



# Cell-free gene expression

David Garenne<sup>1,6</sup>, Matthew C. Haines<sup>2,6</sup>, Eugenia F. Romantseva<sup>3,6</sup>,  
Paul Freemont<sup>2,4,5</sup>, Elizabeth A. Strychalski<sup>3</sup> and Vincent Noireaux<sup>1</sup>

**Abstract** | Cell-free gene expression (CFE) emerged as an alternative approach to living cells for specific applications in protein synthesis and labelling for structural biology and proteomics studies. CFE has since been repurposed as a versatile technology for synthetic biology and bioengineering. However, taking full advantage of this technology requires in-depth understanding of its fundamental workflow beyond existing protocols. This Primer provides new practitioners with a comprehensive, detailed and actionable guide to best practices in CFE, to inform research in the laboratory at the state of the art. We focus on *Escherichia coli*-based CFE systems, which remain the primary platform for efficient CFE. Producing proteins, biomanufacturing therapeutics, developing sensors and prototyping genetic circuits illustrate the broader utility and opportunities provided by this practical introduction to CFE. With its extensive functionality and portability, CFE is becoming a powerful and enabling research tool for biotechnology.

## Portability

Use outside a controlled laboratory setting.

<sup>1</sup>Physics and Nanotechnology, University of Minnesota, Minneapolis, MN, USA.

<sup>2</sup>Section of Structural and Synthetic Biology, Department of Infectious Disease, Imperial College London, London, UK.

<sup>3</sup>National Institute of Standards and Technology, Material Measurement Laboratory, Gaithersburg, MD, USA.

<sup>4</sup>London Biofoundry, Imperial College Translation & Innovation Hub, London, UK.

<sup>5</sup>UK DRI Care Research and Technology Centre, Imperial College London, London, UK.

<sup>6</sup>These authors contributed equally: David Garenne, Matthew C. Haines, Eugenia F. Romantseva.

<sup>✉</sup>e-mail:

p.freemont@imperial.ac.uk;  
elizabeth.strychalski@nist.gov; noireaux@umn.edu  
<https://doi.org/10.1038/s43586-021-00046-x>

Cell-free gene expression (CFE) harnesses the transcription and translation machinery of living cells to enable protein synthesis *in vitro* through the expression of natural or synthetic DNA. In the 1960s, the process of translation was extracted from living cells to establish the genetic code embedded into living cells. This also revealed that the complex parts of the machinery of life could be reconstituted in test tubes, without the requirements of a cell membrane or cell reproduction<sup>1–3</sup>. Coupled cell-free transcription–translation systems based on *Escherichia coli* were devised soon thereafter<sup>4</sup>. The major steps needed to prepare and use CFE systems, although considerably optimized, have not changed (FIG. 1).

DNA-dependent protein synthesis by CFE requires isolating the molecular components for transcription and translation from living cells by preparing a cytoplasmic lysate stripped of genetic material and membranes. The lysate is supplemented with buffers that provide the ribonucleosides and amino acids, an ATP regeneration system to sustain translation and several biochemical cofactors to enhance CFE. Such an open and accessible environment offers unrivalled flexibility to adjust the biochemical environment and engineer biological systems by executing synthetic genetic circuits assembled in the laboratory.

CFE already offers powerful capabilities, and areas for further improvement have been clearly identified. The ability to express genetic circuits in the absence of the rest of the genome allows experiments to be carried out in isolation without endogenous DNA. Current CFE technology uses a rapid experimental set-up for high-throughput operations able to test and

analyse many more genetic constructs simultaneously than would be possible using an *in vivo* system, allowing scalability from the femtolitre to litre scale over a broad variety of reaction volumes and containers<sup>5</sup>. A further benefit of CFE is its demonstrated portability<sup>6</sup>. As a result, the capabilities offered by CFE are becoming attractive for a broad range of applications<sup>7</sup>. However, the state of the art allows the preparation of highly efficient CFE systems (>1 mg ml<sup>-1</sup> of synthesized proteins considered an informal metric for highly efficient CFE) from only approximately half a dozen organisms<sup>8</sup>. With a reaction lifetime of less than a day, CFE cannot yet support the synthesis of long-lived genetically programmed autonomous biological systems.

## The new generation of CFE systems

A new generation of CFE systems has emerged in the past 20 years that provides competitive advantages for engineering biology towards practical applications beyond the laboratory<sup>6</sup>. Modern CFE systems, with protein yields up to several milligrams per millilitre<sup>9</sup>, offer broad versatility, scalability and portability<sup>5</sup>. Current CFE systems also extend fundamental studies in other research fields, particularly biology, chemistry and physics<sup>10–12</sup>. These capabilities required several key technical advances. First, the amount of protein produced using modern CFE systems allows synthetic DNA constructs with biological relevance to be expressed, from single short pieces of DNA to long genetic programmes encoding >100 genes<sup>5</sup>. Increased understanding and new methods to provide more biochemical energy to lysates has led to CFE systems capable of routine protein synthesis at the milligram level<sup>13</sup>. Second, although far from

## The PURE system

(The Protein synthesis Using recombinant Elements system). A system that reconstitutes the *Escherichia coli* translational machinery with fully recombinant proteins, with the exception of ribosomes and tRNAs.

understood, the large biochemical parameter space of a CFE reaction is now amenable to machine learning approaches<sup>14</sup>. Finally, proteomic studies have revealed the content of extract-based CFE systems, which may be exploited to improve protein synthesis or enzymatic pathways<sup>15,16</sup>.

CFE can also offer a more convenient prototyping platform for testing engineered biological functions than a living chassis. Although some early demonstrations have shown good transferability between CFE and cell-based function<sup>17–19</sup>, further work is required to predict and design function *in vitro* before transferring to *in vivo* systems. Although still widely used and one of the most efficient, the T7 transcription system is no longer the only option to express genes *in vitro*<sup>9,17,20,21</sup>. The repertoire of DNA parts that can be expressed in CFE extends to natural bacterial regulatory elements<sup>22,23</sup>. CFE is also compatible with modern laboratory automation at both academic and industrial

scales. This enables high-throughput reaction assembly and execution, offering the potential to improve reproducibility, as well as more rapid workflows to screen constructed DNA designs<sup>24</sup> or the activity of enzymes in parallel<sup>25</sup>, for example. CFE systems are gaining increasingly widespread use through their integration into the design–build–test–learn (DBTL) engineering workflow, common to synthetic and engineering biology research. CFE also offers an alternative bioprocessing environment for biomanufacturing, both in the conventional sense of the production of biologics using cell-free bioreactors and for future distributed biomanufacturing<sup>6</sup>. Finally, the emerging efforts to interface CFE with synthetic substrates, such as unnatural amino acids<sup>26,27</sup>, as well as soft and hard materials<sup>28–30</sup>, seek to conceive novel hybrid smart materials, creating innovative opportunities to expand the capabilities of biological and material systems.

## Scope and goals of this review

Achieving cell-free protein synthesis in optimal conditions requires specific skills and knowledge. This Primer offers practical approaches to CFE, emphasizing the first principles of its workflow (FIG. 1) and allowing both new and existing practitioners to assimilate best practices. We provide metrics to benchmark the successful implementation of the major steps in the preparation of CFE systems to facilitate troubleshooting. Although the synthetic yield of a reporter protein remains a typical metric of CFE performance, other metrics can be considered, such as the response time and dynamic range of a biosensor. Other recent reviews provide complementary discussions focused, for example, on applications<sup>7</sup>, choosing the type of cells for specific applications<sup>8</sup> and the history of cell extract methodology<sup>31</sup>.

*E. coli* remains — both historically and currently — the most common model system for CFE and, indeed, for molecular biology broadly. It is also one of the most studied and best characterized organisms<sup>32</sup>. The growth of *E. coli* cultures is rapid and inexpensive. Additionally, the panoply of applications using *E. coli*-based CFE systems is vast and growing. In this Primer, we focus on the basic workflow for preparing and using lysate-based *E. coli* CFE systems, anticipating that mastering these serves as a solid foundation for many current and future applications. The PURE system, a highly popular CFE system reconstituted from purified *E. coli* components<sup>33</sup>, is not the primary focus of this article, although several examples related to this platform are given. To orient the reader, this Primer also presents exemplary CFE applications in biomanufacturing proteins and therapeutics, prototyping genetic parts and engineering biological-based sensors. This is followed by a discussion of reproducibility in CFE experiments and a consideration of current limitations, both for CFE broadly and *E. coli*-based CFE specifically. Finally, we describe exciting opportunities for CFE systems over the next decade.

## Experimentation

*E. coli* CFE systems can synthesize a broad variety of proteins, including soluble proteins such as enzymes or reporters, proteins with disulfide bonds such as

## CFE workflow

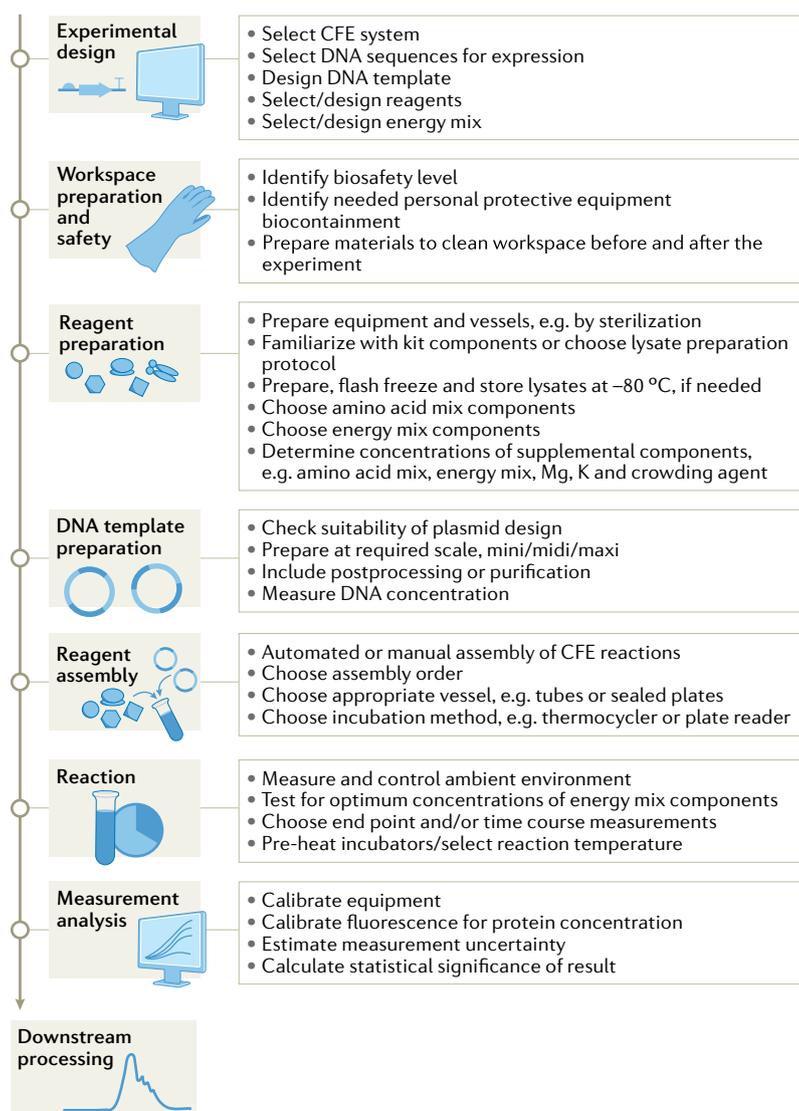


Fig. 1 | **Overview of a typical CFE workflow.** General considerations for users at each step are included. CFE, cell-free gene expression.

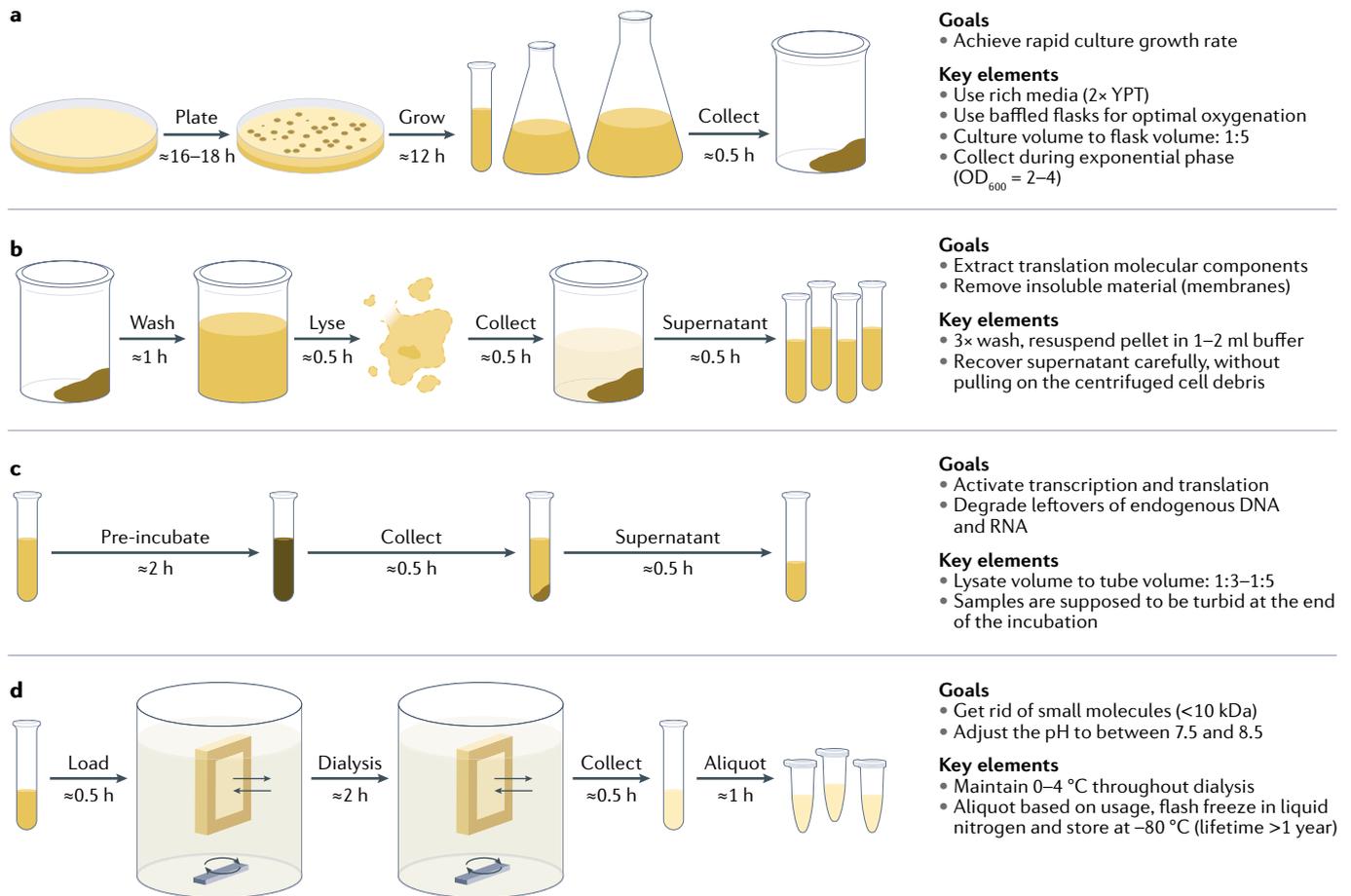


Fig. 2 | **Key steps for *Escherichia coli* lysate preparation.** Process typically requires 3 days and involves four steps. **a** | Cell growth to  $OD_{600}$  of 2–4 in 2.5-l baffled flask and collection by centrifugation. **b** | Washing cells three times in an iso-osmotic buffer at 4 °C, and cell lysis performed at 4 °C using the same iso-osmotic buffer. **c** | Incubation of the lysate in 14-ml culture tubes at 37 °C in a shaker incubator for 2 h. **d** | Dialysis of the lysate at 4 °C in a 2-l beaker with a magnetic bar to stir the buffer, and storage at –80 °C. The dialysis step is optional, providing a lysate with greater efficacy for cell-free gene expression (CFE).

antibodies and membrane proteins such as channels or transporters<sup>34</sup>. *E. coli* CFE systems are unable to produce proteins with post-translational modifications, such as glycosylation, ubiquitylation, lipidation, nitrosylation, methylation and acetylation. These types of protein modification are typically found in eukaryotic organisms. Because *E. coli* lysate is rich in chaperones to assist protein folding<sup>15,35</sup>, soluble proteins produced in *E. coli* CFE systems generally fold robustly. In all cases, folding is a protein-dependent mechanism that can be characterized using technologies such as NMR, crystallography or circular dichroism.

The synthesis of enhanced green fluorescent protein (eGFP) in an *E. coli* CFE system offers an ideal experimental workflow to rapidly master the basic techniques and tools of this technology and provides a good foundation for various applications. *E. coli* CFE systems are ubiquitous, technically accessible and efficient platforms for protein production, producing  $>0.5\text{ mg ml}^{-1}$  of eGFP, for example. Moreover, eGFP is a common, commercially available, fluorescent reporter protein with a maturation time of 5–10 min, facilitating quantitative measurements using typical laboratory instruments. We present this illustrative example to describe

the rudiments of CFE preparation and execution, focusing on meeting important performance metrics and avoiding common mistakes.

Many *E. coli* CFE systems are available commercially (Supplementary Table 1). Although CFE generally presents few safety concerns for the user, the technology itself is sensitive to contaminants, and so standard personal protective equipment is recommended to avoid contamination of the reaction.

#### Custom-made CFE systems

A CFE system requires several components: the lysate, whose preparation spans 3–4 days<sup>19,36</sup>; the aqueous solutions including the energy mix and the amino acid mix, which each take approximately 0.5 days to prepare<sup>19,36</sup>; magnesium, potassium and polyethylene glycol (PEG), which require approximately 1 h to prepare; and the DNA template, which is prepared over approximately 2 days.

***E. coli* lysate preparation.** Lysate is a cytoplasmic extract that provides the necessary molecular components for translation, composed of the ribosomes and translation cofactors<sup>15,35</sup>. Its preparation comprises four major steps (FIG. 2): cell growth, lysis, pre-incubation and dialysis.

A 1-l cell culture usually yields 4–8 ml of lysate, suitable for 12–20 ml of CFE reactions. All flasks, tubes and glassware should be cleaned and sterilized; the use of diethyl pyrocarbonate (DEPC)-treated water during sterilization is not necessary.

Rapid culture growth rates ( $OD_{600} \sim 3$  after 3–4 h) are essential to obtain productive lysates rich in ribosomes<sup>37</sup>. This is achieved using a rich media<sup>38</sup>, such as 2× YPT. The addition of phosphates (22 mmol l<sup>-1</sup> potassium phosphate monobasic and 40 mmol l<sup>-1</sup> potassium phosphate dibasic) to the 2× YT growth medium maintains a constant pH, and lowers the phosphatase activity of the lysate<sup>38</sup>. After assembling the CFE reaction, a lower phosphatase activity prevents rapid degradation of ribonucleosides, ATP and GTP, which are important precursors for nucleic acid and energy metabolites. *E. coli* BL21 DE3 strains offer the convenience of producing T7 RNA polymerase before cell lysis, along with the genetic deletion of several proteases (Supplementary Table 2). Details about cell growth and washing as well as cell lysis methods can be found in Supplementary Tables 3 and 4.

After cell lysis, the supernatant is recovered with a pipette without disturbing debris from the pellet. The supernatant is placed in a new sterile tube and incubated to degrade endogenous mRNAs. The pre-incubation also allows endogenous transcription and translation to come to completion, which activates the lysate<sup>39</sup>. Centrifugation removes DNA (Supplementary Table 5). Although not strictly necessary, a third step uses dialysis prior to storing the lysate to remove molecular components smaller than 10 kDa, which improves protein yields by up to 50%<sup>36</sup> (Supplementary Table 5).

The most productive lysates have a protein concentration of 30–50 mg ml<sup>-1</sup> (REF.<sup>36</sup>), although the preferred lysate protein concentration is approximately 30 mg ml<sup>-1</sup>, representing a tenfold dilution of the *E. coli* cytoplasm<sup>40</sup>. For a lysate with a protein concentration out of this range, preparing a new lysate is recommended. The lysate is added to a CFE reaction to a final concentration of approximately 10 mg ml<sup>-1</sup>, or 33% of the reaction volume if the lysate protein concentration is 30 mg ml<sup>-1</sup>. A cell extract may be verified as cell-free by plating a portion on agar plates without antibiotic; any living cells will grow, enabling estimation of the concentration of living cells remaining in the lysate.

**Supplemental aqueous solutions for a CFE reaction.** Recipes and protocols for both the energy mix and the amino acid mix have been reported (Supplementary Table 6). The energy mix mainly contains the four ribonucleosides and a phosphate donor for ATP regeneration. Although not always necessary<sup>13</sup>, tRNAs may be added to the energy mix, which may improve but does not decrease cell-free protein synthesis. Three other aqueous stock solutions are necessary for CFE reactions and should be prepared fresh at least every month: magnesium glutamate (1 mol l<sup>-1</sup>), potassium glutamate (3 mol l<sup>-1</sup>) and PEG 8000 (40% w/v). Magnesium and potassium are essential ions. Acetate (for example, magnesium acetate) and glutamate (for example, magnesium glutamate) counterions are preferred for *E. coli*

CFE systems, based on empirical results. PEG emulates molecular crowding *in vivo* and is often necessary, due to the substantially lower density of biomolecules in a CFE reaction than in a cell<sup>41</sup>. Ficoll and Dextran can be used as alternative molecular crowding agents. Although increasing molecular crowding is not always necessary, it is recommended to verify whether a molecular crowding agent increases CFE performance.

### Preparation of the DNA template

The proper choice, preparation and storage of DNA templates is critical to achieving high-performance CFE. The DNA is first prepared using a commercially available DNA prep kit. To obtain high-quality DNA for CFE, subsequent clean up is recommended using a PCR purification kit, followed by elution with autoclaved 18.2 MΩ cm<sup>-1</sup> water. A high-quality DNA template is indicated by ratio absorbance measurements at  $A_{260/280}$  of 1.8–2.0. Although both circular plasmids and linear DNA may be used<sup>42–44</sup>, plasmids enable up to two-fold<sup>9</sup> greater CFE activity than linear PCR products. Plasmids are more stable in CFE reactions than linear DNA, because linear DNA can be degraded by protease RecBCD. Nonetheless, linear DNA templates, especially from PCR, offer attractive advantages for many applications, owing to their more rapid and less costly preparation over circular plasmid DNA. Furthermore, simple affordable methods are now available to limit degradation in CFE reactions<sup>43,44</sup>.

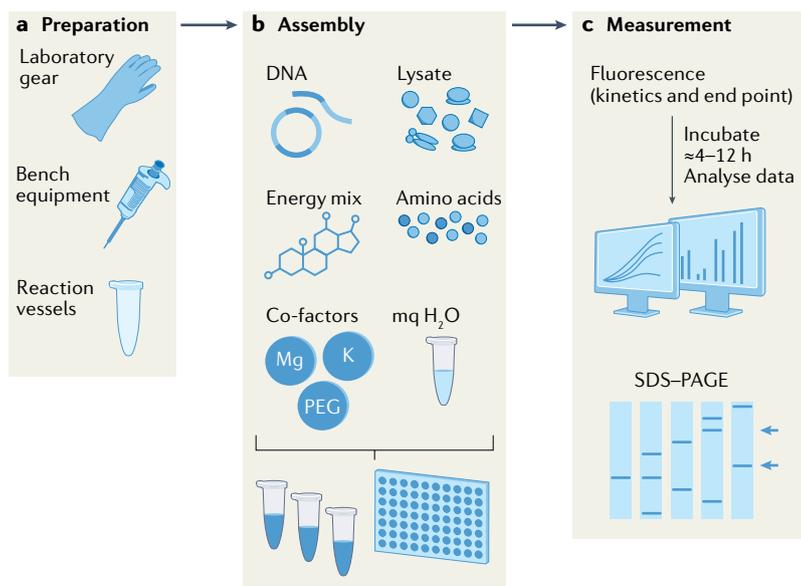
Components of the T7 transcription system — the T7 RNA polymerase and T7p14 promoter — offer universal, powerful and accessible CFE. Therefore, a plasmid with the gene *egfp* cloned under the T7p14 promoter/untranslated region and including the T7 terminator is ideal for assessing the efficiency of a CFE system (Supplementary Table 7). A DNA stock solution at >100 nmol l<sup>-1</sup> is ideal, as a plasmid concentration in a CFE reaction varies typically from 0 to 20 nmol l<sup>-1</sup>. To aid in later protein purification, for example by affinity chromatography, a 6-histidine tag or other affinity tag may be added to the DNA coding sequence. The protocol for such purification in CFE does not differ from standard cell-based purification protocols.

### Storage of CFE components

To avoid degradation of CFE components, each of the components must be stored at their appropriate storage temperature. The lysate, energy mix and amino acid mix solutions should typically be stored at –80 °C. At this temperature, these solutions have a shelf-life of >1 year. None of these components should be autoclaved. The potassium glutamate, magnesium glutamate and PEG stock solutions can be stored at room temperature and should be autoclaved prior to use. The DNA stock solutions should be aliquoted in small fluid volumes appropriate to their usage and stored at –20 °C.

### CFE reaction planning

In addition to the reaction components themselves, performing a CFE reaction requires basic benchtop equipment (FIG. 3). When prepared by hand, the volume of a CFE reaction ranges from 10–20 μl to 1 ml. Preparing



**Fig. 3 | Equipment needed for CFE a | Experiment preparation.** Necessary equipment includes disposable gloves, face masks, safety glasses, calibrated pipettes, tube racks, a vortexer with adjustable speed allowing gentle vortexing (<math><4,000\text{ rpm}</math>) for short time periods (<math><10\text{ s}</math>) to mix the reaction, a minifuge to spin solutions down briefly after vortexing and 70% ethanol to clean and sterilize the work area. **b** | A custom-made cell-free gene expression (CFE) system typically including six sets of different aqueous solutions including the DNA template, lysate, energy mix, amino acid mix, cofactors and magnesium. The reactions are assembled in 1.5-ml tubes or microwell plates. **c** | For time course fluorescence measurements, a plate reader should be programmed before assembling the reaction to reach the desired incubation temperature. Once loaded onto a well plate, reactions are directly incubated and monitored in real time on the plate reader. For end point fluorescence measurements, the tubes or well plate are placed in an incubator. After incubation, data are analysed and plotted. For non-fluorescent proteins or peptides, sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) offers a suitable technique to analyse CFE products. PEG, polyethylene glycol.

CFE reactions smaller than  $10\ \mu\text{l}$  by hand is not advised because the volume of some components is smaller than  $1\ \mu\text{l}$ , which introduces errors and variability in the reaction. Oxygen is necessary for most of the components of CFE and for eGFP maturation<sup>45</sup>. The ratio of the reaction volume to the available air should be in the order of 100, such as a  $10\text{-}\mu\text{l}$  reaction placed in a 1.5-ml vial. Reaction volumes of  $2\text{--}5\ \mu\text{l}$  in wells of  $300\text{--}500\ \mu\text{l}$  on a V-bottom well plate offer the convenience of not requiring shaking to oxygenate the CFE reaction; larger reaction volumes  $>100\ \mu\text{l}$  usually require shaking at approximately 100 rpm for adequate oxygenation. For custom-made kits, a typical reaction is assembled on ice, to better approximate the same start time for the reaction for all of the wells when their temperature is raised simultaneously to the reaction temperature (FIG. 3). Most conditions for CFE reactions have been empirically established or optimized (Supplementary Table 8). The reaction vessels should be tightly closed to avoid evaporation. The incubation device should provide uniform temperature to the whole reaction vessel to avoid condensation. It is therefore not recommended to incubate CFE reactions on a dry block or in a water bath, where, for example, the cap is not at the same temperature as the rest of the vessel. Rather, the CFE reactions should be placed inside an incubator at uniform temperature.

CFE reactions have been observed to be most sensitive to magnesium and DNA template concentrations, and a reaction with weak activity should be checked for concentrations that differ significantly from optimal values. For magnesium, a few millimoles per litre above or below the optimal concentration can result in at least a twofold decrease in protein synthesis (Supplementary Fig. 1). For plasmids, a few nanomoles per litre below the optimal concentration can also result in a several-fold decrease in protein synthesis (Supplementary Fig. 2). Magnesium is essential in living systems — in stabilizing ribosomes<sup>46</sup>, acting as a metal cofactor in enzymatic reactions or in electrostatic shielding of negatively charged polymers such as DNA. Finding the optimal magnesium concentration of a custom-made CFE system is essential and done by systematically varying the magnesium concentration, typically in increments of  $1\text{--}2\ \text{mmol l}^{-1}$  between  $0$  and  $15\ \text{mmol l}^{-1}$ , to screen for an optimal value in the resulting CFE (Supplementary Fig. 1). Determining the optimal plasmid concentration is achieved similarly, with plasmid concentrations typically ranging from  $0$  to  $20\ \text{nmol l}^{-1}$  in increments of  $2\text{--}4\ \text{nmol l}^{-1}$  and fixing magnesium at the optimal concentration (Supplementary Fig. 2).

#### Data collection

Fluorescence measurements using a plate reader and sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) are the most convenient measurement methods to determine whether a CFE is active or a protein has been produced. eGFP, along with the truncated derivative deGFP<sup>9</sup> or superfolder GFP (sfGFP)<sup>47</sup>, has a typical limit of detection of  $10\ \text{nmol l}^{-1}$  for a  $2\text{-}\mu\text{l}$  reaction on a 96-well plate measured using a Biotek H1m or Neo2 plate reader, or similar. A well may be measured every 1–5 min, depending on the needs of the study or application. The choice of signal amplification, sometimes defined as a photomultiplier tube gain, for the plate reader depends on the activity of CFE and should be tested to achieve a detector signal above background and below saturation of the detector. We recommend testing different amplifications using samples of purified eGFP (Cell Biolabs #STA-201) at different concentrations in volumes equal to the CFE reaction volume. For instance, a gain of 50–60 on a Biotek H1m is suitable for measuring protein synthesis of  $>1\ \mu\text{mol l}^{-1}$  eGFP for a reaction volume of  $2\text{--}3\ \mu\text{l}$ . End point measurements taken after the CFE reaction reaches steady state may be executed similarly by testing different gains. The negative control consists of a blank CFE reaction with no DNA template, which allows the measurement of background fluorescence levels. A positive control consists of a CFE reaction without DNA template and  $1\ \mu\text{mol l}^{-1}$  of purified eGFP. SDS–PAGE measurements verify the production of the target protein in a CFE reaction, with a typical detection limit of  $\sim 100\ \text{ng}$  with Coomassie blue staining and  $\sim 10\ \text{ng}$  with silver staining. The percentage of acrylamide is adjusted based on the size of the CFE product: 20% for peptides of  $2\text{--}5\ \text{kDa}$ , 16% for small ( $5\text{--}15\ \text{kDa}$ ) proteins, 12% for eGFP at  $27\ \text{kDa}$  and 7% for proteins  $>100\ \text{kDa}$ .

## Results

Although many techniques are available to determine whether a protein or a peptide has been produced in a CFE reaction, SDS-PAGE and fluorescence are by far the most convenient and affordable, in addition to being quantitative. Here, we describe the advantages and most important properties of these two techniques that are the most used in CFE research. Besides these two methods, the function and activity of a CFE reaction product has to be characterized by assays specific to the protein or peptide synthesized. For instance, SDS-PAGE and fluorescence can be accompanied by sophisticated techniques such as mass spectrometry<sup>15,35</sup>.

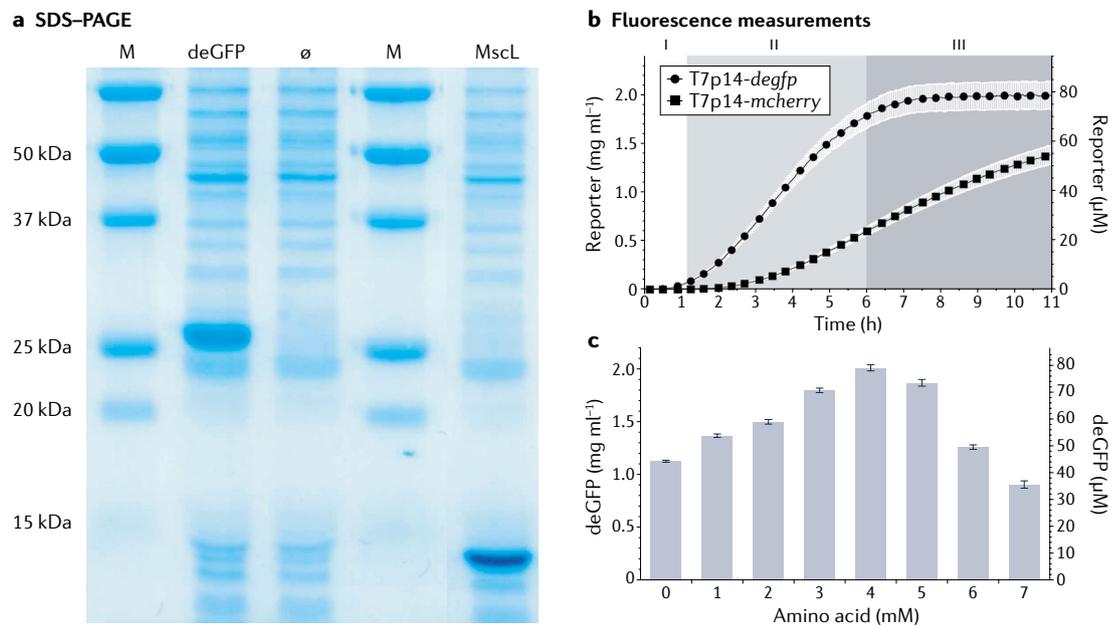
### SDS-PAGE

**Protein quantification on SDS-PAGE.** SDS-PAGE remains one of the most useful techniques to determine whether a protein has been produced in a CFE reaction, when the product is not fluorescently labelled (FIG. 4). Peptides from 10 amino acids (1 kDa) up to proteins of 250 kDa can be visualized by SDS-PAGE with 20% and 5% acrylamide, respectively. Basic Coomassie SDS-PAGE has a protein detection limit of 100 ng. Thus, a protein concentration as low as 20  $\mu\text{g ml}^{-1}$  (5  $\mu\text{l}$  of the CFE reaction) can be detected. A silver staining gel has a protein detection limit of 10  $\text{ng}\mu\text{l}^{-1}$ . Relatively precise quantification is possible using either a protein calibrant

such as bovine serum albumin (BSA)<sup>48</sup> or the pure protein of interest if available. The BSA quantification approach is inexpensive as it requires a basic computer scanner and free software such as ImageJ to analyse the bands on the image produced by the scanner. For proteins synthesized at low concentrations, the protein band can be excised from the gel and processed using advanced techniques such as mass spectrometry<sup>49</sup>.

Proteins or peptides synthesized in CFE systems can be purified prior to SDS-PAGE analysis. A protein having a 6-histidine tag in the amino terminus or carboxy terminus, for instance, can be purified by affinity chromatography using standard procedures<sup>50</sup>. Other tags, such as GST or Strep-tag, can also be employed<sup>51–54</sup>.

**Validation of the results.** Choosing the right polyacrylamide percentage is essential to achieve the desired separation of the protein band of interest from the CFE reaction background. A rough estimation of the protein size and amount produced by SDS-PAGE is done on a single gel only, as comparing the intensity of a band from one gel with another is not recommended, owing to variation in staining between gels. Quantification requires diluting the completed CFE reaction in a blank CFE reaction containing all of the components except plasmids and running SDS-PAGE using aliquots of different dilution strength in order to create several bands



**Fig. 4 | Representative CFE results.** Cell-free gene expression (CFE) kit: myTXTL (Arbor Biosciences) contains all necessary components for CFE except the DNA. All genes expressed through the T7 promoter ( $2\text{ nmol l}^{-1}$  T7p14 gene). **a** | Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) shows production of the reporter protein deGFP (25 kDa) and the *Escherichia coli* membrane protein MscL (15 kDa), marker (M) and Ø blank CFE reaction (no gene expressed). Reactions incubated for 12 h on a V-bottom well plate in a static incubator. **b** | Time course and end point fluorescence measurements of deGFP (circles) and mCherry (squares) production. Fluorescence signal for each well in arbitrary units averaged over five reads and converted into the protein concentration using the calibration shown in Supplementary Fig. 5. Bars correspond to one standard deviation of measurements for triplicate wells (triplicate technical replicates for one biological sample) at each time point. Reactions incubated on a V-bottom well plate in a Biotek H1m plate reader. Shaded regions I, II and III mark the three different phases of cell-free protein synthesis for deGFP. **c** | End point measurement of deGFP production as a function of the amino acid concentration added to the CFE reaction. Reactions incubated for 12 h on a V-bottom well plate in a static incubator and measured on a Biotek H1m plate reader. Equimolar 20-amino acid mix prepared according to REF.<sup>190</sup>. eGFP, enhanced green fluorescent protein.

of the protein of interest at different intensities with the same background.

### Fluorescence measurements

**Kinetics and end points.** Fluorescence offers a practical, rapid, accessible and scalable measurement method for CFE (Supplementary Table 9). Beyond proteins in solution, fluorescence measurements may also quantify and reveal the location of membrane proteins. For example, only membrane proteins folded correctly and embedded in a lipid membrane can result in an increased fluorescence signal<sup>55</sup>. A fluorescent reporter protein, such as eGFP, also allows measurements of dynamics in CFE, as well as the performance of regulatory elements<sup>18</sup> and gene circuits<sup>56–58</sup>.

The most important factors in determining the length of time during which protein synthesis remains active in a CFE reaction are the incubation temperature and the strength of the ATP regeneration system (Supplementary Table 6). When a reporter protein, such as eGFP, with a folding time of 5–10 min is synthesized at an optimal temperature (usually 29–37 °C; Supplementary Table 1), the typical reaction time course includes three phases (FIG. 4): an initial phase of 0–60 min, during which the synthesized mRNAs are produced; a phase of 1–8 h that appears as a quasilinear accumulation of the produced protein; and a plateau phase that indicates the reaction has run to completion. CFE typically stops after several hours for multiple reasons: a finite amount of essential nucleic acid and protein building blocks in the reaction; a decrease of the adenylate energy charge index affects translation<sup>59</sup>; and a drop in reaction pH caused by accumulation of by-products of transcription–translation<sup>60</sup>.

Proteins have different maturation times that can affect the length of each of the three phases of a CFE reaction. Red fluorescent proteins, for example, are known to fold slowly (~0.5–2 h) compared with eGFP<sup>61</sup>. When mCherry is synthesized in the same conditions as eGFP, the three phases span different lengths of time due to the slow maturation time of the former (FIG. 4). In the case of mCherry, the plateau occurs after ~16 h. CFE of a water-soluble protein follows dynamics within the time spans of eGFP and mCherry. Consequently, a CFE reaction may require incubation for up to 16 h to ensure completion.

Mathematical models describe the first two parts of the typical time course, during which the CFE reaction is not limited by resources or the reaction physiology, such as a decrease in pH. In particular, CFE observed in the linear accumulation phase from 1 to 8 h reaches a maximum reaction rate, limited by either the strength of the ATP regeneration system<sup>62</sup> or saturation of the translation machinery<sup>63</sup>. The potential for computational modelling to provide information necessary to optimize and augment lysate-based *E. coli* CFE systems is not currently being explored, however. The consumption of biochemical nutrients (for example, ribonucleosides, amino acids) affects the duration of CFE, and thus how long protein synthesis remains active. Because a CFE reaction contains between 50 and 100 different biochemical species, analysing the use of a particular component, such as

ATP, requires different assays to assess their concentrations. We anticipate that technologies enabling simultaneous characterization of multiple components, such as metabolomics, will soon provide thorough descriptions of biochemical nutrient utilization in CFE reactions. We also anticipate time course measurements to aid in understanding and optimization of biochemical nutrient utilization.

Because it is a purified system, the PURE system offers an easier platform for modelling the biochemical reactions associated with the dynamics of transcription and translation. The PURE system was recently computationally modelled by a biochemical reaction network of about 1,000 reactions<sup>64</sup>. Although still involving many adjustable parameters, the PURE system is considered more tractable than lysate-based systems. The analysis of such a complex dynamical system revealed, for instance, the robustness and insensitivity of CFE to variations in transcription–translation kinetic parameters<sup>65</sup>. Independently, a mechanistic computational model of CFE revealed some of the factors limiting CFE<sup>66</sup>, including inefficient usage of ribosomes and depletion of key initiation factors.

End point measurements are often useful to analyse how a particular biochemical or other parameter impacts protein synthesis. The end point synthesis of eGFP as a function of plasmid template concentration follows two phases: a linear function of protein yield as a function of time from 0 to 5 nmol l<sup>-1</sup> of DNA template, and a saturated plateau in the function of protein yield as a function of time at larger DNA template concentrations (Supplementary Fig. 2). Above a threshold amount of DNA template, the protein synthesis rate and yield do not increase, owing to an apparent saturation in translation machinery<sup>63</sup>. Protein production typically peaks at amino acid concentrations of ~3–4 mmol l<sup>-1</sup> each (FIG. 4). As for magnesium, template DNA and amino acids, any biochemicals not necessary for CFE but essential for studying a particular gene product or gene circuit function may be assayed and optimized by testing a range of concentrations and measuring the impact on synthesis of a reporter protein. As an example, chemicals such as NADPH of nickel ions that are not part of the CFE recipe can be tested by doing such a range of concentrations. The results show that CFE remains strong up to a concentration of 1 mmol l<sup>-1</sup> and 2.5 mmol l<sup>-1</sup> of NADPH and nickel ions, respectively (Supplementary Figs 3 and 4).

Expressing a fluorescent RNA aptamer encoded into the mRNA of the reporter is also useful to test for active transcription. Spinach<sup>67,68</sup>, broccoli<sup>69,70</sup> and malachite green<sup>63,71</sup> aptamers all function in *E. coli* CFE systems and offer a practical resource for troubleshooting CFE efficiency, especially when a reporter protein is not present or detectable in a CFE reaction. Fluorescent aptamers also allow real-time measurements of transcription; when coupled to reporter protein measurements, this allows simultaneous measurement of both transcription and translation.

**Fluorescence quantification.** Many measurements of CFE rely on fluorescence assays, which require calibration to calculate meaningful quantities from arbitrary

or relative units. Calibration of equipment and measurements increases confidence in the results, aids in comparability between measurements, and can allow conversion from relative units to absolute units or the International System of Units (SI) units, such as nanomoles per litre or micromoles per litre. For example, fluorescence intensity measured using a plate reader is often calibrated to the concentration of a fluorescent protein in solution using a commercially available reporter protein, such as eGFP (Cell Biolabs #STA-201 or Biovision #4999-100). First, the concentration of the undiluted stock of eGFP, usually provided at  $1 \text{ mg ml}^{-1}$ , is confirmed by measurement with a calibrated spectrometer, to see whether these values agree to within the uncertainty<sup>72</sup>. The values can differ significantly, and the manufacturer's value may be disregarded in favour of more careful measurement of the stock concentration for subsequent use in calibrating the plate reader. Second, an eGFP calibration of the measurement device is carried out (Supplementary Fig. 5). Although Microsoft Excel is sufficient to analyse the data, other scientific computing software may also be used, such as MATLAB, Prism, KaleidaGraph or Origin.

**Confidence.** Confidence in a result is determined through measurement assurance and uncertainty analysis. Ideally, an experiment should be designed to fit the intended purpose with the greatest confidence in the results. Good experimental design includes, for example, positive and negative controls, appropriate numbers of technical and biological replicates, use of formal design of experiment approaches, consideration of systematic and random error, and statistical analyses. Although a quantitative result with uncertainty to meet target specifications may require careful measurements, a rapid, qualitative result to determine whether a reaction yields any measurable fluorescence at all is sometimes sufficient. Specific steps in a CFE workflow and data analysis should be guided by ensuring fitness for the intended purpose.

The ability to reproduce a measured result provides a good check on confidence in CFE performance. A CFE experiment might be repeated using the same reagents and protocol but performed by a different researcher. Alternatively, the experiment can be repeated by a different researcher and using reagents, including cell lysate, prepared separately but using the same protocols. Consider the example of determining the magnesium concentration that maximizes the protein yield (Supplementary Fig. 1). Depending on the intended use of the CFE system, agreement of the optimal magnesium concentration to within  $\sim 1 \text{ mmol l}^{-1}$  can indicate good reproducibility, as this has been shown to result in identical protein yields to within the experimental uncertainty<sup>73</sup>.

An outlier measurement is characterized by a data point that stands far apart from other data points in the same set of repeats. Outlier measurements occur frequently in CFE, especially during fluorescence measurements. An experiment with outlying measurements should be repeated, especially when the outlying result cannot be attributed to an obvious deviation from the

protocol or error. Dust and manual pipetting errors, for example, commonly lead to unexpected results.

## Applications

Producing proteins is the most basic application of CFE. Compared with other methods for protein production, CFE has several desirable attributes that suit a range of applications (FIG. 5). These include, for example, rapid timescales for protein synthesis<sup>74,75</sup>, consistent behaviour from microlitre to 100-l scales<sup>76,77</sup>, desiccation tolerance, direct access to the reaction environment and ease of purification<sup>78</sup>. In this section, we discuss how these and other attributes enable applications accessible to those entering the field.

## Membrane and other proteins

The study of proteins is critical to our understanding of living processes. This often entails protein synthesis and downstream analysis yielding rich data sets that report biochemical properties, including structural data<sup>79</sup>, kinetic parameters<sup>80</sup> and equilibrium binding constants<sup>81</sup> amongst others. Recombinant expression of specific proteins can result in low yields or misfolding. This is particularly true of membrane proteins, which are often toxic when overexpressed<sup>82</sup> and yet account for nearly a third of all gene products<sup>83</sup>, illustrating their importance. The ability to directly modify the lipid composition of the CFE reaction and to operate under conditions that are not conducive to life makes CFE an ideal tool for membrane protein overexpression. This application of CFE was highlighted by the synthesis of 120 *E. coli* membrane proteins<sup>84</sup>, in which 63% of the targets were synthesized by CFE, compared with the 44% achieved via synthesis *in vivo*. Subsequent studies have further highlighted the capability of CFE for efficient membrane protein production<sup>85</sup>. In particular, nanodiscs have rendered CFE of membrane proteins much more accessible<sup>86</sup>.

In addition to membrane proteins, CFE has also been applied to the synthesis of proteins that are cytotoxic<sup>87</sup> or form inclusion bodies<sup>88</sup>, making purification challenging. CFE is also ideally suited to high-efficiency <sup>15</sup>N labelling of targeted amino acids for 3D structure determination by NMR<sup>89</sup>, because rates of amino acid metabolism in CFE reactions are lower, leading to a reduction in isotope scrambling<sup>90</sup>. In principle, there are other techniques that are compatible with CFE, such as cryogenic electron microscopy.

## Education

The synthesis of proteins using CFE is now used in educational kits<sup>91,92</sup>. BioBits kits, for example, fill a significant gap in the available resources to teach molecular and synthetic biology to high school students. Considering that BioBits kits use freeze-dried CFE and plasmids, they do not require cold-chain distribution or sterile conditions to function, factors that are often limiting for implementation in educational environments. Concepts such as tuning of gene expression are illustrated by getting users to vary the plasmid DNA concentration in CFE reactions. Different proteins and materials can be produced, including fluorescent reporters, fragrances

### Nanodiscs

Stable particles in aqueous solution containing a phospholipid bilayer encircled by protein belts.

### Inclusion bodies

Aggregates of protein that form within cells; in bacteria, generally the result of recombinant protein overexpression.

### Isotope scrambling

The process by which <sup>15</sup>N atoms are transferred to alternative amino acid types within the protein.

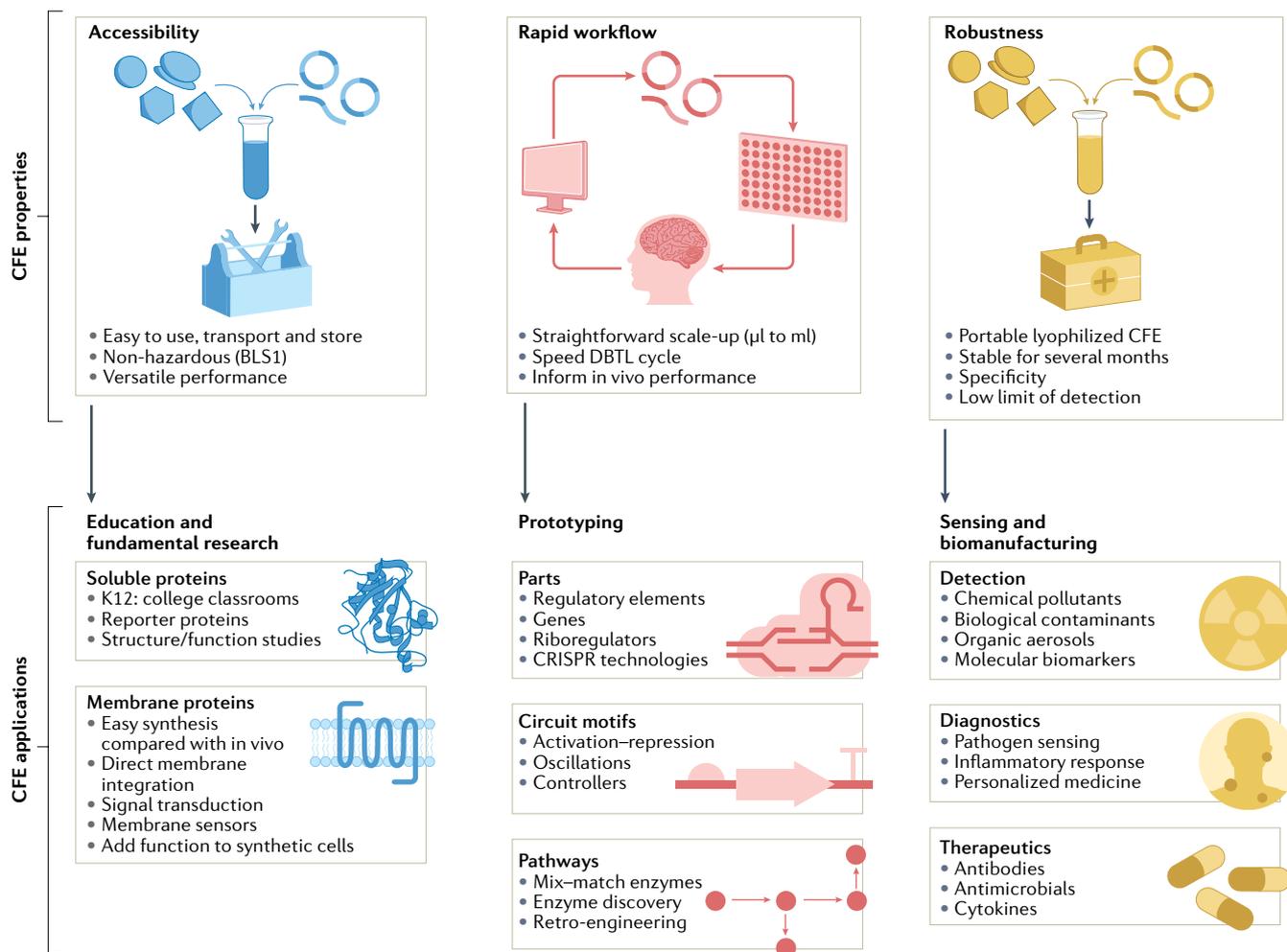


Fig. 5 | **Current scope of applications for CFE.** Advantages of cell-free gene expression (CFE) include accessibility, rapid workflow and robustness, each of which contributes to applications of CFE in various settings. BSL1, biosafety level 1; DBTL, design–build–test–learn.

and hydrogels<sup>91</sup>, each of which stimulate different senses, thus engaging users.

### Prototyping biological devices

A central goal of synthetic biology is the engineering of genetic parts and devices, including safer chimeric antigen receptor (CAR) T cell therapies<sup>93</sup>, enzymatic cascades for biopharmaceutical synthesis<sup>94</sup> and microorganisms that synthesize opioids<sup>95</sup>, to name a few<sup>96,97</sup>. A major bottleneck in this process is the time required to iterate through DBTL cycles, characterizing the effects of genetic variants. Depending on the host used for expression or being engineered, variants require days or weeks to generate prior to their analysis for functionality<sup>57</sup>.

An alternative approach to the characterization of genetic circuits in living cells is to analyse them using CFE. Given direct access to the reaction components, it is feasible to eliminate cloning procedures altogether by encoding variants on linear DNA, whose degradation by RecBCD in *E. coli* CFE systems can be inhibited<sup>42–44</sup>. Linear DNA is produced in less than 8 h, which represents significant time savings in the DBTL cycle. This approach was used in the development of five-node

genetic ring oscillators, of which 95% oscillated for up to 72 h in vivo<sup>57</sup>. CFE is also ideal to rapidly prototype single and combinations of regulatory elements<sup>56</sup>. One limitation of using CFE for this approach pertains to potential growth-associated stress effects when genetic parts or circuits are transferred to in vivo systems, which are difficult to quantify using CFE<sup>24,57</sup>. Furthermore, best practices for ensuring prototyped parts and devices behave equivalently in vivo remain to be established, for example whether multiple or single plasmids should be used when transferring multiple parts.

CFE has also been applied to prototyping CRISPR technologies, such as the efficiency of a guide RNA (gRNA) or the function of a CRISPR enzyme. These technologies have many potential applications in fundamental research and as therapeutics, with treatments for genetic diseases such as sickle cell anaemia and  $\beta$ -thalassaemia<sup>98</sup>. Nevertheless, methods for characterizing CRISPR technologies in terms of protospacer adjacent motif (PAM) sequence requirements, gRNA specificity and the activity of new Cas or anti-CRISPR proteins are time-consuming, requiring days to weeks to perform, and are limited in scale, given the need for protein purification<sup>99,100</sup> or cell

Protospacer adjacent motif (PAM). A DNA sequence of 2–6 bp immediately following the DNA sequence targeted by a CRISPR enzyme.

culture<sup>101</sup>. CFE addresses this issue by enabling researchers to prototype CRISPR technologies at scale using small volumes, via the synthesis of Cas proteins, gRNAs and a reporter<sup>18,102,103</sup>. Specifically, Cas nuclease cleavage of PAM sequences associated with reporter synthesis is measured via a decrease in reporter signal. Good correlation between data generated from these assays and corresponding *in vivo* data has been reported for single gRNAs with different activities ( $R^2 = 0.9$ )<sup>18</sup>. Although it has been suggested that this quantitative relationship may not extend to higher organisms, such as eukaryotes, an inhibitor of CRISPR–Cas12a prototyped in CFE was nevertheless functional in mammalian cells<sup>103</sup>.

### Sensors

There is a growing need to monitor the presence of contaminants and pathogens in our environment using inexpensive, safe and reliable technologies. Biological systems have evolved the capacity to detect analytes using regulators that function at the transcriptional, translational and enzymatic levels. Synthetic biologists aim to leverage this functionality, creating sensors that fulfil the above requirements. Although this was originally achieved using whole-cell biosensors<sup>104</sup>, the application of these sensors is limited, for example, by risks related to releasing genetically modified organisms into the environment<sup>7</sup>. Moving to CFE-based systems has several advantages, including improved safety, avoidance of cold-chain storage through desiccation and the ability for incorporation into inexpensive materials, such as paper or cloth<sup>105</sup>.

CFE-based biosensors have been recently devised to detect compounds such as antibiotics<sup>106</sup>, endocrine-disrupting chemicals<sup>107</sup> and herbicides<sup>108</sup>. These biosensors rely on different strategies to convert analyte concentrations into detectable signals. An example of CFE-based biosensing involved the detection of the quorum sensing molecule *N*-(3-oxododecanoyl)-L-homoserine lactone (3OC12-HSL) isolated from samples from patients with cystic fibrosis<sup>109</sup>. This quorum sensing molecule is produced by *Pseudomonas aeruginosa* infections and has been found at higher concentrations in patients who are hospitalized than in those who are stable<sup>110</sup>. Using a CFE biosensor synthesizing a transcription factor activated by 3OC12-HSL, along with a downstream reporter, this quorum sensing molecule was detected at concentrations of nanomoles per litre, which are comparable with the sensitivity of liquid chromatography–mass spectrometry and suitable for clinical monitoring<sup>109</sup>. Although CFE-based biosensors show clear potential, to the best of our knowledge no examples of regulatory approval exist. Potential reasons for this may include specificity issues or variability in performance<sup>111</sup>.

### Therapeutics

Protein-based biopharmaceuticals have a critical role to play in the prevention and treatment of disease. Between 2015 and 2018, more than 50% of newly approved biopharmaceuticals were monoclonal antibodies, a subset of this class of therapeutics<sup>112</sup>. It is therefore not surprising that CFE is being applied to the development

and production of biopharmaceuticals, including antimicrobial peptides<sup>113</sup>, cytokines<sup>77</sup>, vaccines<sup>114</sup> and antibodies<sup>15</sup>.

A key challenge in the production of biopharmaceuticals using bacterial CFE was the synthesis of proteins with disulfide bonds. The majority of biopharmaceuticals are eukaryotic extracellular proteins requiring this post-translational modification. Disulfide bond formation, however, is incompatible with the native, reducing environment of the bacterial cytoplasm, leading to the synthesis of misfolded, insoluble aggregates<sup>116</sup>. To achieve titres in the range of 100  $\mu\text{g ml}^{-1}$ , several modifications to bacterial CFE were required, including deletion of the glutathione reductase gene<sup>117</sup>, inactivation of remaining reductases via alkylating agents<sup>118,119</sup>, a tailored concentration of glutathione and supplementation of disulfide isomerase to improve folding<sup>120,121</sup>. This approach has yielded protein targets that contain multiple disulfide bonds with 95% correctly folded<sup>77</sup>. Furthermore, conditions for large-scale production have been identified at the microtitre-plate scale, given that the rate of protein production was constant from 250  $\mu\text{l}$  to 100l (REF.<sup>77</sup>). This contrasts with process optimization for biopharmaceutical production using mammalian cell lines, which faces multiple challenges in moving between equivalent scales<sup>122</sup>.

CFE has also been suggested as a platform for point of care biopharmaceutical manufacturing<sup>123</sup>. Considering the production and distribution of biopharmaceuticals traditionally requires many weeks or months, point of care production has the potential to respond more rapidly to changes in demand, as could occur during an infection outbreak or natural disaster. Freeze-drying the components of the CFE reaction ensures stability for an additional 60 days at room temperature<sup>124</sup>, enabling storage and transportation to the end user in the absence of a cold chain. This technology is suited to therapeutic production in challenging areas, such as battlefields or remote regions. Using freeze-dried CFE, antimicrobial peptides, vaccines, antibodies and small molecules have been produced<sup>6</sup>.

### Reproducibility and data deposition

As with other experimental biological systems, reproducibility remains a concern for the preparation, measurement and application of CFE. The adoption of good laboratory practices as well as the development of standards and measurement assurance for CFE are expected to mitigate many of the factors contributing to performance variability. As high-quality data sets become increasingly available for CFE, data and metadata repositories should emerge as important community resources.

### Reproducibility

Researchers in CFE have identified a lack of reproducibility as a significant barrier to adoption and progress (TABLE 1). Good measurement practices, as well as the development of standards for measurements and data reporting, are important for reproducible and comparable CFE measurements within and across laboratories<sup>111</sup>. Indeed, improvements to measurement quality already

**Quorum sensing molecule**  
A molecule used to communicate between individual bacterial cells.

**Monoclonal antibodies**  
Antibodies made by cloning a unique white blood cell.

inform a greater understanding of the components and function of CFE reactions, offering avenues for improved performance<sup>39,56,60,73,76,125–127</sup>, and published protocols and best practices are becoming increasingly available<sup>19,36,76</sup>

(Supplementary Table 9). Addressing issues surrounding reproducibility will help distinguish true variability in biological functions, providing natural limits on the engineerability of CFE systems. This knowledge

Table 1 | **Common problems and sources of variability for each step of the CFE workflow**

Common sources of poor CFE	Troubleshooting
<b>Experimental design</b>	
Inadequate CFE system for desired application	Determine necessary conditions to synthesize desired protein (soluble or membrane protein, disulfide bonds, post-translation modifications)
Incorrect DNA template design: regulatory elements (promoter, operators, untranslated region), genes	Always verify key DNA template regions by sequencing
Poor control over the expression of DNA	Choose correct promoter with respect to CFE system Carefully screen stoichiometry of DNA (plasmid, linear)
<b>Workplace preparation and safety</b>	
Inappropriate or varying ambient environment	Bench cleaned every day; pipette cleaned or autoclaved, tip boxes closed, sterile tubes and double distilled H <sub>2</sub> O used
Contamination from dust or microbial growth	
Non-sterile or non-nuclease-free conditions	
<b>Solution and reagent preparation</b>	
Uncalibrated pipettes and pipetting errors	Calibrate pipettes every 6 months
Contaminated pipette tips	Use sterile, boxed filter pipette tips
Lysate from bacteria harvested at wrong optical density	Collect cells at mid-log phase for lysate preparation
Contaminated or old stock solutions	Always use freshly prepared stock solutions
Improper storage of components	Never use any unsure stock solution; re-prepare any questionable stock solution
Incorrect labelling of containers	
Incorrect stock solution concentrations	
<b>DNA template purification</b>	
Plasmid from overgrown cell culture	Collect cells at late exponential phase
Poor DNA quality	PCR clean up of the DNA stock
Incorrect DNA template concentration	Compare NanoDrop with assays using target-selective fluorescent dyes
Contamination of DNA template	Re-prepare DNA stock
DNA resuspended in wrong solution	
<b>Reaction assembly</b>	
Uncalibrated pipettes and pipetting errors	Calibrate pipettes every 6 months
Insufficient mixing of the CFE reaction	Vortex gently after adding each component
Air gaps or bubbles in well	Remove air with a tip, via centrifugation or gently tap the vessel
Wrong container volume or shape	Reaction to available air ratio should be ~100
Poor vessel sealing	Make sure the reaction vessels are tightly sealed
Degradation of linear DNA templates	Use GamS or Chi6 to inhibit DNA degradation
<b>Reaction</b>	
Inappropriate plate reader settings	Verify wavelength, time lapse, duration, temperature, plate type, top or bottom optics
Evaporation from poorly sealed plate	Make sure the reactions vessels are tightly sealed
<b>Measurement analysis</b>	
Saturated detector signal	Pre-test different device amplification settings (gain)
Too few measurements or repeats for statistical analysis/significance	Carefully choose the time lapse between each data point acquisition depending on expressed DNA; use statistical analysis and design of experiments to determine the number of replicates needed
Positional biases on well plates	Run the same sample across a plate and check for outliers
Wrong range of calibration curve	Avoid saturated signal, work in linear regime of calibration

CFE, cell-free gene expression.

may render biological variability an enabling feature of engineering with CFE systems. Ultimately, more reproducible and comparable CFE measurements will advance CFE capabilities, for example by bridging the gap between *in vitro* and *in vivo* expression in DBTL workflows<sup>128</sup>.

### Calibration

Calibration offers an immediately accessible way to improve measurement quality (Supplementary Fig. 5). Beyond simply aiding in troubleshooting and early detection of instrument problems, calibration provides the supporting data needed to quantify confidence in measurement results<sup>129</sup>. Routine instrument calibration should therefore be included throughout as an integral part of a CFE workflow. Incubation temperatures for cell growth and CFE could be calibrated using a traceable thermometer in an incubator, a temperature-sensitive dye in a multiwell plate or a calibration plate for a multimode microplate reader. A plate reader manufacturer's calibration for measurements of absorbance and fluorescence may include checks on the alignment of optical components, instrument sensitivity, linearity of the detector response for a known concentration of reference standard and co-planar orientation of the microwell plate to the carrier. Researchers may additionally measure calibration plates filled with water, fluorescent solution or absorbance solution, which can indicate bias across the microwell plate, for example, owing to differences in illumination, evaporation or membrane sealing. A calibration curve of a fluorescent reporter produced by the CFE system at known concentrations may also be included on the plate to convert between arbitrary fluorescence units and molar concentrations<sup>9,36</sup>.

### Standardization

Beyond measurement assurance through calibration, standards for measurements can aid in reproducibility and comparability. Recombinant fluorescent reporters — especially versions of green fluorescent protein such as eGFP — have emerged as *de facto* standards for assaying CFE performance, as have CFE systems based on *E. coli*. However, no formal standards exist for CFE, even as the field has issued informal calls for agreement in best practices, measurement assurance and standardization throughout the basic CFE workflow<sup>128</sup>. Researchers may benefit from guidance on best practices for reporting data and methods, similar to minimal information standards for cell culture<sup>130</sup>. A standard protocol for the preparation and use of a standard test plate to assess the composition and performance of a CFE reaction further stands out as particularly promising in the near term<sup>128</sup>. Some researchers advocate the selection of a few representative CFE platforms for comprehensive characterization, data repositories and modelling. Standardization may also facilitate the use of absolute measurements and traceability to the SI for CFE systems, further supporting reproducible and comparable CFE. Indeed, good measurement practices and measurement infrastructure promise to unlock real-world impact for CFE<sup>131</sup>.

## Limitations and optimizations

### Research optimization

In typical research and academic laboratory settings, access to hands-on training with skilled practitioners and quality reagents remains essential for rapid adoption and troubleshooting of CFE. The CFE community works actively to provide instruction, for example, through [courses at the Cold Spring Harbor Laboratory](#) and professional networking at [Build-A-Cell](#) workshops. Many laboratories prepare custom CFE reagents, but characterization can be costly and time-consuming, and the results may not even inform directly on CFE reliability and performance. Commercial kits provide a convenient, if costly, alternative. These could be expanded by manufacturers to include CFE systems based on more organisms, covering more steps in the workflow and with kit components optimized for CFE. New users would also benefit from more complete and accurate specification sheets and instructions, as well as the inclusion of information on benchmarking and troubleshooting.

Additional studies could systematically explore how various methods and materials affect the performance of CFE, beyond those published for select portions of the typical CFE workflow<sup>111</sup> (Supplementary Table 9). For example, development of CFE systems based on new organisms has often proceeded successfully using protocols for *E. coli*<sup>132–136</sup>. Studies examining how specific protocol steps relate to reaction components and performance would aid in optimizing workflows specific to new CFE systems. These studies might consider, for example, various methods for lysate preparation, cell growth conditions, optical density at cell harvest, reagent grade and contaminants, reaction headspace<sup>45</sup>, undissolved components<sup>125</sup>, DNA template preparation<sup>42–44</sup> and ambient humidity and temperature. The growing prevalence of automation for CFE also warrants further investigation and guidance<sup>63,111,137</sup> (BOX 1).

The need for further studies arises, in part, owing to gaps in fundamental biology relevant to CFE. It should be noted that an *E. coli* cell extract comprises up to 800 different proteins, many of which are metabolically active enzymes<sup>15,35</sup>. Given that CFE reactions take place against this complex metabolic background, a lack of clarity surrounding reaction composition, roles for each component and important parameters of a CFE reaction relevant to performance can hinder quantitative modelling<sup>138</sup>, prediction and optimization. The role of biological variability for CFE also remains an open question, especially in relation to other sources of measurement uncertainty.

### Application optimization

Despite the overwhelming popularity of CFE systems based on *E. coli*, these remain limited with respect to some applications. A number of the related technical challenges generalize to CFE systems more broadly, such as improving the longevity of CFE reactions and the efficiency of energy usage. There is also a need for new measurements and methods for dynamic control and optimization of resource utilization during CFE.

**Box 1 | Limitations on the use of CFE systems**

The most prohibitive limitations to using cell-free gene expression (CFE) are inadequate access to resources and knowledge, the need for better measurements and poor prediction of CFE performance.

Broader access to resources and knowledge is most straightforward to address. Despite successful commercial applications of CFE, research in CFE has identified a lack of reproducibility as a barrier to adoption and progress. This disconnect stems from considerable industry investment in optimizing CFE reactions and taking advantage of established best practices in, for example, industrial engineering and manufacturing. Meanwhile, academic researchers pursue projects with smaller scope, fewer resources and a wider variety of materials, techniques and CFE systems. Bridging this capability gap would benefit both communities. Potential paths forward include effective exchange of information (protocols, best practices and training programmes) and tools (community biofoundries, infrastructure for scaling up), and conversations across fields and applications in the form of workshops that include industry and academia.

Better measurements can both inform fundamental understanding and uncover sources of variability in reaction performance. Good measurement practices and new measurement tools may be directed, for example, towards determining the components of a CFE reaction and their function over time. Guidance and standards for reporting measurements, exchanging data and models, experimental design<sup>191–193</sup> and sharing full and detailed protocols will further ensure high-quality data.

Improving the prediction and design of CFE performance requires quantitative prediction of CFE reactions. Constructing good models for both simpler, reconstituted CFE systems and more complex, lysate-based systems would be broadly enabling. Modellers may improve prediction by leveraging better measurements, increased data availability and advances in understanding biomolecular mechanisms and molecular crowding. Quantitative prediction is important in bridging the gap between *in vitro* CFE systems and *in vivo* cell-based systems, approaches to biological control, scaling up to reaction volumes relevant to biomanufacturing and scaling out to CFE systems based on non-model organisms.

For example, analysis of energy consumption suggests that only ~0.6% of the theoretical energy available is expended for protein synthesis<sup>24</sup>. Secondary metabolites, such as acetate<sup>39</sup>, may accumulate and inhibit the reaction. As discussed previously, many protein products relevant to applications require additional steps, such as post-translational modification, beyond the capabilities of the biomolecular machinery in typical *E. coli* CFE systems. The utility of CFE to facilitate prototyping during the DBTL cycle and integration of genetic parts remains unclear, especially considering the challenge of translating and integrating these results into living cells. Scaling up CFE production to industrially relevant volumes may require investment in biofoundry infrastructure, such as fermenters, automated liquid handlers, DNA synthesis tools or downstream processing and purification.

**Outlook**

The potential in the near future of CFE as a major platform for engineering biology enabling biomedical and other applications is tremendous. Automation will yield ever larger data sets, improving our understanding and increasing the impact of CFE (BOX 2). This and additional improvements have the potential to expand the CFE user base, yield systems capable of processing large genetic circuits, produce chemically diverse biopolymers or create abiotic materials with biotic properties. In this section, we compare this outlook with the state of the art and identify opportunities to close technical gaps that need to be overcome to realize the full potential of CFE (FIG. 6).

**Lyophilized CFE**

Attempts to make CFE more user-friendly have largely revolved around lyophilization. For example, applying CFE reactions to paper discs and lyophilizing the resulting product yields paper-based CFE for biosensing, where reactions are initiated by simple rehydration as opposed to combining separate components, such as extract and DNA<sup>105,106</sup>. The raw materials are inexpensive, and the product can be shipped at ambient temperature for increased affordability. Lyophilized CFE was further used to synthesize various biopharmaceuticals<sup>6,140,141</sup>, suggesting this format for on-demand biomanufacturing. Notably, a portable suitcase was built containing all of the components necessary for current good manufacturing practice production of biopharmaceuticals, highlighting the potential for local and distributed production<sup>141</sup>.

However, several issues limit the application of lyophilized CFE, including variability between users<sup>111</sup> and the need to extend its shelf-life<sup>124</sup>. An improved understanding of critical variables in CFE assays and preparations, as well as methods to improve such variability, will allow more reliable lyophilized CFE applications<sup>77,125</sup>. To further reduce variability, there is also the potential to utilize feedback mechanisms within genetic circuits, allowing more homeostatic gene expression<sup>142</sup>. In the case of shelf-life for CFEs, the current 60-day time frame, although useful for certain research applications, may be too short for biopharmaceutical manufacturers, where supply chains and product shelf-life are key commercial considerations.

Attempts to improve shelf-life via the addition of protectants have yet to reach a suitable benchmark for existing biopharmaceutical manufacturing<sup>143</sup>. Further improvements to the shelf-life of lyophilized CFE are, in principle, possible, given that cellular machinery can persist in amber-encased endospores for millions of years<sup>144</sup>.

**Expressing large gene sets**

Expressing large gene sets exhibiting specific regulatory mechanisms and functions remains a major bottleneck for CFE systems because synthetic tailor-made genetic regulations are far from being as effective as those embedded in natural genetic programmes. Circuits composed of five to ten genes can be readily processed in CFE reactions, either in batch mode or in synthetic cells<sup>9,58,142,145–148</sup>. This format is suitable for expressing sections of biosynthetic pathways<sup>149,150</sup>. However, expression-regulated gene circuits composed of more than 10–20 genes are beyond the state of the art. To date, the only examples of large gene sets expressed by CFE are natural bacteriophage genomes, such as T7 (40 kbp, 60 genes)<sup>22</sup> and T4 (169 kbp, 289 genes)<sup>23</sup>. The ability to express large gene sets by CFE would enable a much wider range of applications. For example, clusters encoding large enzyme complexes for the complete synthesis of non-ribosomal peptides<sup>151</sup> or polyketides<sup>152</sup> may be within reach for CFE systems. Many of these compounds have antimicrobial and antitumour properties<sup>153</sup>, and thus the ability to screen for novel bioactive molecules using CFE would be attractive. Engineering minimal cells

**Lyophilization**

The process of water removal via freezing followed by a reduction in pressure and heating.

**Minimal cells**

Engineered cell-sized compartments using natural biomolecules that mimic one or many functions of a biological cell.

## Box 2 | Automating CFE

Automation is expected to play a major role in the next decade for all steps of the cell-free gene expression (CFE) workflow, from lysate preparation to CFE reaction dispensing and measurements to data analysis. Although some robotic solutions already exist, the challenge remains to integrate this equipment into an automated CFE pipeline to, for example, increase throughput and reproducibility and reduce variability and overall cost<sup>194</sup>. As for much research involving biological systems, current automation in CFE workflows typically involves workstations designed for specific tasks. A common example is a stand-alone plate reader for automated fluorescent measurements<sup>24</sup>. Integrating multiple automated tasks within a single workstation occurs more rarely<sup>75</sup>.

The acquisition of larger data sets enables machine learning<sup>14</sup> and statistical inference<sup>24,111</sup> contributing to modelling and predictability for furthering applications of CFE. Novel technologies, such as transfer of nanolitre-size liquid droplets by acoustic dispensing, are emerging as a pivotal alternative to tip-based liquid handling<sup>14,24,63,111,142</sup>. Acoustic dispensing enables the rapid assembly of complex CFE reactions of microlitre size directly onto microtitre plates<sup>24,63,142</sup>. This increases throughput, reduces per-reaction costs and extends limited volumes of batch reagents, such as cell lysates. Although access to acoustic handlers has been limited by their high initial and service costs<sup>195</sup>, the establishment of biofoundries<sup>196</sup> improves access to these and similarly costly automation resources. Access may further increase through cloud-based laboratories that aim to offer a pay per experiment model<sup>197</sup>. Nevertheless, automation is not without its own needs and challenges in calibration, integration and programming. Automation often requires additional and extensive protocol development and optimization, even for workflows already optimized for manual use at the bench.

Although the automation of CFE reaction assembly is present in some laboratories, machine-driven preparation of lysates is far from established. Coupling small culture volumes<sup>76</sup> to an automated extract preparation workflow could help to advance the functionality and understanding of CFE. Integrating CFE with automated oligonucleotide synthesis is another area for development. This was demonstrated recently by the automated assembly of open reading frames into linear expression cassettes, followed by CFE and protein purification on a single platform<sup>75</sup>. Similar efforts could be aided by the integration of enzymatic DNA synthesis platforms<sup>198</sup>, generating genetic diversity in-house and further shortening design–build–test–learn (DBTL) cycles.

is another emerging application that requires the expression of whole metabolic pathways, such as glycolysis or the tricarboxylic acid cycle, both requiring at least ten or more enzymes<sup>154</sup>.

A potential reference landmark for the future requirements for CFE is in bottom-up engineering of a self-supporting CFE-based minimal cell, which remains an outstanding goal of biotechnological research broadly. Such a cell would require the expression of a genome of 400–500 genes, as found in the most primitive bacterial cells<sup>155–157</sup>. The average cytoplasmic protein concentration in *E. coli* is approximately 1  $\mu\text{mol l}^{-1}$  (REF. <sup>158</sup>), which brings the level of protein production to 400–500  $\mu\text{mol l}^{-1}$ . Achieving these concentrations would require yields of 10–13  $\text{mg ml}^{-1}$  for a protein similar to eGFP. CFE protein synthesis of 6  $\text{mg ml}^{-1}$  has been realized with semi-continuous exchange systems<sup>9</sup>. Several strategies have been undertaken to achieve greater protein synthesis. One approach builds on devising alternative energy regeneration systems, for example, photosynthetic ATP regeneration based on physical inputs could be pivotal to achieve greater cell-free protein synthesis<sup>159</sup>. A second approach is based on the development of CFE systems from organisms other than *E. coli* to determine whether greater protein synthesis yields can be achieved<sup>132,136,160</sup>, although these have so far been unsuccessful. The development of more accurate CFE metabolism models is expected to play a critical role in obtaining more powerful CFE systems<sup>62,139,161</sup>.

#### Semi-continuous exchange systems

Systems in which a cell-free gene expression (CFE) reaction is fed with nutrients such as ribonucleosides and amino acids to prolong the CFE reaction, through a dialysis membrane with a 10-kDa molecular mass cut-off.

**Non-canonical amino acids (ncAAs).** Amino acids not found in nature. Examples include *p*-acetyl-L-phenylalanine and canavanine.

**Amber codon suppression** Incorporating an amino acid at a UAG (amber) codon, rather than terminating translation.

**Microsomes** Artificial vesicle-like artefacts derived from pieces of the endoplasmic reticulum when eukaryotic cells are broken up in the laboratory.

#### Increasing biopolymer diversity

Non-canonical amino acids (ncAAs) were incorporated into proteins more than 30 years ago using CFE<sup>162</sup>. Various improvements to this approach have been made in recent years<sup>163,164</sup>, especially incorporating a single ncAA at multiple sites with high efficiency, which is feasible in CFE using methods based on amber codon suppression<sup>165</sup>. It is now possible to incorporate a single ncAA at 40 different loci with an efficiency of 98%<sup>26</sup>. The incorporation of ncAAs into proteins by CFE has found diverse applications in protein structure<sup>79,90</sup>, introduction of post-translational modifications<sup>80,166</sup> and antibody–drug conjugates<sup>163</sup>, in addition to other therapeutics<sup>167</sup>. This strategy cannot accommodate additional codons, however, preventing different types of ncAA at bespoke positions. New methods incorporating different ncAAs at multiple bespoke sites with high efficiency would increase the diversity of proteins and biopolymers synthesized via CFE, with potentially high impact in medical applications. The improved ability to reconstitute functional ribosomes in vitro may aid this objective<sup>168,169</sup>.

Protein glycosylation is another area with promising potential to expand the diversity of proteins produced by CFE. Future applications include the production of glycosylated biopharmaceuticals and pathogen components, such as the heavily glycosylated SARS-CoV-2 spike protein<sup>170</sup>. Native *E. coli* CFE systems lack the machinery for protein glycosylation<sup>8</sup>. One strategy for glycoprotein production uses a lysate from a eukaryotic organism that contains<sup>171</sup> or is enriched with<sup>172</sup> microsomes. Disadvantages of this method include both high cost and lower productivity. An alternative strategy uses glycosyltransferases compatible with bacterial CFE systems<sup>173</sup>. However, the efficiency of this method can be as low as 40%<sup>174</sup>, restricting designs. The next 10 years will see major improvements in these CFE systems.

#### Interfaced CFE

Several novel protein synthesis dynamics and operatory modes have been observed by combining CFE with fluidics. Examples include steady-state protein synthesis, as opposed to batch reactions that run to equilibrium<sup>175</sup>, and control of gene expression via electrical fields applied to reaction chambers<sup>28</sup>. Encapsulation of CFE in droplets<sup>176</sup>, hydrogels<sup>177,178</sup> and liposomes<sup>179</sup> has also yielded new applications. Notable recent examples include light-fuelled CFE encapsulated in liposomes<sup>159</sup>, ultra-high-throughput enzyme engineering in droplets<sup>180</sup> and miniaturized hydrogel structures that maintain steady-state protein synthesis for weeks in microcompartments<sup>178</sup>. A recent review offers a more detailed discussion on this subject<sup>181</sup>.

Several groups are focusing on developing self-replicating systems from biochemical components<sup>168,182–184</sup>, suggesting this self-replication behaviour could be realized in the coming decade. Challenges include complete in vitro ribosome biogenesis and the replication of DNA encoding all components of the CFE system. Although functional 30S ribosomal subunits were recently assembled from CFE-synthesized components<sup>185</sup>, to the best of our knowledge the complete

State of the art  2020	Gaps 	10-year outlook  2030
<b>Automated CFE</b> <ul style="list-style-type: none"> <li>• Human-based CFE preparation</li> <li>• No reproducibility of results from laboratory to laboratory</li> <li>• No CFE materials/data repository</li> </ul>	<ul style="list-style-type: none"> <li>• Automation of the CFE workflow (preparation of lysate, nutrient and cofactor solutions, CFE reaction dispensing and analysis)</li> <li>• CFE central data repository</li> <li>• CFE-specific DNA repository</li> </ul>	<ul style="list-style-type: none"> <li>• Fully automated CFE workflow</li> <li>• Laboratory to laboratory reproducibility</li> <li>• Rapid preparation of CFE systems from a wide variety of cell types</li> </ul>
<b>User-friendly CFE</b> <ul style="list-style-type: none"> <li>• CFE limited to laboratory environments</li> <li>• Variability in CFE performance</li> <li>• Lyophilized CFE systems</li> <li>• CFE-based educational kits</li> </ul>	<ul style="list-style-type: none"> <li>• Year-long shelf-life without the need of a cold chain</li> <li>• User-independent performance (synthesis titre and dynamics)</li> <li>• Time-relevant results (sensing)</li> <li>• Affordable and economical</li> </ul>	<ul style="list-style-type: none"> <li>• Expanded CFE market</li> <li>• At-home and remote CFE (on-tap therapeutics, biomanufacturing in space and remote locations)</li> </ul>
<b>Large genetic sets</b> <ul style="list-style-type: none"> <li>• Circuit parts (CRISPR) and motifs (oscillators, PID controllers)</li> <li>• Simple metabolic pathways</li> <li>• Virus-like particles, whole phages</li> </ul>	<ul style="list-style-type: none"> <li>• Novel, powerful and long-lived ATP regeneration systems</li> <li>• Extracts from microorganisms other than the standard ones</li> <li>• Improved predictability of CFE performance</li> </ul>	<ul style="list-style-type: none"> <li>• Metabolism cycles (Krebs, TCA)</li> <li>• Biosynthetic clusters (non-ribosomal peptides, polyketides)</li> <li>• Minimal cell genome</li> </ul>
<b>Diverse biopolymers</b> <ul style="list-style-type: none"> <li>• Antibody–drug conjugates, AMPs</li> <li>• Bacterial membrane proteins</li> <li>• 1–2 ncAA proteins</li> <li>• Glycosylated proteins (vaccines)</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple genetic code expansions</li> <li>• Humanized glycoproteins that are high-yielding and economical</li> <li>• Engineered ribosomes capable of polymerizing new ncAAs or alternative monomers</li> <li>• Membrane protein insertases</li> </ul>	<ul style="list-style-type: none"> <li>• &gt;10 ncAA proteins, complex therapeutics, new enzyme catalysts</li> <li>• Eukaryotic membrane proteins</li> <li>• Human and pathogen glycoproteins</li> </ul>
<b>Interfaced CFE</b> <ul style="list-style-type: none"> <li>• Liposomes, polymersomes, peptidosomes</li> <li>• CFE on paper, in hydrogels</li> <li>• Microfluidic chips</li> </ul>	<ul style="list-style-type: none"> <li>• Interface CFE with natural, synthetic, hard or soft materials</li> <li>• Characterize CFE biotic–abiotic interfaces</li> <li>• Create physically and chemically robust hybrid CFE materials</li> </ul>	<ul style="list-style-type: none"> <li>• CFE-based smart materials and diagnostics (environment, health)</li> <li>• Self-reproducing synthetic cell based on CFE and minimal genome</li> </ul>

Fig. 6 | **Outlook for CFE.** Each of the five categories highlights opportunities to bridge a technical gap and advance capabilities for cell-free gene expression (CFE). Left column: current state of the art of CFE, including some missing capabilities. Middle column: some of the possible improvements that would expand and improve technology in the short term. Right column: outlook for CFE in 10 years' time. AMP, antimicrobial peptide; ncAA, non-canonical amino acid; PID, proportional-integral-derivative; TCA, tricarboxylic acid.

synthesis and assembly of 50S subunits is beyond current capabilities. Regarding DNA replication, only single DNA species encoding several enzymes have been replicated in individual reactions<sup>183,186</sup>. This must be extended to replicate DNA encoding for more complicated systems containing ribosomes, for instance. Beyond self-replication, other capabilities characteristic of 'smart materials' might be realized using CFE, such as self-healing<sup>187</sup>. Potential routes to this function were demonstrated by the encapsulation of CFE in amphiphilic protein membranes, where monomers were further synthesized by the encased CFE system<sup>188,189</sup>.

### Conclusion

Although much remains to be accomplished to optimize CFE, past and ongoing efforts have already transformed CFE into an accessible technology. Streamlining preparation and use of CFE will remain a major focus

in the next decade, with considerable improvements anticipated in the near future. The development of CFE systems from non-model organisms, the integration of CFE into synthetic materials and the contribution of CFE to synthetic cells all promise to unlock a wealth of unforeseen bioengineering capabilities. As importantly, the CFE community is developing quickly, along with the emergence of many shared resources (Supplementary Table 9). Annual conferences specific to CFE now offer new users excellent opportunities to interact with CFE researchers, gain exposure to the state of the art and see first-hand the versatility of CFE. CFE offers a compelling platform technology for investigating deep questions in biology and impacting nearly every aspect of our daily lives; the discussion here is meant as an open, empowering invitation to those who would take up CFE and help realize this potential.

Published online: 15 July 2021

1. Hoagland, M. B., Stephenson, M. L., Scott, J. F., Hecht, L. I. & Zamecnik, P. C. A soluble ribonucleic acid intermediate in protein synthesis. *J. Biol. Chem.* **231**, 241–257 (1958).
2. Singer, M. F. 1968 Nobel Laureate in Medicine or Physiology. *Science* **162**, 433–436 (1968).
3. Nirenberg, M. Historical review: Deciphering the genetic code — a personal account. *Trends Biochem. Sci.* **29**, 46–54 (2004).
4. Zubay, G. In vitro synthesis of protein in microbial systems. *Ann. Rev. Genet.* **7**, 267–287 (1973).
5. Garenne, D. & Noireaux, V. Cell-free transcription–translation: engineering biology from the nanometer to the millimeter scale. *Curr. Opin. Biotechnol.* **58**, 19–27 (2019).
6. Pardee, K. et al. Portable, on-demand biomolecular manufacturing. *Cell* **167**, 248–259 (2016).
7. Silverman, A. D., Karim, A. S. & Jewett, M. C. Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* **21**, 151–170 (2020).
8. Gregorio, N. E., Levine, M. Z. & Oza, J. P. A user's guide to cell-free protein synthesis. *Methods Protoc.* **2**, 24 (2019).
9. Garamella, J., Marshall, R., Rustad, M. & Noireaux, V. The all *E. coli* TX-TL Toolbox 2.0: a platform for cell-free synthetic biology. *ACS Synth. Biol.* **5**, 344–355 (2016).
10. Kai, L. & Schwille, P. Cell-free protein synthesis and its perspectives for assembling cells from the bottom-up. *Adv. Biosyst.* **3**, 1800522 (2019).
11. Tayar, A. M., Daube, S. S. & Bar-Ziv, R. H. Progress in programming spatiotemporal patterns and machine-assembly in cell-free protein expression systems. *Curr. Opin. Chem. Biol.* **40**, 37–46 (2017).
12. Greiss, F., Daube, S. S., Noireaux, V. & Bar-Ziv, R. From deterministic to fuzzy decision-making in artificial cells. *Nat. Commun.* **11**, 5648 (2020).
13. Dopp, B. J. L., Tamiev, D. D. & Reuel, N. F. Cell-free supplement mixtures: elucidating the history and biochemical utility of additives used to support in vitro protein synthesis in *E. coli* extract. *Biotechnol. Adv.* **37**, 246–258 (2019).
14. Borkowski, O. et al. Large scale active-learning-guided exploration for in vitro protein production optimization. *Nat. Commun.* **11**, 1872 (2020).
15. Garenne, D., Beisel, C. L. & Noireaux, V. Characterization of the all-*E. coli* transcription–translation system myTXTL by mass spectrometry. *Rapid Commun. Mass Spectrom.* **33**, 1036–1048 (2019).
16. Niederholtmeyer, H. et al. The *E. coli* S30 lysate proteome: a prototype for cell-free protein production. *Science* **89**, 3473–3480 (2018).
17. Chappell, J., Jensen, K. & Freemont, P. S. Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res.* **41**, 3471–3481 (2013).
18. Marshall, R. et al. Rapid and scalable characterization of CRISPR technologies using an *E. coli* cell-free transcription–translation system. *Mol. Cell* **69**, 146–157 (2018).
19. Sun, Z. Z. et al. Protocols for implementing an *Escherichia coli* based TX-TL cell-free expression system for synthetic biology. *J. Vis. Exp.* <https://doi.org/10.3791/50762> (2013).
20. Shin, J. & Noireaux, V. An *E. coli* cell-free expression toolbox: application to synthetic gene circuits and artificial cells. *ACS Synth. Biol.* **1**, 29–41 (2012).
21. Shin, J. & Noireaux, V. Efficient cell-free expression with the endogenous *E. coli* RNA polymerase and sigma factor 70. *J. Biol. Eng.* **4**, 8 (2010).
22. Shin, J., Jardine, P. & Noireaux, V. Genome replication, synthesis, and assembly of the bacteriophage T7 in a single cell-free reaction. *ACS Synth. Biol.* **1**, 408–413 (2012).
23. Rustad, M., Eastlund, A., Jardine, P. & Noireaux, V. Cell-free TXTL synthesis of infectious bacteriophage T4 in a single test tube reaction. *Synth. Biol.* **3**, ysy002 (2018).
24. Moore, S. J. et al. Rapid acquisition and model-based analysis of cell-free transcription–translation reactions from nonmodel bacteria. *Proc. Natl Acad. Sci. USA* **115**, E4340–E4349 (2018).
25. Dudley, Q. M., Karim, A. S., Nash, C. J. & Jewett, M. C. In vitro prototyping of limonene biosynthesis using cell-free protein synthesis. *Metab. Eng.* **61**, 251–260 (2020).
26. Martin, R. W. et al. Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat. Commun.* **9**, 1203 (2018).
27. Chemla, Y., Ozer, E., Schlesinger, O., Noireaux, V. & Alfanta, L. Genetically expanded cell-free protein synthesis using endogenous pyrrolysyl orthogonal translation system. *Biotechnol. Bioeng.* **112**, 1663–1672 (2015).
28. Efrat, Y., Tayar, A. M., Daube, S. S., Levy, M. & Bar-Ziv, R. H. Electric-field manipulation of a compartmentalized cell-free gene expression reaction. *ACS Synth. Biol.* **7**, 1829–1833 (2018).
29. Kamat, N. P., Katz, J. S. & Hammer, D. A. Engineering polymersome protocells. *J. Phys. Chem. Lett.* **2**, 1612–1623 (2011).
30. Nallani, M. et al. Proteopolymerosomes: in vitro production of a membrane protein in polymersome membranes. *Biointerphases* **6**, 153–157 (2011).
31. Cole, S. D., Miklos, A. E., Chiao, A. C., Sun, Z. Z. & Lux, M. W. Methodologies for preparation of prokaryotic extracts for cell-free expression systems. *Synth. Syst. Biotechnol.* **5**, 252–267 (2020).
32. Blattner, F. R. et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462 (1997).
33. Shimizu, Y. et al. Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**, 751–755 (2001).
34. Zemella, A., Thoring, E., Hoffmeister, C. & Uebick, S. Cell-free protein synthesis: pros and cons of prokaryotic and eukaryotic systems. *Chembiochem* **16**, 2420–2431 (2015).
35. Foshag, D. et al. The *E. coli* S30 lysate proteome: a prototype for cell-free protein production. *N. Biotechnol.* **40**, 245–260 (2018).
36. Levine, M. Z., Gregorio, N. E., Jewett, M. C., Watts, K. R. & Oza, J. P. *Escherichia coli*-based cell-free protein synthesis: protocols for a robust, flexible, and accessible platform technology. *J. Vis. Exp.* <https://doi.org/10.3791/58882> (2019).
37. Zawada, J. & Swartz, J. Effects of growth rate on cell extract performance in cell-free protein synthesis. *Biotechnol. Bioeng.* **94**, 618–624 (2006).
38. Kim, R. G. & Choi, C. Y. Expression-independent consumption of substrates in cell-free expression system from *Escherichia coli*. *J. Biotechnol.* **84**, 27–32 (2000).
39. Silverman, A. D., Kelley-Loughnane, N., Lucks, J. B. & Jewett, M. C. Deconstructing cell-free extract preparation for in vitro activation of transcriptional genetic circuitry. *ACS Synth. Biol.* **8**, 403–414 (2019).
40. Cayley, S., Lewis, B. A., Guttman, H. J. & Record, M. T. Characterization of the cytoplasm of *Escherichia coli* K-12 as a function of external osmolarity. Implications for protein-DNA interactions in vivo. *J. Mol. Biol.* **222**, 281–300 (1991).
41. Kai, L., Dötsch, V., Kaldenhoff, R. & Bernhard, F. Artificial environments for the co-translational stabilization of cell-free expressed proteins. *PLoS ONE* **8**, e56637 (2013).
42. 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 Synth. Biol.* **3**, 387–397 (2014).
43. Marshall, R., Maxwell, C. S., Collins, S. P., Beisel, C. L. & Noireaux, V. Short DNA containing  $\chi$  sites enhances DNA stability and gene expression in *E. coli* cell-free transcription–translation systems. *Biotechnol. Bioeng.* **114**, 2137–2141 (2017).
44. Yim, S. S., Johns, N. I., Noireaux, V. & Wang, H. H. Protecting linear DNA templates in cell-free expression systems from diverse bacteria. *ACS Synth. Biol.* **9**, 2851–2855 (2020).
45. Voloshin, A. M. & Swartz, J. R. Efficient and scalable method for scaling up cell free protein synthesis in batch mode. *Biotechnol. Bioeng.* **91**, 516–521 (2005).
46. Vernon, W. B. The role of magnesium in nucleic-acid and protein metabolism. *Magnesium* **7**, 234–248 (1988).
47. Pédelacq, 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 (2006).
48. Vincent, S. G., Cunningham, P. R., Stephens, N. L., Halayko, A. J. & Fisher, J. T. Quantitative densitometry of proteins stained with coomassie blue using a Hewlett Packard Scanjet scanner and Scanplot software. *Electrophoresis* **18**, 67–71 (1997).
49. Takemori, N. et al. Top-down/bottom-up mass spectrometry workflow using dissolvable polyacrylamide gels. *Anal. Chem.* **89**, 8244–8250 (2017).
50. Richardson, D., Itkonen, J., Nievas, J., Urtti, A. & Casteleijn, M. G. Accelerated pharmaceutical protein development with integrated cell free expression, purification, and bioconjugation. *Sci. Rep.* **8**, 11967 (2018).
51. Basu, D., Castellano, J. M., Thomas, N. & Mishra, R. K. Cell-free protein synthesis and purification of human dopamine D2 receptor long isoform. *Biotechnol. Prog.* **29**, 601–608 (2013).
52. Laml, T., Stiege, W. & Erdmann, V. A. An improved protein bioreactor: efficient product isolation during in vitro protein biosynthesis via affinity tag. *Mol. Cell. Proteomics* **1**, 466–471 (2002).
53. Bechllars, S. et al. Cell-free synthesis of functional thermostable direct hemolysins of *Vibrio parahaemolyticus*. *Toxicon* **76**, 132–142 (2013).
54. Alimov, A. P., Khmelnit-Sky, A. Y., Simonenko, P. N., Spirin, A. S. & Chetverin, A. B. Cell-free synthesis and affinity isolation of proteins on a nanomole scale. *Biotechniques* **28**, 338–344 (2000).
55. Mü Ller-Lucks, A., Bock, S., Wu, B. & Beitz, E. Fluorescent In situ folding control for rapid optimization of cell-free membrane protein synthesis. *PLoS ONE* **7**, e42186 (2012).
56. Takahashi, M. K. et al. Characterizing and prototyping genetic networks with cell-free transcription–translation reactions. *Methods* **86**, 60–72 (2015).
57. Niederholtmeyer, H. et al. Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife* **4**, e09771 (2015).
58. Karzbrun, E., Tayar, A. M., Noireaux, V. & Bar-Ziv, R. H. Programmable on-chip DNA compartments as artificial cells. *Science* **345**, 829–832 (2014).
59. Matveev, S. V. et al. Effect of the ATP level on the overall protein biosynthesis rate in a wheat germ cell-free system. *Biochim. Biophys. Acta* **1293**, 207–212 (1996).
60. 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 (2014).
61. Balleza, E., Kim, J. M. & Cluzel, P. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nat. Methods* **15**, 47–51 (2018).
62. Vilkhovoy, M. et al. Sequence specific modeling of *E. coli* cell-free protein synthesis. *ACS Synth. Biol.* **7**, 1844–1857 (2018).
63. Marshall, R. & Noireaux, V. Quantitative modeling of transcription and translation of an all-*E. coli* cell-free system. *Sci. Rep.* **9**, 11980 (2019).
64. Matsuura, T., Tanimura, N., Hosoda, K., Yomo, T. & Shimizu, Y. Reaction dynamics analysis of a reconstituted *Escherichia coli* protein translation system by computational modeling. *Proc. Natl Acad. Sci. USA* **114**, E1336–E1344 (2017).
65. Matsuura, T., Hosoda, K. & Shimizu, Y. Robustness of a reconstituted *Escherichia coli* protein translation system analyzed by computational modeling. *ACS Synth. Biol.* **7**, 1964–1972 (2018).
66. Doerr, A. et al. Modelling cell-free RNA and protein synthesis with minimal systems. *Phys. Biol.* **16**, 025001 (2019).
67. Chizzolini, F., Forlin, M., Cecchi, D. & Mansy, S. S. Gene position more strongly influences cell-free protein expression from operons than T7 transcriptional promoter strength. *ACS Synth. Biol.* **3**, 363–371 (2014).
68. Paige, J. S., Wu, K. Y. & Jaffrey, S. R. RNA mimics of green fluorescent protein. *Science* **333**, 642–646 (2011).
69. Weise, L. I., Heymann, M., Mayr, V. & Mutschler, H. Cell-free expression of RNA encoded genes using MS2 replicase. *Nucleic Acids Res.* **47**, 10956–10967 (2019).
70. Filonov, G. S., Moon, J. D., Svendsen, N. & Jaffrey, S. R. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.* **136**, 16299–16308 (2014).
71. Kolpashchikov, D. M. Binary malachite green aptamer for fluorescent detection of nucleic acids. *J. Am. Chem. Soc.* **127**, 12442–12443 (2005).
72. Patterson, G. H., Knobel, S. M., Sharif, W. D., Kain, S. R. & Piston, D. W. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys. J.* **73**, 2782–2790 (1997).
73. Nagaraj, V. H., Greene, J. M., Sengupta, A. M. & Sontag, E. D. Translation inhibition and resource balance in the TX-TL cell-free gene expression system. *Synth. Biol.* **2**, yxs005 (2017).
74. Woodrow, K. A., Airen, I. O. & Swartz, J. R. Rapid expression of functional genomic libraries. *J. Proteome Res.* **5**, 3288–3300 (2006).

75. Chiocchini, C. et al. From electronic sequence to purified protein using automated gene synthesis and in vitro transcription/translation. *ACS Synth. Biol.* **9**, 1714–1724 (2020).
76. Kwon, Y. C. & Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* **5**, 8663 (2015).
77. Zawada, J. F. et al. Microscale to manufacturing scale-up of cell-free cytokine production — a new approach for shortening protein production development timelines. *Biotechnol. Bioeng.* **108**, 1570–1578 (2011).
78. Tuckey, C., Asahara, H., Zhou, Y. & Chong, S. Protein synthesis using a reconstituted cell-free system. *Curr. Protoc. Mol. Biol.* **108**, 16.31.1–16.31.22 (2014).
79. Kigawa, T. et al. Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett.* **442**, 15–19 (1999).
80. Park, H. S. et al. Expanding the genetic code of *Escherichia coli* with phosphoserine. *Science* **333**, 1151–1154 (2011).
81. Villemagne, D., Jackson, R. & Douthwaite, J. A. Highly efficient ribosome display selection by use of purified components for in vitro translation. *J. Immunol. Methods* **313**, 140–148 (2006).
82. Miroux, B. & Walker, J. E. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289–298 (1996).
83. Katzen, F., Chang, G. & Kudlicki, W. The past, present and future of cell-free protein synthesis. *Trends Biotechnol.* **23**, 150–156 (2005).
84. Savage, D. F., Anderson, C. L., Robles-Colmenares, Y., Newby, Z. E. & Stroud, R. M. Cell-free complements in vivo expression of the *E. coli* membrane proteome. *Protein Sci.* **16**, 966–976 (2007).
85. Schwarz, D., Daley, D., Beckhaus, T., Dötsch, V. & Bernhard, F. Cell-free expression profiling of *E. coli* inner membrane proteins. *Proteomics* **10**, 1762–1779 (2010).
86. Yang, J. P., Cirico, T., Katzen, F., Peterson, T. C. & Kudlicki, W. Cell-free synthesis of a functional G protein-coupled receptor complexed with nanometer scale bilayer discs. *BMC Biotechnol.* **11**, 57 (2011).
87. Smith, M. T., Varner, C. T., Bush, D. B. & Bundy, B. C. The incorporation of the A2 protein to produce novel Q $\beta$  virus-like particles using cell-free protein synthesis. *Biotechnol. Prog.* **28**, 549–555 (2012).
88. Salehi, A. S. M. 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 (2016).
89. Ozawa, K., Dixon, N. E. & Otting, G. Critical review cell-free synthesis of <sup>15</sup>N-labeled proteins for NMR studies. *IUBMB Life* **57**, 615–622 (2005).
90. Takeda, M. & Kainosho, M. Cell-free protein production for NMR studies. *Methods Mol. Biol.* **831**, 71–84 (2012).
91. Huang, A. et al. BioBits™ Explorer: a modular synthetic biology education kit. *Sci. Adv.* **4**, eaat5105 (2018).
92. Stark, J. C. et al. BioBits™ Bright: a fluorescent synthetic biology education kit. *Sci. Adv.* **4**, eaat5107 (2018).
93. Giordano-Attianese, G. et al. A computationally designed chimeric antigen receptor provides a small-molecule safety switch for T-cell therapy. *Nat. Biotechnol.* **38**, 426–432 (2018).
94. Huffman, M. A. et al. Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **366**, 1255–1259 (2019).
95. Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science* **349**, 1095–1100 (2015).
96. Voigt, C. A. Synthetic biology 2020–2030: six commercially-available products that are changing our world. *Nat. Commun.* **11**, 10–15 (2020).
97. Meng, F. & Ellis, T. The second decade of synthetic biology: 2010–2020. *Nat. Commun.* **11**, 5174 (2020).
98. Ledford, H. Quest to use CRISPR against disease gains ground. *Nature* <https://doi.org/10.1038/d41586-019-03919-0> (2020).
99. Boyle, E. A. et al. High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. *Proc. Natl Acad. Sci. USA* **114**, 5461–5466 (2017).
100. Tsai, S. Q. et al. Circle-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets. *Nat. Methods.* **14**, 607 (2017).
101. Tsai, S. Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR–Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015).
102. Maxwell, C. S., Jacobsen, T., Marshall, R., Noireaux, V. & Beisel, C. L. A detailed cell-free transcription–translation-based assay to decipher CRISPR protospacer-adjacent motifs. *Methods* **143**, 48–57 (2018).
103. Watters, K. E., Fellmann, C., Bai, H. B., Ren, S. M. & Doudna, J. A. Systematic discovery of natural CRISPR–Cas12a inhibitors. *Science* **362**, 236–239 (2018).
104. Kobayashi, H. et al. Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl Acad. Sci. USA* **101**, 8414–8419 (2004).
105. Pardee, K. et al. Paper-based synthetic gene networks. *Cell* **159**, 940–954 (2014).
106. Duyen, T. T. M. et al. Paper-based colorimetric biosensor for antibiotics inhibiting bacterial protein synthesis. *J. Biosci. Bioeng.* **123**, 96–100 (2017).
107. Salehi, A. S. M. et al. Biosensing estrogenic endocrine disruptors in human blood and urine: a RAPID cell-free protein synthesis approach. *Toxicol. Appl. Pharmacol.* **345**, 19–25 (2018).
108. Silverman, A. D., Akova, U., Alam, K. K., Jewett, M. C. & Lucks, J. B. Design and optimization of a cell-free atrazine biosensor. *ACS Synth. Biol.* **9**, 671–677 (2020).
109. Wen, K. Y. et al. A cell-free biosensor for detecting quorum sensing molecules in *P. aeruginosa*-infected respiratory samples. *ACS Synth. Biol.* **6**, 2293–2301 (2017).
110. Struss, A. K. et al. Toward implementation of quorum sensing autoinducers as biomarkers for infectious disease states. *Anal. Chem.* **85**, 3355–3362 (2013).
111. Cole, S. D. et al. Quantification of interlaboratory cell-free protein synthesis variability. *ACS Synth. Biol.* **8**, 2080–2091 (2019).
112. Walsh, G. Biopharmaceutical benchmarks 2018. *Nat. Biotechnol.* **36**, 1136–1145 (2018).
113. Martemyanov, K. A., Shirokov, V. A., Kurnasov, O. V., Gudkov, A. T. & Spirin, A. S. Cell-free production of biologically active polypeptides: application to the synthesis of antibacterial peptide cecropin. *Protein Expr. Purif.* **21**, 456–461 (2001).
114. Lu, Y., Welsh, J. P. & Swartz, J. R. Production and stabilization of the trimeric influenza hemagglutinin stem domain for potentially broadly protective influenza vaccines. *Proc. Natl Acad. Sci. USA* **111**, 125–130 (2014).
115. Yin, G. et al. Aglycosylated antibodies and antibody fragments produced in a scalable in vitro transcription–translation system. *mAbs* **4**, 217–225 (2012).
116. De Marco, A. Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microbial Cell Factories* **8**, 26 (2009).
117. Knapp, K. G. & Swartz, J. R. Evidence for an additional disulfide reduction pathway in *Escherichia coli*. *J. Biosci. Bioeng.* **103**, 373–376 (2007).
118. Yin, G. & Swartz, J. R. Enhancing multiple disulfide bonded protein folding in a cell-free system. *Biotechnol. Bioeng.* **86**, 188–195 (2004).
119. Kim, D. M. & Swartz, J. R. Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of *Escherichia coli*. *Biotechnol. Bioeng.* **85**, 122–129 (2004).
120. Goerke, A. R. & Swartz, J. R. Development of cell-free protein synthesis platforms for disulfide bonded proteins. *Biotechnol. Bioeng.* **99**, 351–367 (2008).
121. Bundy, B. C. & Swartz, J. R. Efficient disulfide bond formation in virus-like particles. *J. Biotechnol.* **154**, 230–239 (2011).
122. Li, F., Vijayansankaran, N., Shen, A., Kiss, R. & Amanullah, A. Cell culture processes for monoclonal antibody production. *mAbs* **2**, 466–479 (2010).
123. Arnold, C. Who shrank the drug factory? Briefcase-sized labs could transform medicine. *Nature* **575**, 274–277 (2019).
124. 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 (2014).
125. Dopp, J. L., Jo, Y. R. & Reuel, N. F. Methods to reduce variability in *E. coli*-based cell-free protein expression experiments. *Synth. Syst. Biotechnol.* **4**, 204–211 (2019).
126. 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.* **4**, 503–515 (2015).
127. Chizzolini, F. et al. Cell-free translation is more variable than transcription. *ACS Synth. Biol.* **6**, 638–647 (2017).
128. Romantseva, E. & Strychalski, E. A. CELL-FREE (Comparable Engineered Living Lysates For Research Education and Entrepreneurship) Workshop Report. *NIST Special Publication* 1–19 (NIST, 2019).
129. Plant, A. L. et al. How measurement science can improve confidence in research results. *PLoS Biol.* **16**, e2004299 (2018).
130. Hecht, A., Filliben, J., Munro, S. A. & Salit, M. A minimum information standard for reproducing bench-scale bacterial cell growth and productivity. *Commun. Biol.* **1**, 219 (2018).
131. McNerney, M. P. et al. Point-of-care biomarker quantification enabled by sample-specific calibration. *Sci. Adv.* **5**, eaax4473 (2019).
132. Kelwick, R., Webb, A. J., MacDonald, J. T. & Freemont, P. S. Development of a *Bacillus subtilis* cell-free transcription–translation system for prototyping regulatory elements. *Metab. Eng.* **38**, 370–381 (2016).
133. Li, J., Wang, H., Kwon, Y. C. & Jewett, M. C. Establishing a high yielding *Streptomyces*-based cell-free protein synthesis system. *Biotechnol. Bioeng.* **114**, 1343–1353 (2017).
134. Martin, R. W. et al. Development of a CHO-based cell-free platform for synthesis of active monoclonal antibodies. *ACS Synth. Biol.* **6**, 1370–1379 (2017).
135. Arumugam, T. U. et al. Application of wheat germ cell-free protein expression system for novel malaria vaccine candidate discovery. *Expert Rev. Vaccines* **13**, 75–85 (2014).
136. Failmezger, J., Scholz, S., Blombach, B. & Siemann-Herzberg, M. Cell-free protein synthesis from fast-growing *Vibrio natriegens*. *Front. Microbiol.* **9**, 1146 (2018).
137. Marshall, R., Garamella, J., Noireaux, V. & Pierson, A. High-throughput microliter-sized cell-free transcription–translation reactions for synthetic biology applications using the Echo 550 liquid handler. *Labcyte Appl. Note App-G124* (Labcyte, 2018).
138. Siegal-Gaskins, D., Noireaux, V. & Murray, R. M. Biomolecular resource utilization in elementary cell-free gene circuits. *Am. Control Conf.* **2013**, 1531–1536 (2013).
139. Horvath, N. et al. Toward a genome scale sequence specific dynamic model of cell-free protein synthesis in *Escherichia coli*. *Metab. Eng. Commun.* **10**, e00113 (2020).
140. Wilding, K. M. et al. Endotoxin-free *E. coli*-based cell-free protein synthesis: pre-expression endotoxin removal approaches for on-demand cancer therapeutic production. *Biotechnol. J.* **14**, 1800271 (2019).
141. Adiga, R. et al. Point-of-care production of therapeutic proteins of good-manufacturing-practice quality. *Nat. Biomed. Eng.* **2**, 675–686 (2018).
142. Agrawal, D. K., Marshall, R., Noireaux, V. & Sontag, E. D. In vitro implementation of robust gene regulation in a synthetic biomolecular integral controller. *Nat. Commun.* **10**, 5760 (2019).
143. Wilding, K. M., Zhao, E. L., Earl, C. C. & Bundy, B. C. Thermally stable lyophilized enhanced cell-free protein synthesis for on-demand endotoxin-free therapeutic production. *N. Biotechnol.* **53**, 73–80 (2019).
144. Cano, R. J. & Borucki, M. K. Revival and identification of bacterial spores in 25- to 40-million-year-old Dominican amber. *Science* **268**, 1060–1064 (1995).
145. Tayar, A. M., Karzbrun, E., Noireaux, V. & Bar-Ziv, R. H. Synchrony and pattern formation of coupled genetic oscillators on a chip of artificial cells. *Proc. Natl Acad. Sci. USA* **114**, 11609–11614 (2017).
146. Westbrook, A. et al. Distinct timescales of RNA regulators enable the construction of a genetic pulse generator. *Biotechnol. Bioeng.* **116**, 1139–1151 (2019).
147. Adamala, K. P., Martin-Alarcon, D. A., Guthrie-Honea, K. R. & Boyden, E. S. Engineering genetic circuit interactions within and between synthetic minimal cells. *Nat. Chem.* **9**, 431–439 (2017).
148. Lentini, R. et al. Fluorescent proteins and in vitro genetic organization for cell-free synthetic biology. *ACS Synth. Biol.* **2**, 482–489 (2013).
149. Goering, A. W. et al. In vitro reconstruction of nonribosomal peptide biosynthesis directly from DNA using cell-free protein synthesis. *ACS Synth. Biol.* **6**, 39–44 (2017).
150. Blanken, D., Foschepoth, D., Serrão, A. C. & Danelon, C. Genetically controlled membrane synthesis in liposomes. *Nat Commun.* **11**, 4317 (2020).

151. Shen, B. et al. The biosynthetic gene cluster for the anticancer drug bleomycin from *Streptomyces verticillus* ATCC 15003 as a model for hybrid peptide-polyketide natural product biosynthesis. *J. Ind. Microbiol. Biotechnol.* **27**, 378–385 (2001).
152. Schwewe, T. et al. The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl Acad. Sci. USA* **92**, 7839–7843 (1995).
153. Nakabachi, A. & Okamura, K. Diaphorin, a polyketide produced by a bacterial symbiont of the Asian citrus psyllid, kills various human cancer cells. *PLoS ONE* **14**, e0218190 (2019).
154. Voet, D. & Voet, J. G. *Biochemistry* (Wiley, 2015).
155. Hutchison, C. A. et al. Design and synthesis of a minimal bacterial genome. *Science* **351**, aad6253 (2016).
156. Gil, R., Silva, F. J., Pereto, J. & Moya, A. Determination of the core of a minimal bacterial gene set. *Microbiol. Mol. Biol. Rev.* **68**, 518–537 (2004).
157. Forster, A. C. & Church, G. M. Synthetic biology projects in vitro. *Genome Res* **17**, 1–6 (2007).
158. Ishihama, Y. et al. Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genomics* **9**, 102 (2008).
159. Berhanu, S., Ueda, T. & Kuruma, Y. Artificial photosynthetic cell producing energy for protein synthesis. *Nat. Commun.* **10**, 1325 (2019).
160. Des Soye, B. J., Davidson, S. R., Weinstock, M. T., Gibson, D. G. & Jewett, M. C. Establishing a high-yielding cell-free protein synthesis platform derived from *Vibrio natriegens*. *ACS Synth. Biol.* **7**, 2245–2255 (2018).
161. Dai, D., Horvath, N. & Varner, J. Dynamic sequence specific constraint-based modeling of cell-free protein synthesis. *Processes* **6**, 132 (2018).
162. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244**, 182–188 (1989).
163. Yin, C. et al. RF1 attenuation enables efficient non-natural amino acid incorporation for production of homogeneous antibody drug conjugates. *Sci. Rep.* **7**, 3026 (2017).
164. 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 (2014).
165. Gan, Q. & Fan, C. Increasing the fidelity of noncanonical amino acid incorporation in cell-free protein synthesis. *Biochim. Biophys. Acta Gen. Subj.* **1861**, 3047–3052 (2017).
166. Oza, J. P. et al. Robust production of recombinant phosphoproteins using cell-free protein synthesis. *Nat. Commun.* **6**, 8168 (2015).
167. Kapoor, N. et al. Malaria derived glycosylphosphatidylinositol anchor enhances anti-Pf25 functional antibodies that block malaria transmission. *Biochemistry* **57**, 516–519 (2018).
168. Li, J. et al. Cogenerating synthetic parts toward a self-replicating system. *ACS Synth. Biol.* **6**, 1327–1336 (2017).
169. Hammerling, M. J. et al. In vitro ribosome synthesis and evolution through ribosome display. *Nat. Commun.* **11**, 1108 (2020).
170. Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S. & Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330–333 (2020).
171. Buntru, M., Vogel, S., Stoff, K., Spiegel, H. & Schillberg, S. A versatile coupled cell-free transcription–translation system based on tobacco BY-2 cell lysates. *Biotechnol. Bioeng.* **112**, 867–878 (2015).
172. Bulleid, N. J. & Freedman, R. B. Cotranslational glycosylation of proteins in systems depleted of protein disulphide isomerase. *EMBO J.* **9**, 3527–3532 (1990).
173. Guarino, C. & Delisa, M. P. A prokaryote-based cell-free translation system that efficiently synthesizes glycoproteins. *Glycobiology* **22**, 596–601 (2012).
174. Jaroentomeechai, T. et al. Single-pot glycoprotein biosynthesis using a cell-free transcription–translation system enriched with glycosylation machinery. *Nat. Commun.* **9**, 2686 (2018).
175. Swank, Z. & Maerkl, S. J. CFPU: a cell-free processing unit for high-throughput, automated in vitro circuit characterization in steady-state conditions. *BioDesign Res.* **2021**, 2968181 (2020).
176. Courtois, F. et al. An integrated device for monitoring time-dependent in vitro expression from single genes in picolitre droplets. *ChemBioChem* **9**, 439–446 (2008).
177. Whitfield, C. J. et al. Cell-free protein synthesis in hydrogel materials. *Chem. Commun.* **56**, 7108–7111 (2020).
178. Lai, S. N. et al. Artificial cells capable of long-lived protein synthesis by using aptamer grafted polymer hydrogel. *ACS Synth. Biol.* **9**, 76–83 (2020).
179. Noireaux, V. & Libchaber, A. A vesicle bioreactor as a step toward an artificial cell assembly. *Proc Natl Acad Sci USA* **101**, 17669–17674 (2004).
180. Lindenburg, L. & Hollfelder, F. 'NAD-display': ultrahigh-throughput in vitro screening of NAD(H) dehydrogenases using bead display and flow cytometry. *Angew. Chem. Int. Ed.* **60**, 9015–9021 (2021).
181. Laohakunakorn, N. et al. Bottom-up construction of complex biomolecular systems with cell-free synthetic biology. *Front. Bioeng. Biotechnol.* **8**, 213 (2020).
182. Libicher, K., Hornberger, R., Heymann, M. & Mutschler, H. In vitro self-replication and multicentric expression of large synthetic genomes. *Nat. Commun.* **11**, 904 (2020).
183. Van Nies, P. et al. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun.* **9**, 1583 (2018).
184. Lavickova, B., Laohakunakorn, N. & Maerkl, S. J. A partially self-regenerating synthetic cell. *Nat. Commun.* **11**, 6340 (2020).
185. Shimojo, M. et al. In vitro reconstitution of functional small ribosomal subunit assembly for comprehensive analysis of ribosomal elements in *E. coli*. *Commun. Biol.* **3**, 142 (2020).
186. Sakatani, Y., Yomo, T. & Ichihashi, N. Self-replication of circular DNA by a self-encoded DNA polymerase through rolling-circle replication and recombination. *Sci. Rep.* **8**, 13089 (2018).
187. Kelwick, R. J. R., Webb, A. J. & Freemont, P. S. Biological materials: the next frontier for cell-free synthetic biology. *Front. Bioeng. Biotechnol.* **8**, 399 (2020).
188. Voegelé, K. et al. Towards synthetic cells using peptide-based reaction compartments. *Nat. Commun.* **9**, 3862 (2018).
189. Schreiber, A., Huber, M. C. & Schiller, S. M. Prebiotic protocell model based on dynamic protein membranes accommodating anabolic reactions. *Langmuir* **35**, 9593–9610 (2019).
190. Caschera, F. & Noireaux, V. Preparation of amino acid mixtures for cell-free expression systems. *Biotechniques* **58**, 40–43 (2015).
191. Caschera, F. et al. Coping with complexity: machine learning optimization of cell-free protein synthesis. *Biotechnol. Bioeng.* **108**, 2218–2228 (2011).
192. Caschera, F. et al. Automated discovery of novel drug formulations using predictive iterated high throughput experimentation. *PLoS ONE* **5**, e8546 (2010).
193. Spice, A. J., Aw, R., Bracewell, D. G. & Polizzi, K. M. Improving the reaction mix of a *Pichia pastoris* cell-free system using a design of experiments approach to minimise experimental effort. *Synth. Syst. Biotechnol.* **5**, 137–144 (2020).
194. Holland, I. & Davies, J. A. Automation in the life science research laboratory. *Front. Bioeng. Biotechnol.* **8**, 571777 (2020).
195. Young, R., Haines, M., Storch, M. & Freemont, P. S. Combinatorial metabolic pathway assembly approaches and toolkits for modular assembly. *Metab. Eng.* **63**, 81–101 (2021).
196. Hillson, N. et al. Building a global alliance of biofoundries. *Nat. Commun.* **10**, 2040 (2019).
197. Segal, M. An operating system for the biology lab. *Nature* **573**, S112–S113 (2019).
198. Eisenstein, M. Enzymatic DNA synthesis enters new phase. *Nat. Biotechnol.* **38**, 1113–1115 (2020).

#### Acknowledgements

This work was supported by the National Science Foundation (NSF) (MCB-1844152, EF-1934496, CBET-1916030), the Human Frontier Science Program (HFSP) (RGP0037/2015) and the Israel Binational Science Foundation (2018208) (V.N. and D.G.). P.F. and M.C.H. are supported by UKRI Engineering and the Physical Sciences Research Council (EPSRC) (EP/T013788/1).

#### Author contributions

Introduction (D.G., M.C.H., P.F., E.A.S. and V.N.); Experimentation (D.G., E.F.R., P.F., E.A.S. and V.N.); Results (D.G., E.F.R., P.F., E.A.S. and V.N.); Applications (D.G., M.C.H., P.F., E.A.S. and V.N.); Reproducibility and data deposition (E.F.R., P.F., E.A.S. and V.N.); Limitations and optimizations (E.F.R., P.F., E.A.S. and V.N.); Outlook (D.G., M.C.H., P.F., E.A.S. and V.N.); Overview of the Primer (P.F., E.A.S. and V.N.).

#### Competing interests

P.F. sits on the Scientific Advisory Board of Tierra Biosciences. V.N. receives research funds from Arbor Biosciences, a distributor of the myTXTL cell-free protein synthesis kit. E.A.S., E.F.R., D.G. and M.C.H. declare no competing interests.

#### Disclaimer

Certain commercial entities, equipment or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials or equipment are necessarily the best available for the purpose.

#### Peer review information

*Nature Reviews Methods Primers* thanks A. Baccouche, D. Karig, S. Maerkl, S. Mansy, D. Yang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

#### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1038/s43586-021-00046-x>.

#### RELATED LINKS

Build-A-Cell: <https://www.buildacell.org/>  
Courses at the Cold Spring Harbor Laboratory: <https://meetings.csh.edu/courseshome.aspx>

© Springer Nature Limited 2021