-resistance/resistance-mechanisms-in-bacteria/
4. Again, this will depend on the answer to question #3, but a change in sensitivity to the antibiotic can arise through mutation. Acquiring new proteins (degrade antibiotic, pump it out) likely occurs through horizontal gene transfer.

Post-Lab Analysis:
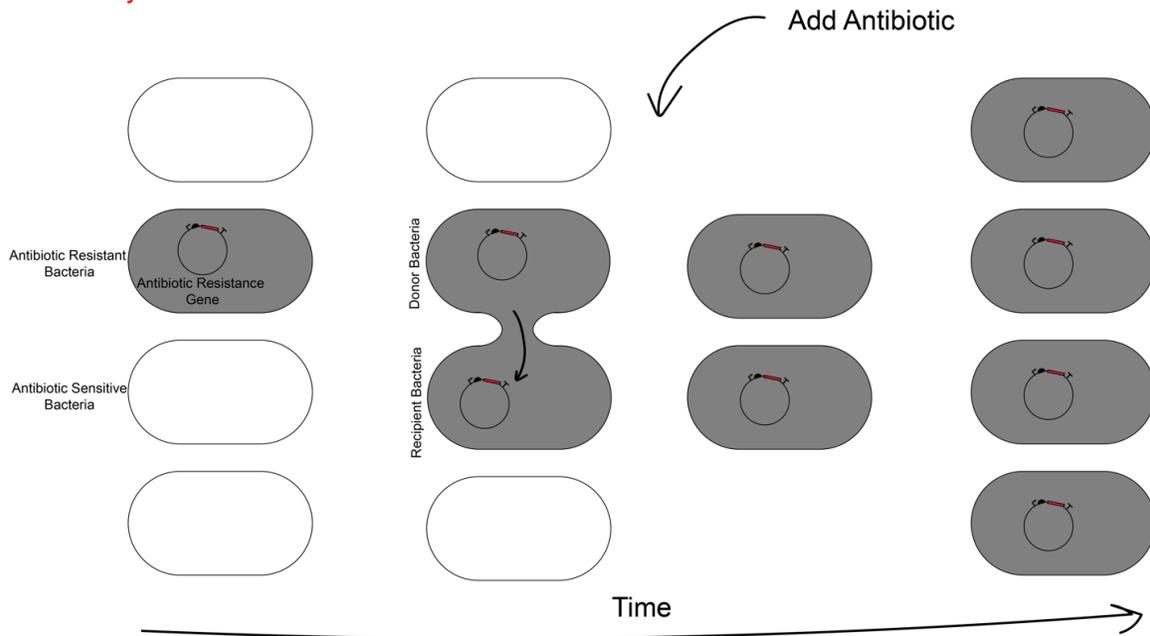
1.
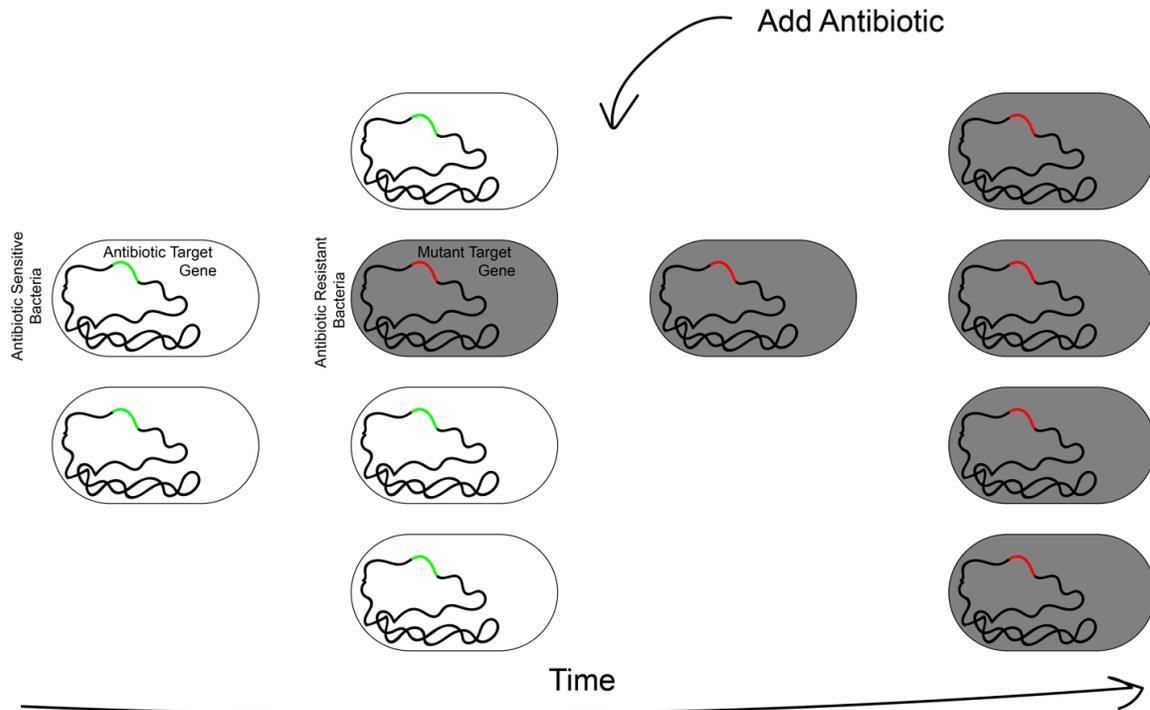
2.

3. In one experiment, pre-expressed AadA was added to the cell-free reaction in the presence of antibiotics. This scenario simulated horizontal gene transfer as a new protein encoded on a plasmid was used to cause antibiotic resistance. In the other experiment, a rpsL R86S mutant cell-free extract was used instead of the wild-type extract in the presence of antibiotic. This scenario simulated random mutation since the target of the antibiotic was mutated and the change came from change in the original bacterium's DNA, not from transfer of new DNA. The wild-type FD-CF reactions without AadA added were the experimental controls in this experiment.

4. Answers here can vary greatly. Students might offer alternatives to antibiotics (phage therapy, antimicrobial peptides) or more effective antibiotic use practices (complete doses of antibiotics as prescribed, proper disposal of antibiotics, use of multiple antibiotics at a time to treat patients).

Exploring Antibiotic Resistance in Cell-Free Systems – Teacher Overview

Purpose:
- Students will explore how bacteria develop resistance to antibiotics using cell-free protein synthesis of the fluorescent protein dTomato.
- Students will process the data collected using ImageJ and Excel.
- Students will demonstrate understanding of how antibiotic resistance spreads and will research behaviors and technologies that can help combat this problem.

Background:
- Development of antibiotic resistance and global health problem
- How bacteria become resistant to antibiotics

Pre-lab activity:
- Review antibiotic mechanisms of action
- Research an example of antibiotic resistance and classify it as horizontal gene transfer or genetic mutation

Antibiotic resistance lab:
- Cell-free protein synthesis using dTomato
- Explore horizontal gene transfer and genetic mutation using a resistance enzyme (AadA) and a streptomycin-resistant mutant lysate, respectively
- Vary concentration of antibiotics
- CFPS lab and measure relative color change (ImageJ)
- Optional use of biological statistics (t-test) to analyze fluorescence data

Post-lab analysis:
- Students discover that both horizontal gene transfer and genetic mutation allow protein synthesis to continue in the presence of streptomycin
- Model horizontal gene transfer
- Model genetic mutation
- Research ways that students' and their families' consumer choices can help prevent antibiotic resistance

Curriculum 10

CRISPR Sort

Task: Match key CRISPR components to their biological functions. Work in groups and compare and discuss your group's results with others!

Explanation of card sort activity:
http://www.theteachertoolkit.com/index.php/tool/card-sort

Categories and pictures:

Cas9

PAM

Guide RNA (gRNA)
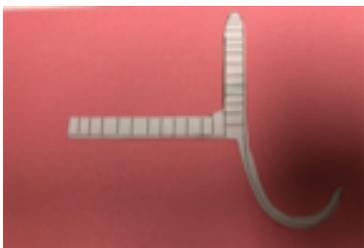
Target DNA

Subheadings/characteristics:

Molecular scissors
Cleaves (cuts) double strands (double strand break)
Endonuclease
Cleaves DNA strands upstream of PAM
Binds PAM/unwinds DNA double helix at site complementing the guide RNA nucleotide sequence
Three-nucleotide sequence motif recognized by Cas9
Cas9 binds
Abundant throughout genome
In humans, occur about every 50 base pairs or less
5'-NGG-3' (AGG, GGG, CGG, or TGG)
Act as GPS
Designed by scientist to compliment target DNA
Usually about 20 nucleotides long
Guides Cas9 to the target
Form complementary base pairs with target DNA
Binds with target DNA triggering Cas9 nuclease activity
Sequence to be cleaved (cut) and removed
Can be any sequence near PAM
Both strands are cleaved (cut)
Repaired by nonhomologous end joining or homology-directed repair
May be a regulatory or coding site
Can be any sequence near PAM
Binding with guide RNA triggers Cas9 nuclease activity
Protospacer Adjacent Motif

CRISPR Sort Answer Key

Cas9	PAM	Guide RNA (gRNA)	Target DNA
Molecular scissors	Three-nucleotide sequence motif recognized by Cas9	Act as GPS	Sequence to be cleaved (cut) and removed
Cleaves (cuts) double strands (double strand break)	Cas9 binds	Designed by scientist to compliment target DNA	Can be any sequence near PAM
Endonuclease	Abundant throughout genome	Usually about 20 nucleotides long	Both strands are cleaved (cut)
Cleaves DNA strands upstream of PAM	In humans, occur about every 50 base pairs or less	Guides Cas9 to the target	Repaired by nonhomologous end joining or homology-directed repair
Binds PAM/ unwinds DNA double helix at site complementing gRNA nucleotide sequence	5'-NGG 3' (AGG, GGG, CGG, or TGG)	Form complementary base pairs with target DNA	May be a regulatory or coding site
		Binding with target DNA triggers Cas9 nuclease activity	Binding with guide RNA triggers Cas9 nuclease activity
Pictures:			
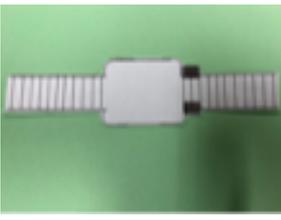
Pictures of CRISPR components adapted from:
https://www.hhmi.org/biointeractive/crispr-cas-9-mechanism-applications

Examples of how the card sort activity might look:
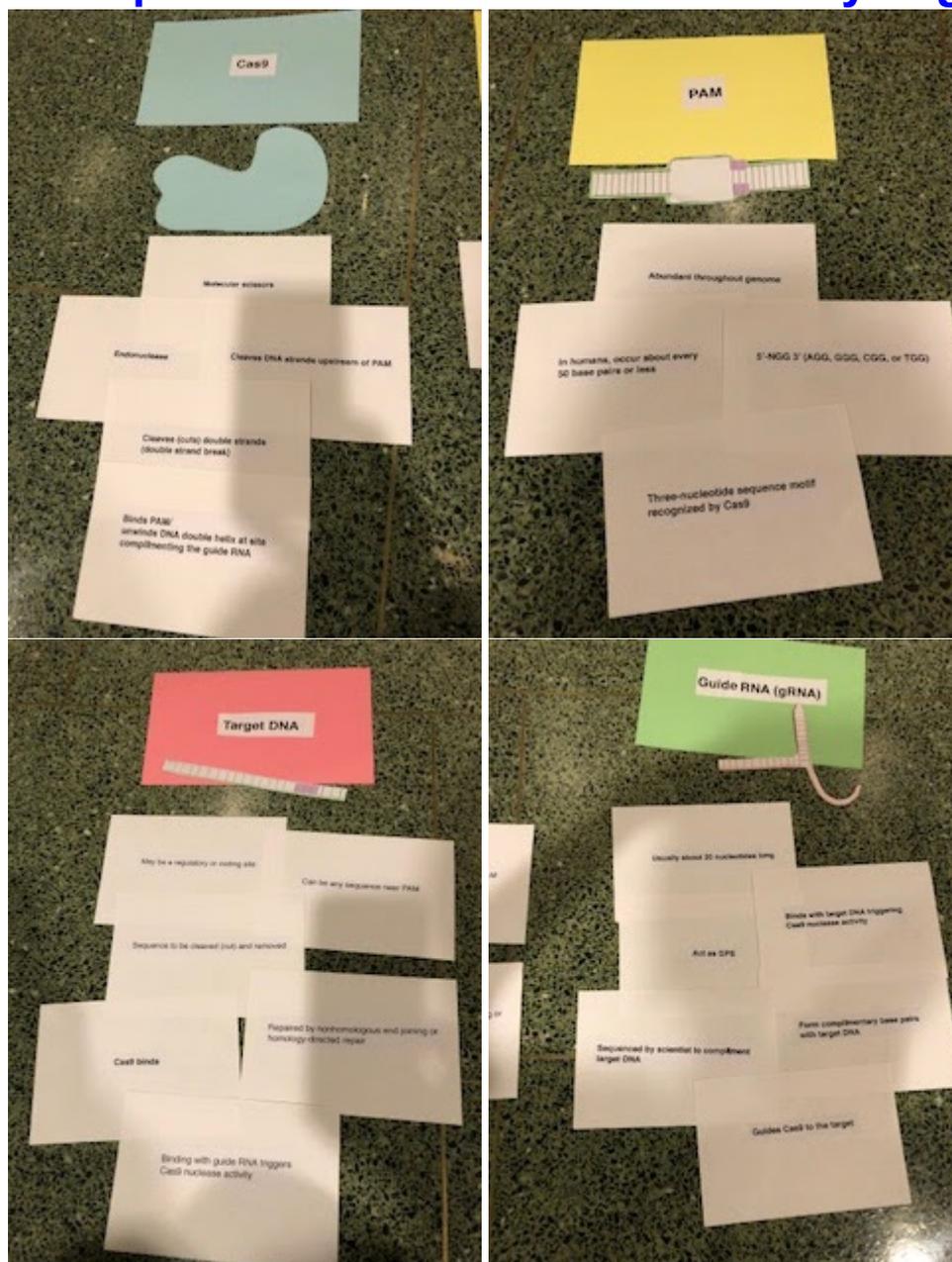
Teacher Notes:
Designed to support a lab activity developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University). Written by Mechelle Washington (Mather High School).

biobits

Exploring CRISPR-Cas9 genome editing with cell-free reactions

BioBits™ Health

Outline

• **What is CRISPR?**

• Investigate how the CRISPR-Cas9 system works

• Consider the ethics behind the use of CRISPR technologies

biobits   BioBits™ Health   2

CRISPR-Cas is the immune system of bacteria and archaea
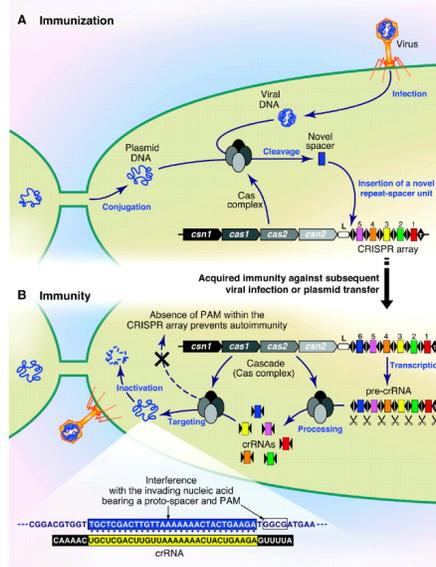
Allows cells to "remember" viruses that have infected them in the past and destroy them before they can infect the cell again

Horvath and Barrangou
Science (2010)

BioBits™ Health

3

CRISPR-Cas systems can be used for targeted genome editing

BioBits™ Health

4
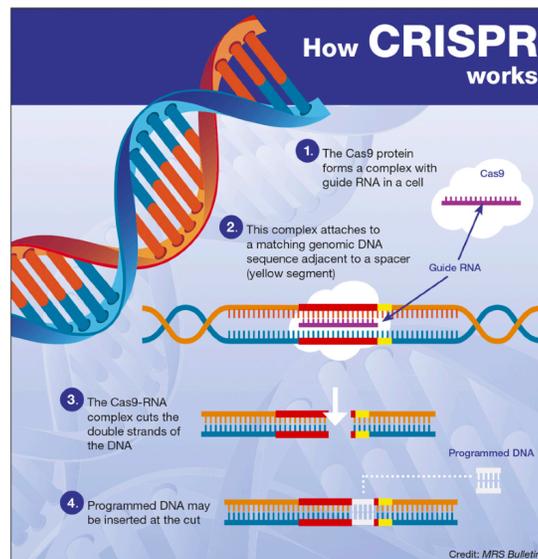
CRISPR technology has many applications for human health

Correct genetic disorders · Create better crop plants · Detect disease at the point-of-care
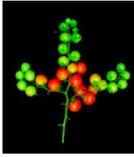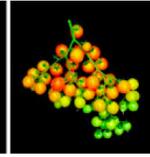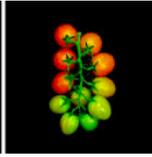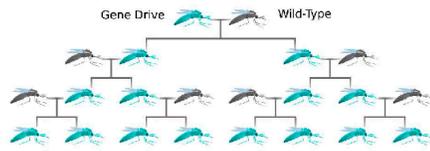
Eliminate mosquito-borne diseases using gene drives · Engineer animal organs to meet transplant needs · Program cells to produce chemicals, materials

https://www.biospace.com/article/-cp8r-researchers-use-crispr-to-treat-duchenne-muscular-dystrophy-in-dogs/
https://www.nationalgeographic.com/environment/future-of-food/food-technology-gene-editing/
https://www.broadinstitute.org/news/sherlock-team-advances-its-crispr-based-diagnostic-tool
https://wyss.harvard.edu/media-post/crispr-cas9-gene-drives/
http://time.com/5159889/why-pig-organs-could-be-the-future-of-transplants/

biobits — BioBits™ Health — 5

Outline

- What is CRISPR?

- **Investigate how the CRISPR-Cas9 system works**

- Consider the ethics behind the use of CRISPR technologies

biobits — BioBits™ Health — 6

3 components are needed for CRISPR-Cas9 activity

You will investigate how these components work together to cut DNA

You will use a guide RNA that targets Cas9 to a fluorescent protein

Fluorescent protein
DNA

Fluorescent protein
mRNA

Fluorescent protein

?
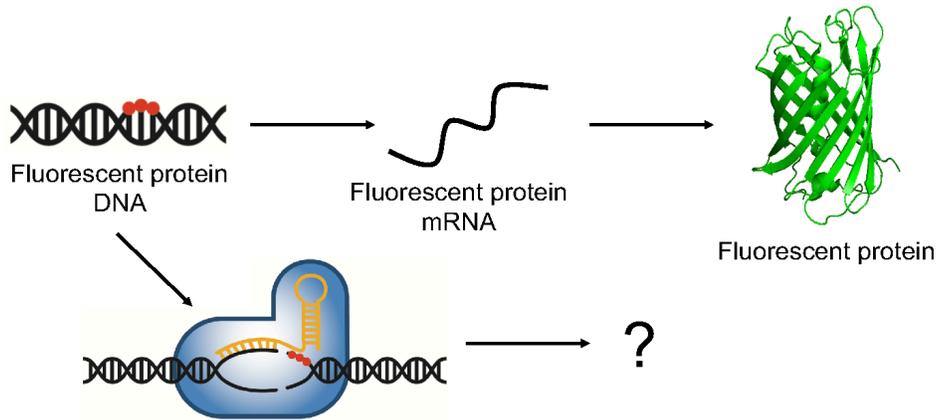
What do you think will happen when Cas9 is used to target the fluorescent protein gene?

Outline

- What is CRISPR?

- Investigate how the CRISPR-Cas9 system works

- **Consider the ethics behind the use of CRISPR technologies**

If you can solve a problem, should you?

Ledford, *Nature* (2015)

CRISPR technology has the potential to transform our ability to treat human diseases like cancer, HIV, and Alzheimer's disease, improve agricultural crops, eliminate malaria from mosquitoes, increase the supply of organs for transplants, and make materials like plastics sustainably.

But scientists are concerned that the ability to edit life could be abused to create, for example, designer babies or simply that CRISPR will be implemented in humans before it is fully understood and the risks are known.

You will research the benefits and potential risks associated with CRISPR technology and decide how you think CRISPR technology should be used.

biobits   BioBits™ Health   11

Curriculum 12

Using fluorescent proteins as reporters of CRISPR-Cas9 activity

Objective: Use freeze-dried, cell-free (FD-CF) reactions and a CRISPR-Cas9 system to determine which fluorescent protein DNA sequence (target DNA) is targeted by a mystery guide RNA.

Pre-lab Activity:

Review the components of the CRISPR-Cas9 system. What role does each component play in the activity of CRISPR-Cas9? Describe the role of each component in the table below:

CRISPR-Cas9 System	
Component	What role does this component play in CRISPR-Cas9 activity?
Cas9	
Guide RNA	
PAM	
Target DNA	

1) How will you know if Cas9 was active?

2) If the Cas9 guide RNA works (binds) your target DNA, what do you expect to observe?

3) If the Cas9 guide RNA does not work, what do you expect to observe?

Procedure:

Using fluorescent proteins as reporters of CRISPR-Cas9 activity lab

Materials needed:
- FD-CF reactions in PCR tube strips (4 reactions per group)
- Large-scale (15 µL) FD-CF reactions in microcentrifuge tubes (5 µL of reaction mixture needed per group)
- 13.33 ng/µL Cas9 plasmid DNA stock solution (provided)
- 13.33 ng/µL fluorescent protein plasmid DNA stock solution (one each of mCherry, mRFP1, dTomato, mOrange,YPet, and sfGFP plasmid DNA provided)
- 33.33 ng/µL gRNA plasmid DNA stock solution (provided)
- Nuclease-free water (provided)
- Sterile pipet tips and micropipettes
- PCR thermocycler OR BioBits^TM portable incubator
- Blue light imager OR BioBits^TM portable imager

DAY 0
22. Rehydrate large-scale FD-CF reactions by adding 15 µL of Cas9 plasmid DNA stock solution to the FD-CF reactions in microcentrifuge tubes. Close the lid and gently flick the side of the tube to dissolve the freeze-dried CFPS pellet. Reactions can be incubated in the BioBits^TM portable incubator or a PCR thermocycler for 20-24 hours at 30°C or on the benchtop at room temperature for 24-48 hours. These reactions will pre-express the Cas9 protein that you will use in your experiment. <u>You will need to prepare 5 µL of Cas9 reaction mix per group.</u>
    This step can be done by the instructor or students.

DAY 1
19. Obtain one of the fluorescent protein plasmid stocks from your instructor. Depending on class size, each group can test the activity of the mystery gRNA on one or multiple of the fluorescent proteins. The class' combined data should test the ability of the gRNA to target each of the 6 fluorescent proteins.
20. Label a strip of four FD-CF reactions with the name of your fluorescent protein plasmid. Label two of these tubes with "-" and two with "+" to indicate which tubes will include gRNA plasmid.
21. Add the appropriate amounts of Cas9 reaction mix, gRNA plasmid stock, and water (refer to the table below) to each of the FD-CF reactions.
    a. Add water first, then fluorescent protein plasmid, then Cas9, then gRNA plasmid.
    b. You can use the same pipet tip for each water addition as long as you do not touch any surface other than the inside of the tubes! For pipetting Cas9, fluorescent protein plasmid, and gRNA plasmid you'll need to switch pipet tips.
    c. *Be careful to avoid touching the white pellet at the bottom of the tubes!*

Tube	Water (µL)	Fluorescent protein plasmid (µL)	Cas9 FD-CF reaction (µL)	gRNA plasmid (µL)
- gRNA	4	1	1	0
+ gRNA	3	1	1	1

22. Seal your FD-CF reactions using an 8-strip PCR tube lid. Gently flick the side of the PCR tubes to dissolve the pellet in the Cas9/DNA/water mixture that you added.
23. Allow the reactions to incubate until the next time you come to class. Reactions can be incubated in a PCR thermocycler for 20-24 hours at 30°C or on the benchtop at room temperature for 24-48 hours. Get excited!

DAY 2
24. Obtain your group's samples from your teacher. Place your strip of reactions in front of a piece of white paper and record your observations.
25. Take a picture of your reactions under blue light (using the BioBits™ blue light imager or other blue light source) and record your observations.
26. Quantify fluorescence produced in each reaction with ImageJ, using the BioBits™ Health ImageJ tutorial. Record the relative fluorescence unit (RFU) values you measure in ImageJ in the table below. Consult with your classmates who tested other fluorescent protein targets to complete the data table for all 6 fluorescent proteins.
27. Calculate the average and standard error of the RFU values you measured using Excel. Record these values in the table below.
28. Plot the average RFU values you calculated as a function of gRNA addition for each fluorescent protein target. Choose an appropriate graph type (bar, scatter, etc.) to display your results. Add error bars showing the standard error of your measurements.
29. *Optional: Use the "t-test: Paired Two Sample for Means" function in Excel to determine if the addition of gRNA results in statistically significant differences in fluorescence for each fluorescent protein target. The following tutorial can help you get started:* https://www.rwu.edu/sites/default/files/downloads/fcas/mns/running_a_t-test_in_excel.pdf

Data:

Fluorescent protein: mCherry				
Guide RNA?	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
-				
+				

Fluorescent protein: mRFP1				
Guide RNA?	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
-				
+				

Fluorescent protein: dTomato				
Guide RNA?	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
-				
+				

Fluorescent protein: mOrange				
Guide RNA?	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
-				
+				

Fluorescent protein: YPet				
Guide RNA?	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
-				
+				

Fluorescent protein: sfGFP				
Guide RNA?	RFU reaction 1	RFU reaction 2	Average RFU	Std Error RFU
-				
+				

Post-lab Analysis:

Which fluorescent protein gene did the gRNA target? How do you know?

Answer Key:

CRISPR-Cas9 System	
Components/Acting Molecules:	**What do you expect this molecule to do? What will happen?**
Cas9	Enzyme that can cut DNA. If guided by gRNA to a specific target cute site, it will cut the DNA at that site.
Guide RNA	A synthetic piece of RNA that consists of a scaffold that can bind Cas9 and a spacer that can be designed to target any DNA sequence. Using this setup, the gRNA guides the Cas9 to the target cut site so the Cas9 enzyme can cut that site.
PAM	Protospacer Adjacent Motif – the target cut site needs to be immediately adjacent to the relevant PAM sequence. The PAM sequence allows for Cas9 to bind and cleave the target DNA sequence.
Target DNA	The DNA molecule that contains the target cut site. The gRNA will bind to this DNA molecule at a specific site for Cas9 to cut there.
Cut sites	The cut site is the target where we want the Cas9 to cut. The spacer on the gRNA will be complementary to this cut site in order to guide the Cas9 to the target cut site on the target DNA.

1) How will you know if Cas9 was effective? The CRISPR-Cas9 system will target one of the fluorescent-protein-encoding DNA sequences. If Cas9 is effective in cutting at the target cut site, then the DNA sequence will be split in two and therefore will not be able to undergo transcription and translation to make the fluorescent protein, leading to no observed fluorescence.

2) If the Cas9 guide RNA works (binds) your target DNA, what do you expect to see?
Once the Cas9 guide RNA binds to the target DNA, the Cas9 will cut the target DNA at that specific cut site. Because the DNA is now cut, transcription cannot happen to make the RNA, and then the reporter protein cannot translated, and thus no fluorescence will be observed.

3) If the Cas9 guide RNA does not work, what do you expect will happen? If the Cas9 guide RNA does not bind, then the Cas9 will not cut the DNA. If the DNA remains uncut, it will be able to undergo transcription and translation to create its encoded fluorescent protein, which we can observe.

Note: This is what is expected assuming that everything went correctly. If student results do not match this table, encourage a discussion where different groups can compare their results and come up with hypotheses why their results were different.

Target Protein	Guide RNA?	
	-	+
mCherry	Red fluorescence	Red fluorescence (same as –gRNA)
mRFP1	Red fluorescence	Decreased or no red fluorescence (compared to –gRNA)
dTomato	Orange fluorescence	Orange fluorescence (same as –gRNA)
mOrange	Orange-yellow fluorescence	Orange-yellow fluorescence (same as –gRNA)
YPet	Yellow fluorescence	Yellow fluorescence (same as –gRNA)
sfGFP	Green fluorescence	Green fluorescence (same as –gRNA)

Which fluorescent protein gene did the gRNA target? How do you know? The gRNA is designed to target the DNA that encodes for mRFP1. This is because for all other proteins, we did not see a difference in protein expression between the reactions where we didn't add gRNA and the reactions where we did add gRNA. This means the CRISPR-Cas9 system did not bind and cleave the DNA because the DNA was still able to be used in transcription and translation to create the fluorescent protein. On the other hand, the mRFP1 reactions show a decrease in fluorescence if the gRNA was added. This means that the CRISPR-Cas9 system did bind and cleave the DNA, causing it unable to be used in transcription and translation to create the mRFP1 fluorescent protein.

Teacher Notes:
Designed to support a lab activity developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University). Written by Mechelle Washington (Mather High School) and Jessica Stark.

Curriculum 13

Exploration: CRISPR HTT Activity

Part I:

Research the HTT gene and Huntington's disease by reading the following articles:

1) HTT Gene
https://ghr.nlm.nih.gov/gene/HTT

2) Huntington disease
https://ghr.nlm.nih.gov/condition/huntington-disease

3) New NIST SRM Helps Improve Diagnosis of Huntington's Disease
https://www.nist.gov/news-events/news/2011/04/new-nist-srm-helps-improve-diagnosis-huntingtons-disease

4) Modern Genome Editing Technologies in Huntington's Disease Research
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5389024/

5) CRISPR takes on Huntington's disease: Gene editing offers the prospect of curing the inherited neurodegenerative condition in a single dose
https://www.nature.com/articles/d41586-018-05177-y

Part II:

Compare the following target DNA sequences for Huntington's Disease. The inherited mutation that causes Huntington disease is known as a CAG trinucleotide repeat expansion. This mutation increases the size of the CAG segment in the *HTT* gene. People with Huntington disease have 36 to more than 120 CAG repeats. People with 36 to 39 CAG repeats may or may not develop the signs and symptoms of Huntington's disease, while people with 40 or more repeats almost always develop the disorder (1).

Based on your research and these criteria, which if any of the following patients have the mutated form of the HTT gene? Annotate on the sequence. Reference specific points in each sequence to support your conclusion.

Patient 1:
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTCCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGAACAGCCG
CCACCGCCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGCCGCAGGCACA
GCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCCACCCGGCCCG
GCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTACC...+9,141
bases

Patient 2:
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTCCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACA
GCCGCCACCGCCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGCCGCAG
GCACAGCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCCACCCG

GCCCGGCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTACC...
+9,141 bases

Patient 3:
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTCCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC
AGAACAGCCGCCACCGCCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGC
CGCAGGCACAGCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCC
ACCCGGCCCGGCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTA
CC...+9,141 bases

Patient 4:
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTCCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA
GAACAGCCGCCACCGCCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGCC
GCAGGCACAGCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCCA
CCCGGCCCGGCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTAC
C...+9,141 bases

Part III: Determine which tools work best!

Examine the following diagram. The diagram shows several editing methods that can be utilized to edit the HTT gene. Which of the genome editing tools would you use to edit a patient's mutated HTT gene? Choose one of the patient's HTT gene sequences from Part II and explain how you would use one of the editing tools that you read about in article 4 above (pictured below) to edit out that patient's CAG trinucleotide repeat expansion.

Depending on which complex you choose, use the terms PAM, RNA guide, target DNA, DNA guide and/or the name of the editing complex you choose, when explaining why that system would work best for your sequence. Reference article 4 link, Figure 2.

HTT editing

Image credit: Malankhanova, et al. *J Huntingtons Dis.* (2017)

Patient # _____ DNA Sequence:

Explanation: (Why did you choose this editing tool? Why is it the best tool? How will it work to edit your target sequence?)

Part IV: Model CRISPR in Six Steps or Less

Illustrate and describe how CRISPR Cas9 would remove the CAG trinucleotide repeat expansion from one of the patients from genome (Part II), in six steps or less.
Be sure to design a guide RNA for the patient's HTT targeted DNA sequences and reference the location of the PAM site.

CRISPR HTT Extension: Using BLAST

BLAST (Basic Local Alignment Search Tool) is a powerful online tool for comparing biological sequences and comparing user input biological sequences to sequences in the National Center for Biotechnology Information's (NCBI) database. This activity will provide a brief walk through to introduce you to BLASTn, a tool for comparing DNA sequences. Here, we will use BLASTn to compare the patient DNA sequences from the above activity to each other and to known human genes in the NCBI database.

1) Navigate to https://blast.ncbi.nlm.nih.gov/Blast.cgi on your web browser. You will see the home page for BLAST. Select Nucleotide BLAST.

2) Copy and paste the following DNA sequences into the enter query sequence box (red dash):

*Note: you may copy and paste multiple sequences into the same box if they are in FASTA format (https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=BlastHelp)

>Patient 1
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTCCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGAACAGCCG
CCACCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGCCGCAGGCACA
GCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCCACCCGGCCCG
GCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTACC

>Patient 2
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTCCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACA
GCCGCCACCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGCCGCAG
GCACAGCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCCACCCG
GCCCGGCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTACC

>Patient 3
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTCCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC
AGAACAGCCGCCACCGCCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGC
CGCAGGCACAGCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCC
ACCCGGCCCGGCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTA
CC

>Patient 4
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTCCAGCAGCAG
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC
AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA
GAACAGCCGCCACCGCCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGCC
GCAGGCACAGCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCCA
CCCGGCCCGGCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTAC
C

3) Once the sequences have been pasted into the enter query sequence box, select "Human genomic + transcript" under the Database heading (red dash). This will limit our search from the entire NCBI database to just sequences found in humans. Press BLAST (green dash) when you have selected the database and are ready to run the algorithm.

4) After the algorithm is done running, you will see the output screen. A description of all alignments to the database for a single query sequence is displayed (red dash). You can toggle between query sequences by selecting them from the drop-down menu (green dash).

5) Within the description, alignments between the NCBI database and the query are listed. The identity of the database hit (red dash) and alignment statistics (green dash) are present for each alignment. The alignment statistics consist of Max score, Total score, Query coverage, E value, and Identity. Max score and Total score are measures of alignment strength between query and database hit based on the BLAST algorithm's criteria (match, mismatch, gap & extension). The same database hit can align to the query multiple times. Max score is a measure of the highest single alignment for a database hit. Total score is a measure of the sum of all alignments for a single database hit. Query coverage is how much of the query sequence has aligned to database hit. 100% Query coverage means that 100% of the query was

involved in the alignment to the database hit. The database hit can be longer than the query. E value is a measure of the likelihood that an alignment is due to random chance. A lower E value represents a low probability that the alignment is due to random chance. Identity is a measure of how similar the nucleotides in query are to the database hit. 100% identity means that the two sequences are identical within the alignment. 50% identity means that 50% of the nucleotides in the query are the same as the nucleotides in the database hit within the alignment.

6) Individual alignments can be seen if you scroll down on the output screen. Within each alignment, the range of the alignment (red dash), alignment statistics (green dash), and alignment visual (blue dash) are displayed. The alignment statistics are similar to those described in step 5, but Gaps and Strand are added. Gaps refers to the number of times one sequence has a nucleotide in a position that the other sequence does not. Gaps decrease the strength of an alignment. Strand refers to which of the two strands of DNA the sequence is found on. DNA forms a double helix and contains a Plus and a Minus strand. The DNA sequence that we entered into BLASTn only represents one of these strands. The Strand metric does not affect alignment strength. The alignment visual (blue dash) allows you to see how the nucleotides align and the location of mismatches and gaps.

7) Compare all the alignment statistics for the 4 patient sequences. Fill in the table below.

Patient	Top Hit	Total Score	Query Cover	E Value	Identity	Gaps
1						
2						
3						
4						

Post-Extension Questions:

1) Which patient alignment was the strongest to the "Homo sapiens huntingtin (HTT), mRNA" database hit? Reference alignment statistics in your answer.

2) Is the HTT gene in the NCBI database the mutant or wild-type gene? Explain how you know.

3) How would you use BLASTn to find the identity of an unknown DNA sequence?

Teacher Resources
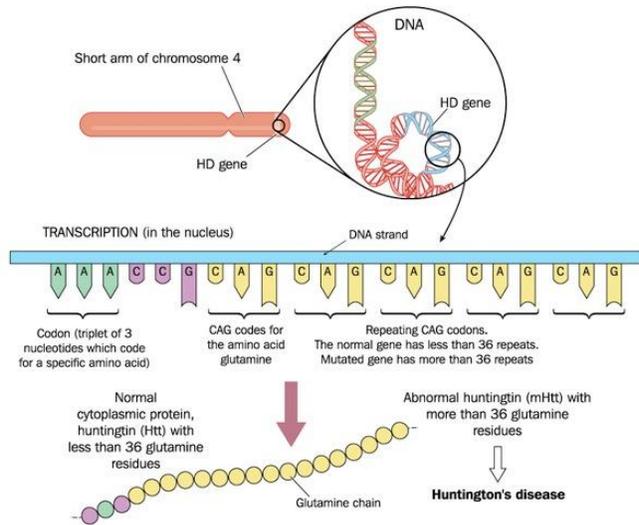
Gene editing tools	Efficiency	Specificity	Design simplicity	Delivery	Sequence limitations	Cost
ZFN	low	low	complicated	medium difficulty	preference for G-rich consensus sequences	high
TALEN	medium	high	complicated	medium difficulty	T at the 5' end of the target	cheap
CRISPR/Cas9	high	high	easy	easy	5'-NGG-3' PAM at the 3' end of the target	cheap
High-fidelity CRISPR/Cas9	high	high	easy	easy	5'-NGG-3' PAM at the 3' end of the target	cheap
CRISPR-Cpf1	high	high	easy	easy	5'-TTN-3' PAM at the 5' end of the target	cheap
NgAgo	high	high	easy	easy	none	cheap
SGN	low	low	easy	medium difficulty	none	cheap

*The paper that describes the use of NgAgo for genome editing was retracted shortly after "Modern Genome Editing Technologies in Huntington's Disease Research" (source 4) was published. NgAgo was removed from the choices in Part III to reflect this.

Repeat expansion mutation

Original DNA code for an amino acid sequence.

DNA → bases

C A T T C A C A G G T A A T C A T G C T A

His — Ser — Gln — Val — Ile — Met — Leu

↑
Amino acid

Repeated trinucleotide (CAG).

C A T T C A C A G C A G C A G G T A A T C

His — Ser — Gln — Gln — Gln — Val — Ile

Repeated trinucleotide adds a string
of glutamines (Gln) to the protein.

target DNA

5'
3'

PAM

3'
5'

guide RNA

Cas9

5'

3'

KEY:

CRISPR HTT Activity

Part II:
Patients 4 will likely develop Huntington's disease as they each have more than 40 repeats. Patient 1 may develop symptoms of Huntington's disease as they have 39 repeats. Patients 2 & 3 have a normal number of repeats and are not at increased risk of Huntington's disease.

Patient 1: (39 repeats)
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTC<u>CAGCAGCAG CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG</u>AACAGCCG CCACCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGCCGCAGGCACA GCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCCACCCGGCCCG GCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTACC...+9,141 bases

Patient 2: (21 repeats)
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTC<u>CAGCAGCAG CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC</u>AACA GCCGCCACCGCCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGCCGCAG GCACAGCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCCACCCG GCCCGGCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTACC...+9,1 41 bases

Patient 3: (23 repeats)
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTC<u>CAGCAGCAG CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC AG</u>AACAGCCGCCACCGCCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGC CGCAGGCACAGCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCC ACCCGGCCCGGCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTA CC...+9,141 bases

Patient 4: (42 repeats)
ATGGCGACCCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCTCAAGTCCTTC<u>CAGCAGCAG CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA G</u>AACAGCCGCCACCGCCGCCGCCGCCGCCGCCGCCTCCTCAGCTTCCTCAGCCGCCGCC GCAGGCACAGCCGCTGCTGCCTCAGCCGCAGCCGCCCCCGCCGCCGCCCCCGCCGCCA CCCGGCCCGGCTGTGGCTGAGGAGCCGCTGCACCGACCAAAGAAAGAACTTTCAGCTAC C...+9,141 bases

Part III:
Answers can vary here. Most common answer will be to use CRISPR. Depending on which CRISPR technique (Cas9 or Cpf1) the student chooses, the site of the PAM may vary. Cas9 needs an NGG PAM at the 3' end of the target, while Cpf1 requires a TTN at the 5' end of the target. Cas9 is likely the easier of the two.
Part IV:

1) Vector containing a repair template, Cas9 DNA, and gRNA DNA is injected into the patient.

2) Cas9 and the gRNA are expressed, assemble together, and cut the mutant gene

3) The repair template homologously combines with the cut patient's DNA

4) The patient's gene is now the non-muta sequence

CRISPR HTT Extension: Using BLAST

Patient	Top Hit	Total Score	Query Cover	E Value	Identity	Gaps
1	Homo sapiens huntingtin (HTT), mRNA	655	100%	4e-120	99.58%	2
2	Homo sapiens huntingtin (HTT), mRNA	538	100%	9e-161	100.00%	0
3	Homo sapiens huntingtin (HTT), mRNA	514	100%	2e-143	98.31%	5
4	Homo sapiens huntingtin (HTT), mRNA	655	100%	4e-120	99.58%	2

1) The alignment between patient 2 and Homo sapiens huntingtin (HTT), mRNA was the strongest alignment. That alignment had the lowest E-value, highest identity and fewest gaps.

2) The HTT gene in the NCBI database is the wild-type gene. The gene aligns 100% with patient 2, who does not have an increased risk of developing Huntington's Disease.

3) To identify an unknown sequence, you could enter the DNA sequence into BLASTn and see what it aligns to. To evaluate which alignment is the most likely, you can use the alignment statistics and look for high query cover and identity, as well as low E-value and fewer gaps.

Teacher Notes:

Designed to support a lab activity developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University). Written by Mechelle Washington (Mather High School) and Grant Rybnicky (NSF Graduate Research Fellow, Ph.D. Student, Interdisciplinary Biological Sciences, Northwestern University).

Curriculum 14

Exploring Solutions to Genetic Problems

<table>
<tr>
<td>1</td>
<td>Chef Complaint
James, 6 years old, was brought in to the clinic experiencing fatigue and joint pain.</td>
</tr>
<tr>
<td>5</td>
<td>Patient Background
James is a quiet 6-year-old boy who leads a relatively sedentary lifestyle. When he is home, he spends most of his time in his room reading books or watching cartoons. Three months ago, he began attending first grade and seems to enjoy it. Recently, his mother (Sheryl) noticed that James has been appearing sluggish
and she has had to work even harder to get him up for school in the morning. While she suspected something was not quite right, she reasoned that he's a growing boy and his body must need the rest.</td>
</tr>
<tr>
<td>15</td>
<td>On Monday, Sheryl received a phone call informing her that James wasn't feeling well and that he needed to be picked up from school. When his mother asked him what was wrong, James indicated that he felt really tired and his arms and legs hurt. While looking him over, Sheryl noticed that his extremities seemed a little swollen. She immediately scheduled a doctor's appointment and brought her son in for further examination.</td>
</tr>
<tr>
<td>20</td>
<td>Upon physical examination his physician confirms that his extremities are swollen and finds that his blood oxygen is also low. A urine and blood sample were taken and sent to the laboratory for analysis.</td>
</tr>
<tr>
<td>25</td>
<td>Diagnosis
James' blood sample came back positive for hemoglobin S, a defective version of the hemoglobin A protein found in red blood cells. His physician diagnosed James with Sickle Cell Anemia.</td>
</tr>
<tr>
<td>30</td>
<td>TASK
Use the following models to better understand James' Sickle Cell condition and explore current methods for treating it. Then, construct a model to illustrate how gene editing tools like CRISPR Cas9 can be used to correct genetic conditions like Sickle Cell Anemia.</td>
</tr>
</table>

<u>Model 1</u>: **Origin of Sickle Cell Anemia**

*Image sources: https://ct-stem.northwestern.edu/curriculum/preview/28/2/;
https://www.studyblue.com/notes/note/n/03-biomolecules-watts/deck/1859402*
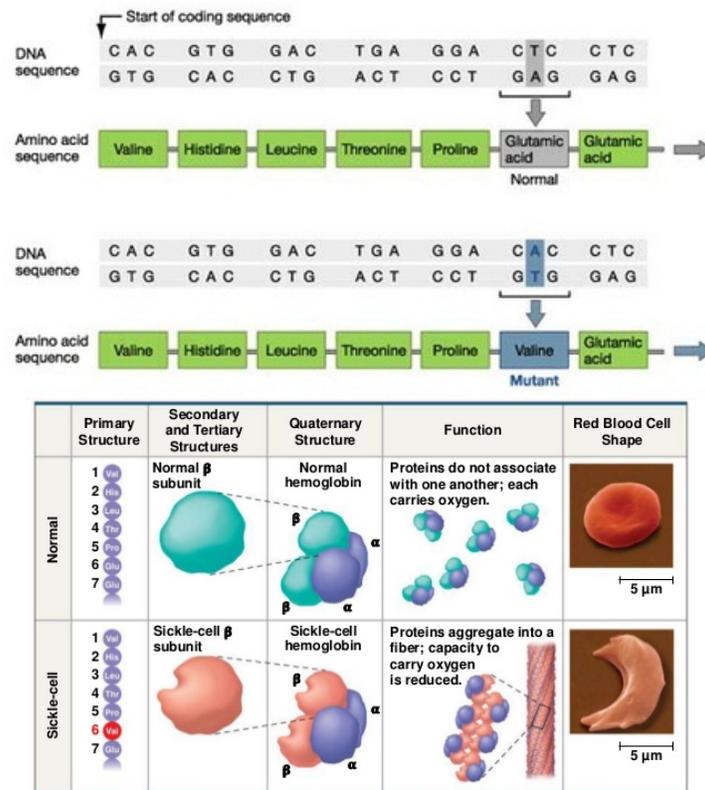
Questions:

1. What is the primary function of red blood cells in your body and why is this function essential to one's survival?

2. How do red blood cells affected by the Sickle Cell condition phenotypically differ from "normal" red blood cells?

3. How might these phenotypic changes account for James' symptoms?

4. What is the source of this phenotypic difference between normal and sickled cells? Elaborate on how this difference can produced a new phenotype using your understanding of the central dogma.

Model 2: **Transfusion**

Test for:
HIV
Hepatitis B
Hepatitis C
HTLV
Syphilis
ABO + RhD
Other phenotype
red cell antibodies
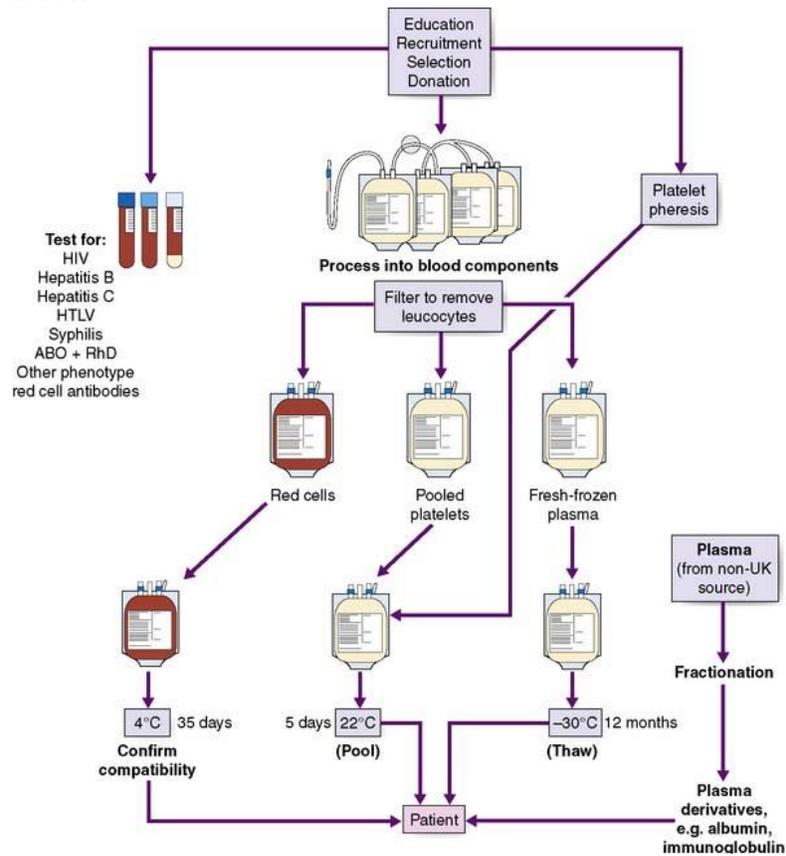
Questions:

5. After blood has been donated, its components can be separated and transfused according to the recipient's needs. Which portion of the donor sample would James require and how will it help treat his condition?

6. Will this treatment address the symptoms or cause of James' condition? (Is this a temporary or long-term solution?)

7. What limitations exist with this treatment?

Model 3: **Transplantation**
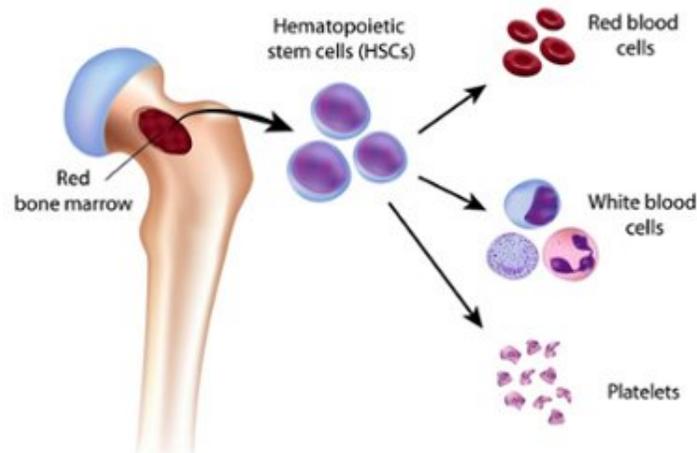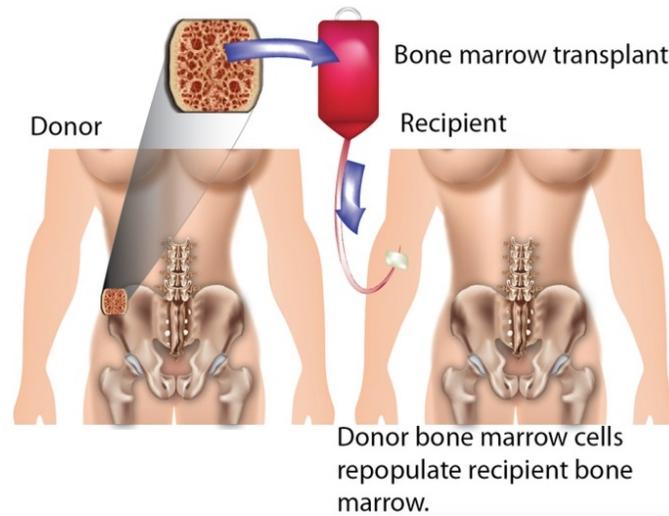
Image sources: Health Library for People (http://www.healthlibrary.com/); lyceum.algonquincollege.com

Questions:

8. How will a bone marrow transplant treat James' condition similarly to a blood transfusion?

9. How will a bone marrow transplant treat James' condition differently from a blood transfusion?

10. Will this treatment address the symptoms or cause of James' condition? (Is this a temporary or long-term solution?)

11. What obstacles might exist in receiving a successful bone marrow transplant?

Model 4: **Gene Therapy**

DIRECT DELIVERY

therapeutic transgene

the therapeutic transgene is packaged into a delivery vehicle such as a virus

... and injected into the patient

target organ

CELL-BASED DELIVERY

therapeutic transgene

the therapeutic transgene is packaged into a delivery vehicle such as a virus

the therapeutic transgene is introduced into a delivery cell such as a stem cell that is often derived from the patient

the genetically modified cells (e.g. stem cells) are multiplied in the laboratory

... and readministered to the patient

12. What therapeutic transgene would you expect to find in a plasmid designed to treat Sickle Cell Anemia?

13. How will this therapeutic gene be introduced into the host/recipient of the treatment?

14. Compare and contrast treatment via gene therapy with that of a bone marrow transplant.

<u>Model 5</u>: **Gene Editing**

15. Create a model demonstrating how the CRISPR Cas9 complex could treat James' Sickle Cell condition. *Hint: You may find Model 1 helpful in completing this task.

16. How would using CRISPR technology provide an advantage in treating individuals with Sickle Cell Anemia over each of the previous treatments modeled?

KEY:

1) Red blood cells carry oxygen inhaled by the lungs to tissues throughout the body. They also carry carbon dioxide produced by tissues to the lung to be exhaled. This is critical to survival as all cells need oxygen for cellular respiration and would be poisoned by carbon dioxide otherwise.
2) Sickle cell affected cells have a different cell shape and a reduced ability to carry oxygen.
3) James experienced fatigue, which may be due to his body's inability to make enough energy (ATP). Low amounts of oxygen in the tissue would prevent respiration and allow for little ATP to be made. Sickle cell affected cells have a lower capacity for carrying oxygen.
4) The Sickle cell phenotype is caused by a mutation in the hemoglobin gene. A point mutation in the DNA encoding hemoglobin causes a different RNA transcript to be made, ultimately resulting in a valine amino acid in position 6 of hemoglobin instead of a glutamic acid amino acid.
5) The red cells component.
6) Transfusing red blood cells will help alleviate James' symptoms, but is a short-term solution since his body will continue to make Sickle cell affected red blood cells.
7) Limitations with this treatment are that it does not fix the source of the problem (the Sickle cell mutation) and the treatment is dependent on the availability of donated blood that passes through quality control.
8) A bone marrow transplant will also give James functional red blood cells.
9) A bone marrow transplant will continually give James functional red blood cells and not require transfusions.
10) A bone marrow transplant is more permanent solution than red blood cell transfusion.
11) A bone marrow transplant also requires a donor, but the donor must be compatible with the recipient. It also requires invasive surgery rather than a simple transfusion.
12) A corrected copy of the hemoglobin gene, one that does not contain any mutations.
13) Viral delivery directly to patient OR implantation of genetically modified cells (similar to bone marrow transplant)
14) Gene therapy is independent of tissue compatibility since the tissue used is the patient's own tissue.
15) See Next Page
16) Using CRISPR to edit a patient with Sickle Cell Anemia allows the patient a more permanent solution as their own cells' DNA is edited in the process. This also avoids problems with finding donors and donors being compatible with the patient (bone marrow transplant and red blood cell transfusion).
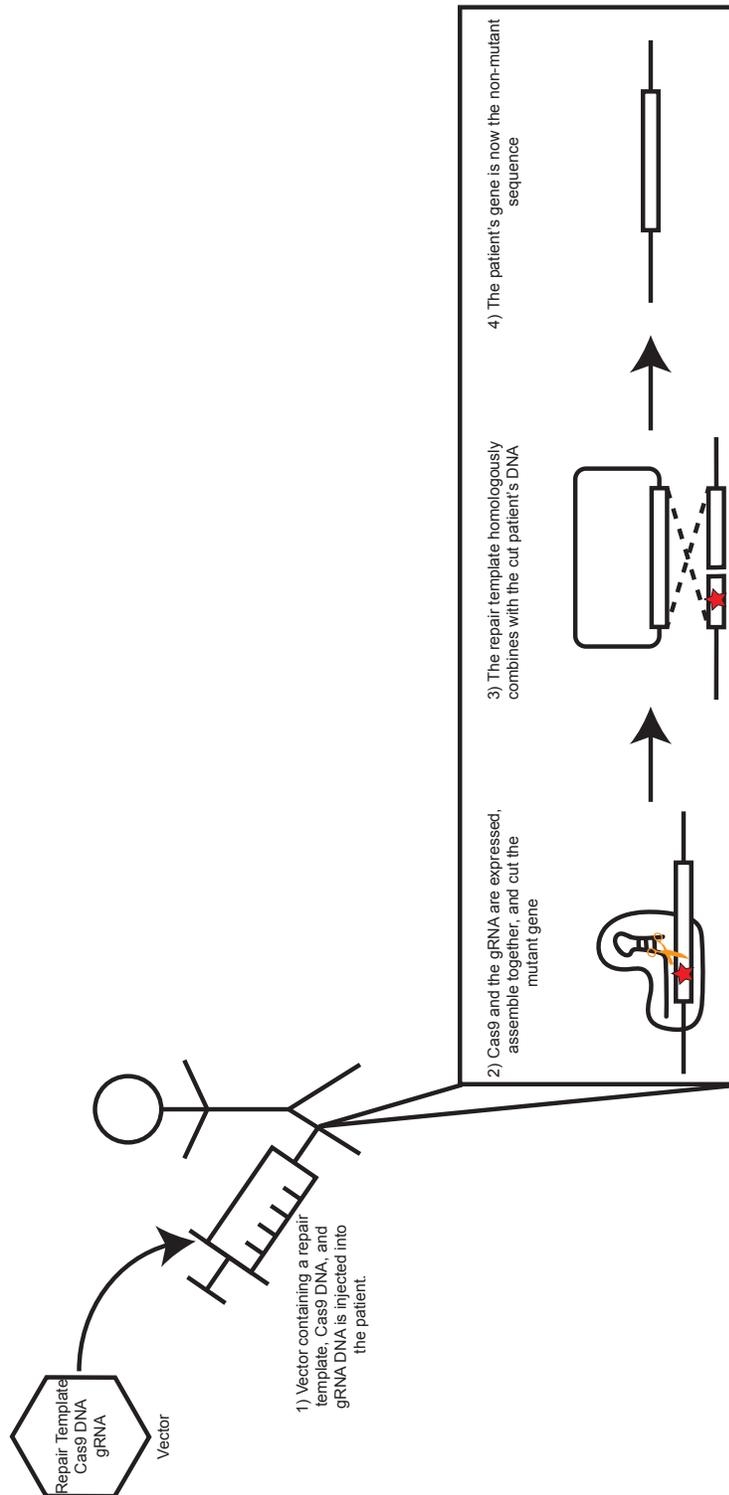
Repair Template
Cas9 DNA
gRNA

Vector

1) Vector containing a repair template, Cas9 DNA, and gRNA DNA is injected into the patient.

2) Cas9 and the gRNA are expressed, assemble together, and cut the mutant gene

3) The repair template homologously combines with the cut patient's DNA

4) The patient's gene is now the non-mutant sequence

Teacher Notes:
Designed to support a lab activity developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University). Written by Jason Forbrook (Waukegan High School).

Curriculum 15

Applications of CRISPR

Driving Question:

How can CRISPR technology be used to tackle current problems in society?

Task:

1. Select a current societal challenge that CRISPR technology has the potential to solve. The following resources can help you select a topic area, but you will likely need to do additional research:
    ○   https://www.nytimes.com/2015/11/15/magazine/the-crispr-quandary.html
    ○   http://time.com/4379503/crispr-scientists-edit-dna/
    List at least one additional source that you used to research your application area:

2. Work with your group to answer the consideration questions (below) that will help you better understand the benefits and concerns of using CRISPR technology to solve this problem.

    ○   Why do we need to solve this problem?

    ○   What strategies are currently being implemented to solve this problem? (address strengths and limitations of these strategies)

    ○   How can CRISPR technology be applied to solve this problem? (address advantages over current strategies as well as limitations)

    ○   Which people/groups would benefit from the application of CRISPR to solve this problem? How?

    ○   Which people/groups might be opposed to the application of CRISPR to solve this problem? Why?

3. Construct a presentation to communicate your findings to your colleagues. You will be evaluated primarily on your communication skills this time, but you may want to use the content of this presentation as part of your final presentation (*see Ethics of CRISPR activity*).

    Applications of CRISPR – Presentation Rubric (see below)

★   As your colleagues present, record information that will help you establish your position on whether or not CRISPR technology should be used to tackle current societal problems (you will present your position in your final presentation). You may use the template below to help you with this task.

    Applications of CRISPR - Structured Notes (see below)

Applications of CRISPR – Presentation Rubric

Presentation Criteria	No Evidence	Approaching	Proficient	Advanced	Evidence
Content Fluency	Presenter does not demonstrate any evidence of understanding of content.	Presenter demonstrates some evidence of understanding of content	Presenter demonstrates understanding of content by speaking fluently about the material for most of the presentation.	Presenter demonstrates understanding of content by speaking fluently throughout and can simplify/elaborate/ make connections/ answer questions related to content when appropriate.	
Visual Aid	Visual aid does not assist at all in communicating information.	Visual aid assists in communicating some information shared by the presenter with the audience	Visual aid assists in communicating the same information shared by the presenter with the audience.	Visual aid clarifies main ideas for the audience, while allowing the presenter to share additional detail.	
Voice	Presenter does not speak with appropriate volume, clarity, or pace for any of the presentation	Presenter speaks with appropriate volume, clarity, and pace for some of the presentation.	Presenter speaks with appropriate volume, clarity, and pace for most of the presentation.	Presenter speaks with appropriate volume, clarity, and pace throughout the presentation.	
Body Language	Presenter's body language does not engage the audience via appropriate eye contact, gestures, and body movement throughout for any of the presentation.	Presenter's body language engages the audience via appropriate eye contact, gestures, and body movement throughout for some of the presentation.	Presenter's body language engages the audience via appropriate eye contact, gestures, and body movement throughout for most of the presentation.	Presenter's body language engages the audience via appropriate eye contact, gestures, and body movement throughout the presentation.	
Reflection					

Applications of CRISPR - Structured Notes

Group	Application: Why is this a problem that needs to be solved?	Current Strategies: What are the strengths and limitations?	Why CRISPR? Describe advantages over current strategies.	People/groups in favor: Who would benefit from this? How?	People/groups opposed: Who might oppose this? Why?
1					
2					
3					
4					

Curriculum 16

Ethics of CRISPR

Essential Question

If you can solve a problem, should you?

You and your colleagues will research the ethics associated with the use of CRISPR technologies to answer our driving question as to whether the scientific community should continue to develop CRISPR technologies to tackle current societal challenges.

Understand ethical considerations associated with the use of CRISPR technologies
Use the following articles and video to begin your research into the ethical considerations surrounding the use of CRISPR technology:

- https://www.genome.gov/27569225/what-are-the-ethical-concerns-about-genome-editing/
- https://www.youtube.com/watch?v=8Ijr1ccYPtI
- https://www.washingtonpost.com/news/speaking-of-science/wp/2017/02/14/ethicists-advise-caution-in-applying-crispr-gene-editing-to-humans/?utm_term=.0ea7704583c5
- https://www.broadinstitute.org/news/licensing-crispr-agriculture-policy-considerations
- http://www.bu.edu/khc/files/2018/10/CRISPR-Ethics-reading.pdf
- https://www.omicsonline.org/open-access/ethical-issues-in-genome-editing-using-crisprcas9-system-2155-9627-1000266.pdf

Choose one example of a societal challenge that CRISPR technology has the potential to solve. Research the ethics surrounding this application and explain at least two ethical concerns associated with using CRISPR to address this need. Cite at least two additional sources to support your explanation.

Some concerns about using CRISPR you may consider:
- Ecological: How might CRISPR edited organisms or biological systems impact the environment?
- Ethics: How should scientists or companies regulate CRISPR?
- Morals: How do individuals' own principles regarding right and wrong influence how CRISPR can or should be used?
- Legal: Are there laws or policies that are or should be in place to regulate CRISPR technology?
- Safety: Is CRISPR safe to use in humans or other biological systems? How can the scientific community ensure public safety?

Case Study: CRISPR babies reported in China
In November 2018, a Chinese scientist named He Jiankui announced the birth of two twin girls whose genomes had been edited using CRISPR, representing the first instance of CRISPR editing in humans. He edited the babies' genomes to delete a portion of the CCR5 gene, based on research that suggests that this deletion could make the girls immune to infection by HIV. Their edited genomes would protect the twins from being infected by the virus by their father, who is HIV positive.

Though it sounds like Dr. He had the babies' best interests in mind, other scientists are skeptical. By editing the twins' DNA, Dr. He violated an agreement among CRISPR scientists that they would not edit human embryos, as experts feel that we don't know enough about the effects of genome editing in humans to edit genes that could be inherited. There are also established ways to prevent transmission of HIV from fathers to children that don't involve genome editing, so some question if the girls' CRISPR edits were necessary. Others suspect that Dr. He's efforts were designed to make him famous, rather than to help the twins.

Use the following articles and video to begin your research into the recent editing of humans in China. You can also reference some of the articles you read as part of your general research into CRISPR ethics outlined on the previous page.
- https://www.youtube.com/watch?v=th0vnOmFltc
- https://www.nytimes.com/2018/11/26/health/gene-editing-babies-china.html
- https://www.nature.com/articles/d41586-018-07545-0
- https://www.nytimes.com/2018/12/05/health/crispr-gene-editing-embryos.html
- https://interestingengineering.com/crispr-might-have-made-chinas-designer-babies-smarter

Do you think Dr. He's actions were ethical? Do you think CRIPSR editing of humans should be allowed now or in the future? If so, are there certain edits that should or should not be allowed? Cite at least two additional sources to support your arguments.

Present your findings

Finally, you will present your group's findings to the class. Use your independent research and our in class discussions to prepare a presentation that includes:

1. An explanation of how CRISPR technology can be used to edit genomes and potentially solve current problems in society.
2. Three examples of challenges CRISPR technology has the potential to solve and discuss how its successful implementation would provide an advantage over current approaches to these challenges.
3. A discussion of ethical considerations for at least one of your three examples from part 2 as well as the recent example of CRISPR editing of humans in China. Include at least one ethical consideration related to each of the examples discussed
4. Finally, state your position (claim) on the scientific community should continue to develop CRISPR technologies to tackle current societal challenges. Support your position weighing the benefits of CRISPR technology against concerns for using this technology.

Ethics of CRISPR – Presentation Rubric

	No Evidence	Approaching	Proficient	Advanced
CRISPR Basics	The presentation does not meet <u>any</u> of the proficient level criteria.	The presentation meets <u>some</u>, but not all of the proficient level criteria.	The presentation… ❏ **describes** the intended function ❏ **discusses** role of components of system ○ Cas9 complex ○ guide RNA ○ target DNA ○ PAM sequence	The presentation meets all of the proficient level criteria <u>AND</u> … ❏ **discusses** more than one of the following ethical concerns of using CRISPR for one or both examples discussed ○ ecological ○ ethical ○ legal ○ moral ○ safety
CRISPR example 1	The presentation does not meet <u>any</u> of the proficient level criteria.	The presentation meets <u>some</u>, but not all of the proficient level criteria.	The presentation… ❏ **identifies** an application of CRISPR (problem to be solved) ❏ **describes** a current method for addressing problem ❏ **discusses** advantages of CRISPR over current method	
CRISPR example 2	The presentation does not meet <u>any</u> of the proficient level criteria.	The presentation meets <u>some</u>, but not all of the proficient level criteria.	The presentation… ❏ **identifies** an application of CRISPR (problem to be solved) ❏ **describes** a current method for addressing problem ❏ **discusses** advantages of CRISPR over current method	
CRISPR example 3	The presentation does not meet <u>any</u> of the proficient level criteria.	The presentation meets <u>some</u>, but not all of the proficient level criteria.	The presentation… ❏ **identifies** an application of CRISPR (problem to be solved) ❏ **describes** a current method for addressing problem ❏ **discusses** advantages of CRISPR over current method	
Ethical considerations of CRISPR: CRISPR babies	The presentation does not meet <u>any</u> of the proficient level criteria.	The presentation meets <u>some</u>, but not all of the proficient level criteria.	The presentation… ❏ **discusses** concerns of using CRISPR for engineering humans ○ ecological ○ ethical ○ legal ○ moral ○ safety	
Ethical considerations of CRISPR: Example 2	The presentation does not meet <u>any</u> of the proficient level criteria.	The presentation meets <u>some</u>, but not all of the proficient level criteria.	The presentation… ❏ **discusses** concerns of using CRISPR for chosen application ○ ecological ○ ethical ○ legal ○ moral ○ safety	

Teacher Notes:
Designed to support a lab activity developed by Jessica Stark (NSF Graduate Research Fellow, Ph.D. Candidate, Chemical Engineering, Northwestern University). Written by Jason Forbrook (Waukegan High School) and Jessica Stark.