SYNTHETIC BIOLOGY

Synthetic Biology, 2022, **7(1)**, 1–17
DOI: https://doi.org/10.1093/synbio/ysac015
Advance access publication date: 13 August 2022
Research Article

Effects of DNA template preparation on variability in cell-free protein production

Eugenia Romantseva*, Nina Alperovich, David Ross, Steven P. Lund, and Elizabeth A. Strychalski

National Institute of Standards and Technology, Gaithersburg, MD USA
*Corresponding author: E-mail: eugenia.romantseva@nist.gov

Abstract

DNA templates for protein production remain an unexplored source of variability in the performance of cell-free expression (CFE) systems. To characterize this variability, we investigated the effects of two common DNA extraction methodologies, a postprocessing step and manual versus automated preparation on protein production using CFE. We assess the concentration of the DNA template, the quality of the DNA template in terms of physical damage and the quality of the DNA solution in terms of purity resulting from eight DNA preparation workflows. We measure the variance in protein titer and rate of protein production in CFE reactions associated with the biological replicate of the DNA template, the technical replicate DNA solution prepared with the same workflow and the measurement replicate of nominally identical CFE reactions. We offer practical guidance for preparing and characterizing DNA templates to achieve acceptable variability in CFE performance.

Key words: DNA template; cell-free expression; cell-free protein production; TX-TL; CFE; variability; reproducibility; automation

1. Introduction

Cell-free expression (CFE) increasingly impacts biomanufacturing as an enabling *in vitro* tool. At the state of the art, for example, CFE can provide a high-throughput platform for on-demand biomanufacturing of therapeutics (1–11), rapid prototyping of genetic circuits (12–17), expansion of the genetic toolbox (18–23) and development and optimization of biological sensors (24–27). CFE systems leverage isolated and purified cellular components (28, 29) or cell extracts containing native transcription and translation machinery (30), in an open reaction environment that facilitates the direct addition of enzymes, cofactors, energy molecules and recombinant DNA templates (31). Systems made from the extract of lysed *Escherichia coli* cells remain the most common model system for CFE and can be purchased commercially from a variety of vendors. The synthesis of a fluorescence reporter, such as enhanced green fluorescence protein (eGFP), typically serves to test the performance of a CFE system with regard to its ability to manufacture a protein of interest (32).

Despite the utility and growing popularity of CFE, variability in reaction performance remains a common and persistent challenge (33). To begin to address this challenge, researchers have called for shared protocols and methods containing best practices for typical CFE systems—along with improvements in data accessibility, sharing expertise across laboratories and methods and tools for measuring individual components and whole CFE systems throughout a typical workflow (32, 33). While some variability may be attributed to the inherent complexity of biological systems, studies increasingly explore aspects of this challenge to characterize and improve reproducibility. For example, one study investigated the native bacterial transcription machinery in *E. coli*-based CFE systems to quantify how common extract preparation methods affect extract functionality and performance (34). Other work examined batch-to-batch variability between CFE reactions when extract preparation methods and lysate composition were held constant (35) and evaluated the impact of subsequent variations on RNA circuit characterization (17). Exploration of the variability in transcription and translation in both lysate-based and reconstituted CFE systems revealed that greater variability was present in transcription rather than translation (36). Variability within experimental replicates of CFE was reduced by optimizing extract preparation, solubilizing master mix components and improving the mixing of reaction components (37). Finally, interlaboratory studies quantified the effects of the source and preparation of the cell extract, the composition of the supplemental reaction buffer, the facility at which experiments were conducted and the human operator on variability in CFE (38) and identified the site and operator as significant contributions to variability. Missing from these studies, variability in protein production resulting from the DNA template remains unexplored.

CFE relies on high-concentration, intact and biologically active DNA templates for *in vitro* protein production (33, 39, 40), and common methods to prepare DNA templates may impact the variability in protein titer and the rate of protein produced by the CFE reaction for many reasons. DNA preparation techniques, such as those using spin columns or magnetic beads to extract DNA from cells, may yield damaged templates, due to vigorous pipetting, vortexing or centrifugation, as well as a DNA solution contaminated with phenol, salts, RNA, proteins or polysaccharides (41, 42).

Submitted: 13 January 2022; **Received (in revised form):** 1 July 2022; **Accepted:** 12 August 2022

Commercial DNA preparation kits, which vary in size based on the volume of the bacterial culture and the desired DNA yield, may produce DNA of inconsistent quality, with time-consuming midiprep and maxiprep kits preferred anecdotally over more rapid minipreps (33, 43). While the additional polymerase chain reaction (PCR) cleanup of the extracted DNA template has led to greater protein titer compared to unpurified templates (39), this postprocessing step is not performed routinely. Elution of DNA templates in buffer instead of water, as recommended by many DNA extraction kits, can adversely affect the carefully controlled buffering capacity of CFE systems, contributing further to variability in protein production (39, 40). Rapid DNA quantitation using spectrophotometry, without proper calibration, can yield less accurate measurements of stock DNA concentration than values obtained using fluorometry; this can potentially lead to an unknown and inappropriate concentration of DNA in the CFE reaction, affecting protein production. Storage of the prepared DNA templates, especially at temperatures below 4°C and for periods longer than 1 week, can lead to DNA molecules adhering to the walls of the storage container and reduced concentration of DNA in the CFE reaction (33, 40).

Automation is increasingly integrated into engineering biology and biomanufacturing workflows, in part because it may reduce variability by removing the human element from the biology workflow (44). For example, automation can improve pipetting precision and accuracy during liquid handling steps (45). Automation can also reduce contamination by decreasing the number of manual handling steps during sample preparation (44). Automation enables the collection of large datasets and identification of previously uncharacterized sources of variability (45). CFE systems are amenable to automated workflows with common biofoundry equipment (45, 46), such as automation accessible to academic institutions through the use of community biofoundries and cloud laboratories (47, 48). Several recent examples highlight the utility of acoustic liquid handlers for high-throughput, low-volume CFE, including semi-automated workflows for prototyping and characterizing gene expression elements in nonmodel CFE systems (49). Despite these apparent advantages, the ability for automation to reduce variability in CFE workflows remains uncharacterized.

Here, we explore the effects of DNA template preparation on CFE in the first study of its kind. We measure the concentration of the DNA template and assess the physical damage to the DNA template and purity of the DNA solution resulting from eight DNA preparation workflows. We focus on two common DNA extraction methodologies: one that utilizes filter columns in a bind–wash–elute process to prepare plasmid DNA and a magnetic-bead-based approach that lacks centrifugation-dependent steps. We evaluate the effects on the quality of the DNA template of removing contaminants, such as denatured proteins, genomic DNA and cellular debris from the extracted DNA solution through a postprocessing step. We quantitate the concentration of the prepared DNA in solution using a fluorometric measurement, assess contamination in the DNA solution using spectrophotometry and visualize physical damage to the DNA template with gel electrophoresis. We assess the variance in protein titer and rate of protein production in the CFE reactions containing DNA prepared with eight DNA preparation workflows. We use biological, technical and measurement replicates to isolate and measure sources of variance. To address the growing interest in automating CFE protocols for high-throughput applications and to explore the effects of the human element on variation in titer and rate of protein produced, we compare manual workflows to workflows automated using biofoundry equipment.

2. Materials and methods

A detailed description of the protocol used for this work can be found elsewhere (40).

2.1 Instrument calibration

To characterize and reduce instrument noise, we calibrated all laboratory equipment on the day of use, including incubators used for cell growth, the plate reader used to measure concentration of DNA templates resulting from each DNA preparation workflow and eGFP produced by the CFE reactions, and the automated liquid handler used for all automated workflows (SI Figures S2–S5 and S8–S10). To improve the pipetting accuracy and precision of our automated liquid handler, custom liquid classes were developed (SI Table S1) using a gravimetric approach with the Hamilton Liquid Verification Kit (Hamilton Robotics, 95843-01) for: QIAGEN buffers P1, P2, N3, PB and PE; Omega Bio-tek buffers Solution I, Solution II, N3, ETR Bind, ETR Wash, VHB Wash and SPM; PureLink buffers B2 and W1; Quant-iT Broad-Range double-stranded DNA (dsDNA) working solution and λ DNA standards and nuclease-free water.

Benchtop incubators were calibrated for temperature using a traceable digital bottle thermometer (Traceable Products, 4426 or 4428). The fluorescence plate reader was calibrated in three ways. First, we verified the alignment of the optics, sensitivity of the instrument, the linearity of the signal and leveling of the plate carrier in a pass–fail measurement using the Biotek Fluorescence Test Plate (BioTek Instruments, Inc., 1400501). Second, to identify the systematic bias in fluorescence for measurements of DNA concentration, we prepared three internal calibration standards in 96-well plates (Greiner Bio-One, 655096): a blank plate with empty wells, a nuclease-free water plate with 200 µl of nuclease-free water in each well and a fluorescein plate with 200 µl of 10 nmol/l of National Institute of Standards and Technology (NIST) traceable fluorescein (Thermo Fisher Scientific, Inc., F36915) in each well. Plates were sealed with an aluminum foil seal (Beckman Coulter, 538619) and centrifuged at 2500 *g* for 15 s to ensure that all the liquid was at the bottom of the wells. The aluminum foil was removed prior to imaging at room temperature with excitation and emission wavelengths of 485/20 nm and 528/20 nm, respectively. Third, to identify systematic bias in the fluorescence for measurements of eGFP produced by the CFE reactions, we prepared three internal calibration standards in 384-well plates (Greiner Bio-One, 781892): a blank plate with empty wells, a nuclease-free water plate with 15 µl of nuclease-free water in each well and a fluorescein plate with 15 µl of 1 µmol/l of NIST traceable fluorescein in each well. Plates were sealed with a low-autofluorescence seal (Thermo Fisher Scientific, 232701) and centrifuged at 2500 *g* for 15 s to ensure that all the liquid was at the bottom of the wells. Plates were imaged at 30°C with excitation and emission wavelengths of 485/20 nm and 516/20 nm, respectively. Maintenance of the Hamilton STAR automated liquid handler was performed using a 'Daily maintenance' script in the 'Microlab STAR Maintenance & Verification' application and the maintenance 96 Total Aspiration and Dispense Monitoring (TADM) tool (Hamilton Robotics, 199199101). The performance of the positive pressure module of the Hamilton STAR automated liquid handler was verified using a custom script that rinsed the manifold with water and cycled through a range of pressures to ensure optimal performance.

2.2 Cell culture

For this work, we used a circular DNA template, pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 (SI Figure S1), with a modified bacteriophage

Lambda Cro promotor, ampicillin resistance gene, eGFP gene and T500 transcriptional terminator (50). This DNA template has previously served as a positive control plasmid for the myTXTL CFE system (51) and produces a strong fluorescence signal around an emission wavelength of 488 nm. pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 was a gift from Vincent Noireaux (Addgene plasmid #40019; http://n2t.net/addgene:40019; RRID:Addgene_40019).

Briefly, the bacterial stab containing pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 was streaked across LB (Luria-Broth) agar ampicillin (100 ug/mL) culture plates, and the plates were incubated at 30°C for 16 h. Starter liquid culture was prepared in 14-ml culture tubes (Fisher Scientific, 14-959-11B) by inoculating 5 ml of rich M9 media (3 g/l KH_2PO_4, 6.78 g/l Na_2HPO_4, 0.5 g/l NaCl, 1 g/l NH_4Cl (Millipore Sigma M6030-1 KG), 0.1 mmol/l $CaCl_2$ (Millipore Sigma, 21115), 2 mmol/l $MgSO_4$ (Millipore Sigma, 230391), 4% glycerol (Thermo Fisher Scientific, Inc., 15514011) and 20 g/l casamino acids (Millipore Sigma, 2240-500GM)) supplemented with 100 μg/ml ampicillin (Millipore Sigma, A0166) with a single colony. In this study, each single colony of *E. coli* containing pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 was considered its own biological replicate and, therefore, a biologically distinct sample (52). Starter liquid culture was incubated at 30°C with 300 RPM shaking for 20 h or until the optical density at 600 nm (OD_{600}) of each culture was between 4.0 and 4.5. Eight starter liquid cultures, each from a different colony, were grown, and the five with the highest OD_{600} were chosen for further growth. Each of these five starter liquid cultures was grown into a large liquid culture, giving five biological replicates for the experiment. Each large liquid culture was prepared in 250-ml culture flasks by inoculating 120 ml of rich M9 media supplemented with 100 μg/ml ampicillin with 120 μl of starter liquid culture. Large liquid cultures were incubated at 30°C with 300 RPM shaking for 24 h or until OD_{600} of each culture was between 9.0 and 10.0. Cells were cooled to 4°C for 15 min to stop growth and then aliquoted using an automated liquid handler into 96-well plates (Eppendorf, 951033405) for automated DNA preparation workflows and 2-ml microcentrifuge tubes (CELLTREAT Scientific Products, 229446) for manual DNA preparation workflows. These aliquots or technical replicates were used as repeated measurements of the same biological sample (52). A total of 200 technical replicates of liquid culture were prepared for this study, including 100 tubes and 100 wells across two 96-well plates. This constituted 40 technical replicates of each biological replicate, 20 in tubes and 20 in wells. Cells were pelleted by centrifugation at 4500 *g* for 15 min at room temperature. Supernatant was removed from the 96-well plates by aspiration using an automated liquid handler and from the tubes by inverting the tubes by hand over a waste container. Pelleted cells were frozen at −20°C overnight.

2.3 DNA preparation workflows

DNA solutions were prepared with eight DNA preparation workflows (MFP-, MFP_+, MBP-, MBP_+, AFP-, AFP_+, ABP- and ABP_+) that explored two sample preparation methodologies (manual and automated), two DNA extraction methodologies (filter column or plate and magnetic bead) and the presence or absence of a postprocessing step (Figure 1).

MFP- workflow: manual filter column-based DNA extraction using the QIAprep Spin Miniprep kit (QIAGEN, 27106) without additional postprocessing The QIAGEN QIAprep Spin Miniprep kit was chosen to investigate previous inconsistent results in the concentration of DNA and quality of the DNA, both in terms of physical damage to the DNA template and purity of the DNA solution, prepared with miniprep kits, as opposed to midi and maxi prep kits (33); this kit uses filter columns in a bind–wash–elute procedure. Buffers P1 and PE were prepared according to manufacturer's instructions. Buffer N3 was cooled to 4°C for at least 30 min and kept on ice for the duration of the workflow. Nuclease-free water for DNA elution was warmed to 65°C for at least 30 min and kept at that temperature in a dry bath for the duration of the workflow.

Pelleted cells were removed from −20°C and allowed to come to room temperature. Cells were resuspended with 250 μl of Buffer P1 by vortexing and pipetting. Resuspended cells were lysed with 250 μl of Buffer P2, and the contents of each tube were mixed by inverting the tubes 10 times. Lysed cells were neutralized with 350 μl of cold Buffer N3, and the contents of each tube were mixed by inverting the tubes 20 times. Neutralized cells were incubated on ice for 5 min and then centrifuged for 30 min at 18 000 *g* to pellet the cell debris. A total of 700 μl of clear supernatant was transferred to each QIAprep 2.0 spin column, which was then centrifuged at 18 000 *g* for 1 min. The flow through was discarded. A total of 500 μl of Buffer PB was added to each spin column, and the spin columns were centrifuged at 18 000 *g* for 1 min. The flow through was discarded. A total of 750 μl of Buffer PE was added to each spin column, and the spin columns were centrifuged at 18 000 *g* for 1 min. The flow through was discarded. The spin columns were dried by centrifugation at maximum speed for 1 min. DNA was eluted from the spin columns into DNA LoBind tubes (Eppendorf, 022431021) by adding 55 μl of nuclease-free water at 65°C to the center of each spin column. Spin columns were incubated for 5 min and centrifuged 18 000 *g* for 1 min. Eluted DNA was stored at 4°C for up to 1 week.

MBP- workflow: manual magnetic bead-based DNA extraction using the Mag-Bind Ultra Pure Plasmid DNA 96 kit (Omega Bio-tek. Inc., M1258-01) without additional postprocessing The Omega Bio-tek Mag-Bind Ultra Pure Plasmid DNA 96 kit was chosen for its lack of centrifugation-dependent steps and associated high shear forces that may lead to breakage, especially for larger plasmids and genomic DNA; this kit uses magnetic beads to selectively bind DNA from other cellular components. Solution I, VHB Buffer and SPM Wash Buffer were prepared according to manufacturer's instructions. N3 Buffer was cooled to 4°C for at least 30 min and kept on ice for the duration of the workflow. Nuclease-free water for DNA elution was warmed to 65°C for at least 30 min and kept at that temperature in a dry bath for the duration of the workflow.

Pelleted cells were removed from −20°C and allowed to come to room temperature. Cells were resuspended with 250 μl of Solution I by vortexing and pipetting. Resuspended cells were lysed with 250 μl of Solution II, and the contents of each tube were mixed by inverting the tubes 10 times. Lysed cells were neutralized with 125 μl of cold N3 Buffer, and the contents of each tube were mixed by inverting the tubes 20 times. Neutralized cells were incubated on ice for 5 min and then centrifuged for 30 min at 18 000 *g* to pellet the cell debris. A total of 500 μl of clear supernatant was transferred to 2.0-ml microcentrifuge tubes (CELLTREAT Scientific Products, 229 446). A total of 500 μl of ETR Binding Buffer and 30 μl of MagBind Particles RQ were added to each microcentrifuge tube. Tubes were vortexed for 10 s to mix the contents and incubated for 5 min at room temperature. The tubes were transferred to the magnetic separation rack (New England Biolabs. Inc., S1509S) and incubated for 5 min to clear the beads. The clear supernatant was aspirated and discarded. Beads were washed with ETR Wash Buffer, by adding 500 μl of buffer to each tube and vortexing the tubes for 6 s to resuspend the beads. The tubes were transferred

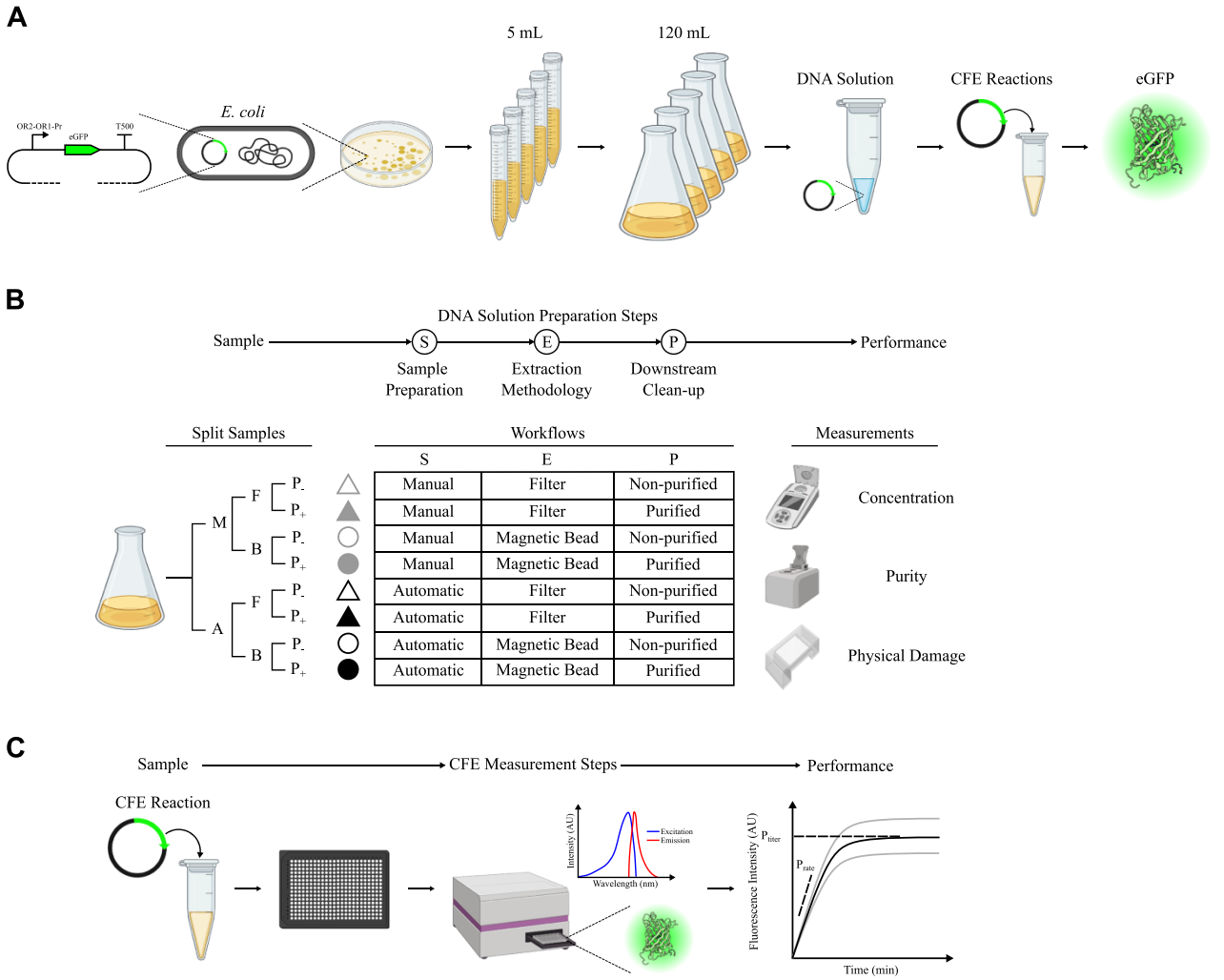

Figure 1. Experimental design. This study, characterized the effects of DNA template preparation on variability in eGFP produced by the myTXTL *E. coli*-lysate-based CFE system. (A) *E. coli* cells containing the pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 plasmid were grown from individual colonies in large liquid cultures. Plasmids were extracted using eight different DNA preparation workflows. The resulting DNA solutions were used to prepare CFE reactions that produced eGFP upon incubation at 30°C. (B) Each cell culture was divided (split) to perform eight different DNA preparation workflows (MFP-, MFP_+, MBP-, MBP_+, AFP-, AFP_+, ABP- and ABP_+) that explored two sample preparation methodologies (manual and automated), two DNA extraction methodologies (filter column or plate and magnetic bead) and the presence or absence of a postprocessing step. The concentration of the DNA in the prepared DNA solutions was quantified using fluorometry, and the purity of the DNA solutions were measured using UV spectrophotometry. Physical damage to the DNA template was assessed using gel electrophoresis. (C) CFE reactions were assembled in PCR tubes and then transferred to 384-well plates for incubation at 30°C. eGFP production was measured using a plate reader. The resulting fluorescence intensity measurements were used to calculate protein titer (P_{titer}), protein rate (P_{rate}), variance in P_{titer} and variance in P_{rate}.

to the magnetic separation rack and incubated for 90 s to clear the beads. The clear supernatant was aspirated and discarded. The beads were washed with VHB Buffer, by adding 700 µl of buffer to each tube and vortexing the tubes for 3 s to resuspend the beads. The tubes were transferred to the magnetic separation rack and incubated for 90 s to clear the beads. The clear supernatant was aspirated and discarded. A total of two washes with VHB Buffer were performed. The beads were washed with SPM Wash Buffer, by adding 700 µl of buffer to each tube and vortexing the tubes for 3 s to resuspend the beads. The tubes were transferred to the magnetic separation rack and incubated for 90 s to clear the beads. The clear supernatant was aspirated and discarded. The magnetic beads were dried in a dry bath at 75°C for 7 min. DNA was eluted by adding 60 µl of nuclease-free water at 65°C to the center of each bead pellet. The beads were resuspended by vortexing for 6 s and then incubated for 5 min. The tubes were transferred to the magnetic separation rack and incubated for 90 s to clear the beads. The clear supernatant containing the DNA solution was aspirated and transferred to DNA LoBind (Eppendorf, 022431021) tubes for storage. Eluted DNA was stored at 4°C for up to 1 week.

AFP- workflow: automated filter column-based DNA extraction using the QIAprep 96 Turbo Miniprep kit (QIAGEN, 27191) without additional postprocessing The QIAGEN QIAprep 96 Turbo kit was the matching counterpart to the QIAprep Spin Miniprep kit used for manual filter-based extraction of plasmid DNA. This kit uses a filter plate to clear the cell lysate of debris and a binding plate for DNA extraction in a bind–wash–elute procedure. Buffers P1 and PE were prepared according to manufacturer's instructions. Buffer N3 was cooled to 4°C for at least 30 min and kept at this temperature for the duration of the workflow. Nuclease-free water for DNA elution was warmed to 65°C for at least 30 min and kept at this temperature for the duration of the workflow. Buffers were aliquoted into multi-well reservoir plate(s)

for automated liquid handling as follows: 15 ml of Buffer P1 was added to Column 1 of a 6-column reservoir (Agilent, 204284-100); 20 ml of Buffer P2 was added to each column of a 2-column reservoir (Agilent, 204359-100); 15 ml of Buffer N3 was added to each column of a 2-column reservoir; 100 ml of Buffer PB was added to Column 1 and 100 ml of Buffer PE was added to Column 2 of a 2-column reservoir and 15 ml of nuclease-free water was added to Column 1 of a 6-column reservoir. All reservoirs were covered with a sterile lid (Agilent, 202497-100). All pipetting steps were performed using an automated liquid handler.

The 96-well plates containing pelleted cells were removed from –20°C and allowed to come to room temperature. Cells were resuspended with 200 µl Buffer P1 by pipetting. Resuspended cells were lysed with 250 µl of Buffer P2, and the contents of each well were mixed by shaking the plate at 90 RPM for 120 s. Lysed cells were neutralized with 350 µl of cold Buffer N3, and the contents of each well were mixed by shaking the plate at 90 RPM for 120 s and pipetting using widebore tips. Neutralized cells were transferred to a 96-well filter plate using widebore tips. The contents of the filter plate were processed into a binding plate using an automated positive pressure filter press at 34.5 kPa (5 psi) for 7 min. The filter plate was then discarded. The contents of the binding plate were processed to waste at 275.8 kPa (40 psi) for 1 min. A total of 900 µl of Buffer PB was added to each well of the binding plate, and the contents of the binding plate were processed to waste at 275.8 kPa (40 psi) for 1 min. A total of 900 µl of Buffer PE was added to each well of the binding plate, and the contents of the binding plate were processed to waste at 275.8 kPa (40 psi) for 1 min. The binding plate was dried by processing the contents of all the wells to waste at 448.2 kPa (65 psi) for 10 min. DNA was eluted from the binding plate by adding 110 µl of nuclease-free water at 65°C to the center of each well. The binding plate was then incubated for 5 min, and its contents were processed into a DNA LoBind 96-well plate (Eppendorf, 951032000) at 448.2 kPa (65 psi) for 10 min. Eluted DNA was stored at 4°C for up to 1 week.

ABP- workflow: automated magnetic bead-based DNA extraction using the Mag-Bind Ultra Pure Plasmid DNA 96 kit (Omega Bio-tek. Inc., M1258-01) without additional post-processing The Omega Bio-tek Mag-Bind Ultra Pure Plasmid DNA 96 kit can be used for manual or automated workflows. Solution I, VHB Buffer and SPM Wash Buffer were prepared according to manufacturer's instructions. N3 Buffer was cooled to 4°C for at least 30 min and kept at that temperature for the duration of the workflow. Nuclease-free water for DNA elution was warmed to 65°C for at least 30 min and kept at that temperature for the duration of the workflow. Buffers were aliquoted into multi-well reservoir plate(s) for automated liquid handling as follows: 15 ml of Solution I was added to Column 2, 30 ml of ETR Binding Buffer to Column 5 and 30 ml of ETR Wash Buffer to Column 6 of a 6-column reservoir; 20 ml of Solution II was added to each column of a 2-column reservoir; 15 ml of N3 Buffer was added to each column of a 2-column reservoir; 40 ml of VHB Buffer was added to Columns 1 and 2 and 40 ml of SPM Wash Buffer to Column 6 of a 6-column reservoir and 15 ml of nuclease-free water was added to Column 3 of a 6-column reservoir. MagBind Particles RQ were mixed by vortexing for 10 s and aliquoted into 1.5-ml microcentrifuge tubes by adding 500 µl of particles to each tube. The tubes were loaded into a 32-tube carrier with 1.5-ml tube inserts (Hamilton Robotics, 182238) on the deck of the liquid handler. All reservoirs were covered with a sterile lid. A universal magnetic plate (Alpaqua, A00400) was used for pelleting the magnetic beads in each well. A filter plate (QIAGEN, 27191) was used to clear the cell lysate of debris. All pipetting steps were performed using an automated liquid handler.

The 96-well plates containing pelleted cells were removed from –20°C and allowed to come to room temperature. Cells were resuspended with 200 µl of Solution I by pipetting. Resuspended cells were lysed with 250 µl of Solution II, and the contents of each well were mixed by shaking the plate at 90 RPM for 120 s. Lysed cells were neutralized with 125 µl of cold N3 Buffer, and the contents of each well were mixed by shaking the plate at 90 RPM for 120 s and pipetting using widebore tips. Neutralized cells were transferred to a 96-well filter plate using widebore tips. The contents of the filter plate were processed into a 96-well lysate plate (Thermo Fisher Scientific, AB-1127) using an automated positive pressure filter press at 34.5 kPa (5 psi) for 7 min; the filter plate was then discarded. A total of 500 µl of clear supernatant was transferred from each well of the lysate plate into a 96-well receiving plate (Thermo Fisher Scientific, AB-1127); the lysate plate was then discarded. A total of 500 µl of ETR Binding Buffer and 30 µl of Mag-Bind Particles RQ were added to each well, and the contents of each well were mixed by pipetting. The plate was transferred to the magnetic separation module and incubated for 5 min to clear the beads. The clear supernatant was aspirated and discarded. The beads were washed with ETR Wash Buffer by adding 500 µl of buffer to each well. The contents of the wells were mixed by pipetting at 400 µl/s for 10 cycles, followed by shaking at 800 RPM for 60 s. The plate was transferred to the magnetic separation module and incubated for 90 s to clear the beads. The clear supernatant was aspirated and discarded. The beads were washed with VHB Buffer by adding 700 µl of buffer to each well. The contents of the wells were mixed by pipetting at 600 µl/s for 10 cycles, followed by shaking at 700 RPM for 60 s. The plate was transferred to the magnetic separation module and incubated for 90 s to clear the beads. The clear supernatant was aspirated and discarded. A total of two washes with VHB Buffer were performed. The beads were washed with SPM Buffer by adding 700 µl of buffer to each well. The contents of the wells were mixed by pipetting at 500 µl/s for 10 cycles, followed by shaking at 700 RPM for 60 s. The plate was transferred to the magnetic separation module and incubated for 90 s to clear the beads. The clear supernatant was aspirated and discarded. The magnetic beads were dried by incubating the plate in a heater–shaker module at 75°C for 7 min. DNA was eluted by adding 70 µl of nuclease-free water at 65°C to the center of each well. The beads were resuspended by shaking at 900 RPM for 60 s and then incubated for 5 min. The plate was transferred to the magnetic separation module and incubated for 120 s to clear the beads. The clear supernatant from each well was aspirated and transferred to a DNA LoBind 96-well plate. Eluted DNA was stored at 4°C for up to 1 week.

Manual postprocessing of DNA solutions using the PureLink PCR Purification kit (Thermo Fisher Scientific, K310002) as part of the MFP_+ and MBP_+ workflows Additional postprocessing of the DNA solutions was included, because it has been previously reported to improve the quality of the DNA templates for CFE applications (39). The Invitrogen PureLink PCR Purification kit uses filter columns to remove contaminants, such as primers, Deoxyribonucleotide triphosphates (dNTPs), enzymes and salts, from PCR amplicons. Binding Buffer B2 and Wash Buffer W1 were prepared according to manufacturer's instructions. Nuclease-free water for DNA elution was warmed to 65°C for at least 30 min and kept at that temperature in a dry bath for the duration of the workflow.

DNA solutions prepared with manual filter-based and bead-based extraction workflows were removed from 4°C and allowed to come to room temperature. In DNA LoBind tubes, 50 µl of eluted DNA solution was combined with 200 µl of Binding Buffer B2 using widebore pipette tips. The contents of the tubes were mixed by inverting 10 times and transferred to PureLink PCR spin columns using widebore tips. Spin columns were then centrifuged at 10 000 *g* for 1 min, and the flow through was discarded. The spin columns were washed with Wash Buffer W1 by adding 650 µl of buffer to each column. Spin columns were centrifuged at 10 000 *g* for 1 min, and the flow through was discarded. The spin columns were dried by centrifugation at maximum speed for 3 min. For DNA elution, spin columns were transferred to clean DNA LoBind tubes, and 50 µl of nuclease-free water at 65°C was added to the center of each column. Spin columns were incubated for 5 min and centrifuged at maximum speed for 2 min. The spin columns were discarded, and eluted DNA solution in DNA LoBind tubes was stored at 4°C for up to 1 week.

6. Automated postprocessing of DNA solutions using the PureLink Pro 96 PCR Purification kit (Thermo Fisher Scientific, K3100-96A) as part of the AFP_+ and ABP_+ workflows

The Invitrogen PureLink Pro 96 PCR Purification kit was the matching counterpart to the PureLink PCR Purification kit used for postprocessing DNA solutions manually. Binding Buffer B2 and Wash Buffer W1 were prepared according to manufacturer's instructions. Nuclease-free water for DNA elution was warmed to 65°C for at least 30 min and kept at that temperature for the duration of the workflow. Buffers were aliquoted into multi-well reservoir plate(s) for automated liquid handling as follows: 25 ml of B2 Binding Buffer was added to Column 1 and 75 ml of W1 Wash Buffer was added to Column 2 of a 2-column reservoir. The reservoir was covered with a sterile lid. All pipetting steps were performed using an automated liquid handler.

DNA solutions prepared with automated filter-based and bead-based extraction workflows were removed from 4°C and allowed to come to room temperature. In a DNA LoBind 96-well plate, 50 µl of eluted DNA solution was combined with 200 µl of Binding Buffer B2. The contents of the LoBind plate were mixed by shaking at 300 RPM for 120 s and transferred to a PureLink filter plate. The contents of the filter plate were processed to waste using an automated positive pressure filter press at 275.8 kPa (40 psi) for 1 min. A total of 600 µl of Wash Buffer W1 was added to each well of the filter plate, and the contents of the plate were processed to waste at 275.8 kPa (40 psi) for 1 min. The filter plate was dried by processing the contents of all the wells to waste at 448.2 kPa (65 psi) for 10 min. DNA was eluted from the filter plate by adding 60 µl of nuclease-free water at 65°C to the center of each well. The plate was incubated for 5 min, and its contents were processed into a DNA LoBind 96-well plate at 448.2 kPa (65 psi) for 10 min. Eluted DNA was stored at 4°C for up to 1 week.

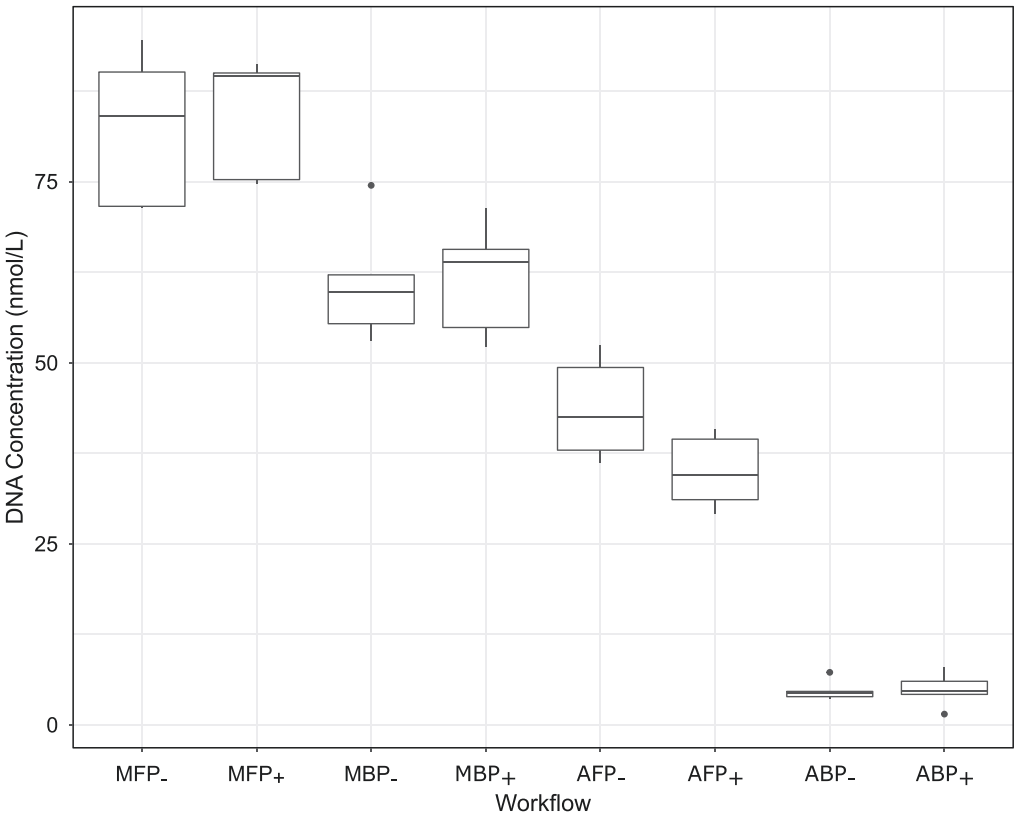

Figure 2. DNA concentration. Box plots of DNA concentration for each DNA preparation workflow. Each box plot represents DNA concentration from five biological replicates. The line inside the box designates the median concentration. Upper and lower box boundaries correspond to the upper quartile or 75th percentile of the data and the lower quartile or 25th percentile of data, respectively. Upper and lower whiskers designate the maximum and minimum values, respectively, excluding the outliers. Filled circles are outliers that are >1.5 times the inner quartile range above the upper quartile or <1.5 times the inner quartile range below the lower quartile, respectively.

2.4 Fluorometric analysis of DNA concentration using the Quant-iT dsDNA Broad-Range Assay Kit (Thermo Fisher Scientific, Q33130)

The concentration of DNA in each of the DNA solutions prepared with the eight DNA preparation workflows (Figure 2) was quantified in 96-well plates using an automated fluorometric assay and a plate reader. To determine the concentration of DNA in each DNA solution, each 96-well plate contained two eight-point calibration curves relating known concentrations of λ DNA standards to measured fluorescence intensity. To mitigate potential degradation of the DNA sample during storage, all DNA solutions were analyzed within 24–72 h of the completion of each DNA preparation workflows. DNA concentration was not quantified using ultraviolet (UV) spectrophotometry because this measurement technique often yields an overestimate due to the absorption of UV light at 260 nm by common impurities, such as RNA and protein, along with the purine and pyrimidine bases in DNA (42). The Quant-iT Broad-Range dsDNA working solution was prepared by combining 600 µl of Component A with Component B to a final volume of 200 ml in a 250-ml Nalgene Polyethylene terephthalate glycol (PETG) bottle. The working solution was then transferred to a black single-cavity reservoir (Agilent, custom order) for automated liquid handling and covered with a black sterile lid (Agilent, custom order). All pipetting steps were performed using an automated liquid handler.

Prepared DNA solutions were removed from 4°C and allowed to come to room temperature. The Quant-iT working solution was aliquoted across four 96-well black quantitation plates (Greiner Bio-One, 655096) by pipetting 190 µl/well into all wells in Columns 1 and 12 and 195 µl/well to all wells in Columns 2 through 11. λ DNA standards (0 ng/µl, 5 ng/µl, 10 ng/µl, 20 ng/µl, 40 ng/µl, 60 ng/µl, 80 ng/µl and 100 ng/µl) were added to Columns 1 and 12 of all quantitation plates by pipetting 10 µl/well. The prepared DNA solutions were added to Columns 2 through 11 of all quantitation plates by pipetting 5 µl/well. The contents of each quantitation plate were mixed by shaking at 200 RPM for 2 min. DNA concentration in each well was measured from the bottom using a fluorescent plate reader, with excitation and emission wavelengths set to 485/20 nm and 528/20 nm, respectively. Mean fluorescence intensity of 10 measurements was reported by the plate reader from each sample well; standard deviation from these measurements was not reported.

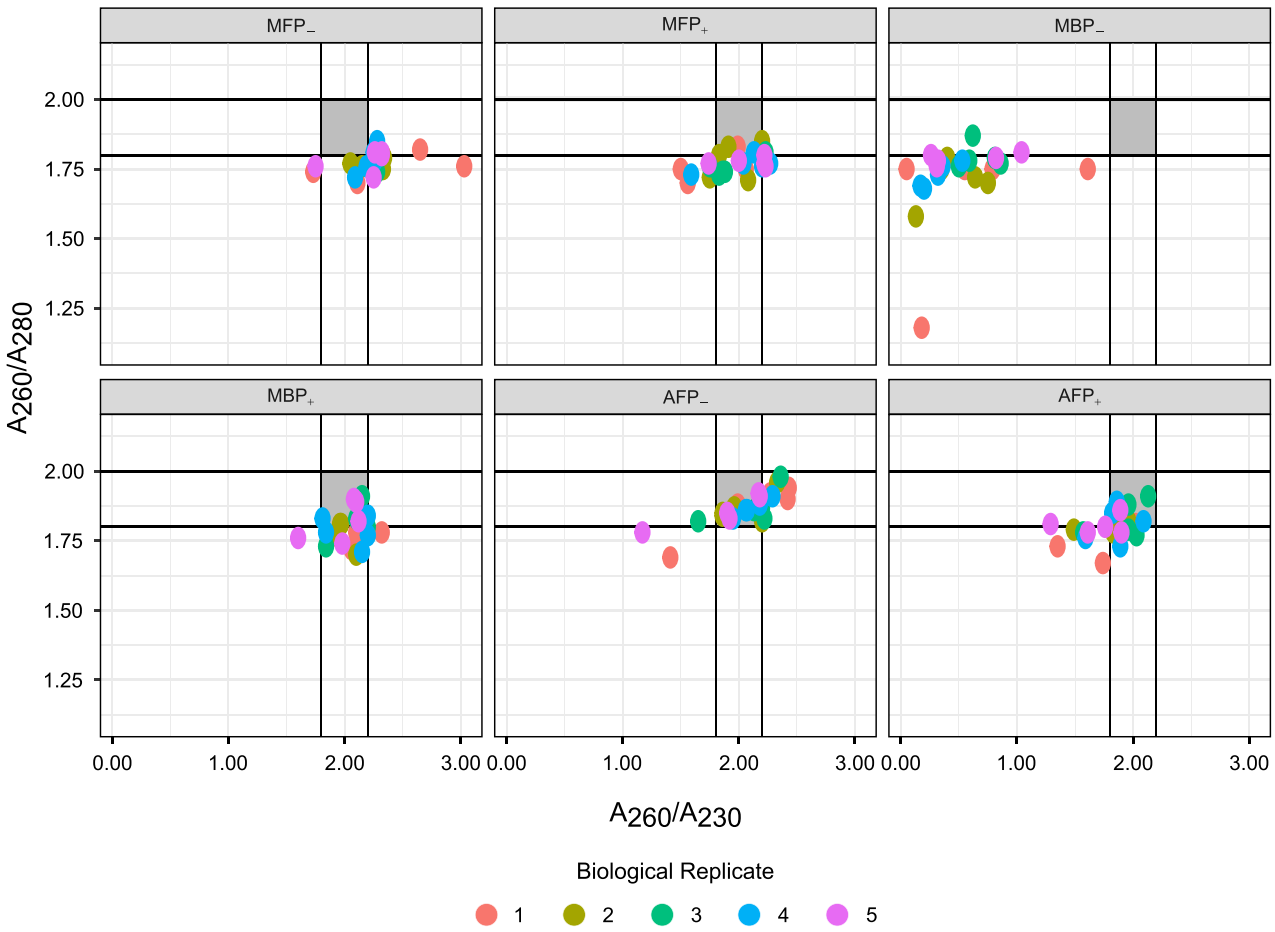

Figure 3. DNA solution purity. Scatter plots of DNA solution purity for each DNA preparation workflow taken from measurements of A260/A280 and A260/A230 ratios. Lines and shaded areas indicate upper and lower limits of community accepted purity ratios(42), respectively. Purity ratios for DNA templates prepared with ABP- and ABP+ workflows are not reported as measurements of DNA purity with UV spectrophotometry are unreliable for solutions containing less than 8.60 nmol/L of DNA. MFP-: manual, filter extracted, non-purified; MFP+: manual, filter extracted, purified; MBP-: manual, magnetic bead extracted, non-purified; MBP+: manual, magnetic bead extracted, purified; AFP-: automated, filter extracted, non-purified; AFP+: automated, filter extracted, purified; ABP-: automated, magnetic bead extracted, non-purified; ABP+: automated, magnetic bead extracted, purified.

2.5 Manual spectrophotometric analysis of DNA quality

Contamination in the DNA solution (Figure 3), for example, due to the presence of RNA, proteins, phenols, salts and polysaccharides, was assessed manually using UV spectrophotometry (Denovix, DS-11), by measuring each solution's absorbance spectrum between 200 nm and 320 nm and calculating the A_{260}/A_{230} and A_{260}/A_{280} purity ratios (42, 53). A_{260}/A_{230} purity ratios indicate DNA suitability for microarray applications, and values from 1.8 to 2.2 are the community-accepted range for pure DNA. A_{260}/A_{280} ratios indicate DNA suitability for PCR applications, and values from 1.8 to 2.0 are the community-accepted range for pure DNA. Absorbance measurements were performed at the manufacturer's default settings using the 'dsDNA' application. To mitigate the potential degradation of the DNA sample during storage, all DNA solutions were analyzed within 24–72 h of the completion of each DNA preparation workflows.

The prepared DNA solutions were removed from 4°C and allowed to come to room temperature. The stage of the spectrophotometer was washed with Deionized (DI) water and wiped dry using Kim wipes. A blank sample composed of 1 μl nuclease-free water was measured prior to measurements of DNA solutions. Each prepared DNA solution was measured by adding 1 μl of the solution to the sample stage, closing the lid and pressing the 'Measure' button.

2.6 Gel electrophoresis

Physical damage to the DNA template was assessed using manual gel electrophoresis, with specific attention to breakage and shearing, for example, due to centrifugation, vortexing, pipetting and passage through a separation column. Because of the large number of DNA samples, we looked at physical damage in only one technical replicate for each biological replicate prepared with each DNA preparation workflow.

Briefly, dilute DNA solutions were prepared in 200-μl PCR tubes (USA Scientific, 1402-4700) by combining 5 μl of prepared DNA solution with 15 μl of nuclease-free water. The contents of the tubes were mixed by gentle tapping and then the tubes were briefly spun down. 1.2% agarose E-gels (Thermo Fisher Scientific, G501801) were loaded by adding 20 μl of 1-kb ready-to-use DNA ladder (Thermo Fisher Scientific, SM1333) to Lanes 1 and 12 and 20 μl of each diluted DNA solution to Lanes 2 through 11. Gels were run for 45 min on the E-gel Power Snap Electrophoresis Device (Thermo Fisher Scientific, G8100) and then imaged using a gel reader.

2.7 CFE assay

We chose the *E. coli*-lysate-based CFE system myTXTL (Daicel Arbor Biosciences), because *E. coli* remains the most common biological system for CFE. Working with a single lot of the commercially available myTXTL for all experiments removed the CFE system from consideration as a possible source of variability in protein titer and rate of protein production in our study. DNA templates prepared with the eight DNA preparation workflows were used to assemble a total of 450 CFE reactions. To achieve the recommended concentration of DNA for this CFE system by the manufacturer, nuclease-free water was used to adjust DNA concentration in each reaction to 5 nmol/l, except in the case of CFE reactions containing DNA prepared with ABP- and ABP_+ workflows, as described below. Reactions were prepared 24–72 h after DNA quantitation from DNA solutions stored at 4°C. This ensured an accurate measurement of DNA concentrations and reduced the chance of DNA breakage due to storage at lower temperatures, multiple freeze–thaw cycles and adsorption of the DNA to the storage containers.

The CFE assay consisted of four types of samples: (i) a 4-point recombinant eGFP calibration curve (0 μmol/l, 5 μmol/l, 10 μmol/l and 30 μmol/l); (ii) blank samples (nuclease-free water and 1× phosphate-buffered saline (PBS; Thermo Fisher Scientific, 10010023)); (iii) a negative control reaction and (iv) CFE reactions composed of myTXTL Sigma 70 Master Mix (Daicel Arbor Biosciences), 5 nmol/l of DNA template and nuclease-free water. Each of the four types of samples were prepared in 200-μl PCR tubes and stored on ice protected from light. Once prepared, samples were loaded at 10 μl/well into a 384-well CFE plate (Greiner Bio-One, 781892) for incubation and analysis (SI Table S3 and SI Figure S11). Because of the large number of DNA solutions in this study, a total of three CFE plates were used to assay all 450 CFE reactions. CFE plates were prepared and assayed individually over 72 h. On each CFE plate, CFE reactions were added in triplicate. The eGFP calibration curve, the blank samples and the negative control were added to each CFE plate in triplicate in three separate locations (SI Figure S11) to assess the variation in protein production across the plate. The CFE plate was kept on ice and protected from light until all samples had been added. The plate was then sealed with a low-autofluorescence seal, centrifuged at 2500 *g* for 15 s and immediately transferred to a preheated plate reader for incubation at 30°C. eGFP produced by each sample was measured every 5 min for 24 h, without shaking and with measurements taken through the bottom of the plate with excitation and emission wavelengths of 485/20 nm and 516/20 nm, respectively.

A recombinant eGFP (Cell Biolabs, Inc.) calibration curve was included on each CFE plate to determine the concentration of eGFP in units of μM produced by the CFE reactions. Recombinant eGFP standard was removed from storage in −80°C and defrosted on ice, protected from light. The concentration of the standard was verified using spectrophotometry (Denovix, DS-11, 'Protein A280' application) as follows: the stage of the spectrophotometer was washed with DI water and wiped dry using Kim wipes; a blank sample composed of 1 μl 1× PBS was measured and the concentration of the eGFP standard (29 kDa, 55 000 M^{-1} cm^{-1}) was measured by adding 1 μl of the standard to the sample stage, closing the lid and pressing the 'Measure' button. The mean concentration from five replicate measurements was used to prepare a four-point calibration curve in 1× PBS in 200-μl PCR tubes: 0 μmol/l, 5 μmol/l, 10 μmol/l and 30 μmol/l.

A blank sample consisting of nuclease-free water was included on each CFE plate to account for background fluorescence in nuclease-free water used to dilute DNA solutions to 5 nmol/L in each CFE reaction. Nuclease-free water blank samples were prepared by aliquoting 36 μl of nuclease-free water into three 200-μl PCR tubes.

Another blank sample consisting of 1× PBS was included on each CFE plate to account for background fluorescence in 1× PBS used to prepare the eGFP calibration curve. PBS blank samples were prepared by aliquoting 36 μl of 1× PBS into three 200-μl PCR tubes.

A negative control sample was included on each CFE plate to account for background fluorescence from the myTXTL Sigma 70 Master Mix in each CFE reaction. The negative control was prepared in a 200-μl PCR tube by adding 27 μl of nuclease-free water to Sigma 70 Master Mix to a final volume of 108 μl; the DNA template was omitted.

CFE reactions containing DNA prepared with MFP-, MFP_+, MBP-, MBP_+, AFP- and AFP_+ workflows were assembled in 200-μl

PCR tubes by adding the DNA template and nuclease-free water to 27 µl Sigma 70 Master Mix, for a final concentration of 5 nmol/l DNA per 36 µl of total sample volume. CFE reactions containing DNA templates prepared with ABP- and ABP_+ workflows contained a variable concentration of DNA <5 nmol/l, because of the insufficient DNA concentrations resulting from these workflows. For those CFE reactions containing DNA templates prepared with ABP- and ABP_+ workflows, DNA templates were added to 27 µl Sigma 70 Master Mix to a total sample volume of 36 µl.

2.8 Data analysis and visualization

To calculate the stock concentration of each DNA solution, we used an ordinary linear regression model to fit the observed fluorescence intensity measurements versus nominal DNA concentration for the mean of two series of calibration samples of λ DNA standards on each quantitation plate (SI Figure S6).

The myTXTL CFE system itself has a non-zero fluorescence signal when measured in the same fluorescence channel as eGFP, and we performed a background subtraction of this signal from all CFE reactions. The mean fluorescence intensity of all nine negative control CFE reactions on each CFE plate at time point 0 was subtracted from the fluorescence intensity of each CFE reaction at each time point. CFE reactions containing DNA templates prepared with ABP- and ABP_+ workflows contained less than the nominal 5 nmol/l of DNA added to all other CFE reactions. For this reason, fluorescence intensity measurement for CFE reactions containing DNA prepared with ABP- and ABP_+ workflows were first scaled by the DNA concentration added to each reaction to facilitate comparison across all CFE reactions prior to background subtraction.

The resulting fluorescence intensities were used to characterize the rate of protein production (P_{rate}) and protein titer (P_{titer}). For each CFE reaction, P_{rate} was evaluated as the slope from an ordinary least-squares regression line fit to fluorescence intensity measurements over the first 250 min. P_{titer} was represented as the fluorescence intensity at 960 min, to match the recommended incubation time for measuring the endpoint yield of protein produced with myTXTL CFE reactions (51). We compared the magnitude and relative variance of P_{rate} and P_{titer} for each pair of workflows. To compare magnitudes for feature P (P_{rate} or P_{titer}), we computed the mean value of P for each of the five biological replicates for each DNA preparation workflow. To identify statistically significant differences in the mean values of P between the workflows, we applied a Friedman test to the collection of mean values of P, treating workflow as the factor of interest and biological replicate as a blocking factor. Pairwise workflow comparisons were conducted using Nemenyi's all-pairs comparison tests (54) using the PMCMRplus package (55) within **R** (56).

Because our study investigated reproducibility in P_{rate} and P_{titer}, our analysis focused on comparing the relative variance—rather than the magnitudes—of P_{rate} and P_{titer} for each workflow. We assessed the relative variance for each combination of DNA preparation workflow and feature P using a random-effects model with effects for the biological replicate, technical replicate and measurement replicate. Let P_{ijkl} denote the value of P for measurement replicate l (l = 1, 2, 3) from technical replicate k (k = 1, 2, 3, 4, 5) and biological replicate j (j = 1, 2, 3, 4, 5) prepared using workflow i (i = 1,…, 8). To account for the observed differences in the magnitude of the fluorescence intensity measurements between workflows, we normalized the collection of P values for each workflow by the median value of P from that workflow, i.e. $P^*_{ijkl} = \frac{P_{ijkl}}{m_i}$, where m_i is the median of P_{ijkl} across j, k, and l. For workflow i, the following random-effects model was fit to the normalized feature values: $P^*_{ijkl} = \mu_i + \alpha_{ij} + \beta_{ijk} + \epsilon_{ijkl}$, where μ_i represents the mean median-normalized value of P for workflow i, α_{ij} is a random effect corresponding to biological replicate j, β_{ijk} is a random effect corresponding to technical replicate k within biological replicate j and ϵ_{ijkl} is the random error associated with measurement replicate l from CFE reactions containing DNA technical replicate k from biological replicate j. We assumed all random effects α_{ij}, β_{ijk}, and ϵ_{ijkl} were mutually independent and distributed normally with mean 0 and variances $\sigma^2_{\alpha i}, \sigma^2_{\beta i}$ and $\sigma^2_{\epsilon i}$, respectively. $\sigma^2_{\alpha i}, \sigma^2_{\beta i}$ and $\sigma^2_{\epsilon i}$ represent variance associated with the biological replicate of the plasmid DNA, the technical replicate of DNA template prepared with a given DNA preparation workflow and measurement replicate of nominally identical CFE reactions containing the DNA prepared with the same DNA preparation workflow, respectively. Parameters for this random-effects model were estimated using the ***lme4*** (57) package within **R** (56). Ninety-five percentage confidence intervals were then evaluated for each mean and variance in the fitted model using profile likelihoods.

To test whether observed differences in variance estimates across DNA preparation workflows were statistically significant, we used drop-in-deviance tests. These tests used log-likelihoods as a measure of goodness-of-fit and changes in the log-likelihood under different constraints on model parameters as a measure of statistical significance. For example, to assess whether the biological variance estimates for workflows i and i', namely $\sigma^2_{\alpha i}$ and $\sigma^2_{\alpha i'}$, respectively, differ with statistical significance, the goodness-of-fit for a model in which each workflow took on its own value for the biological variance was compared to the goodness-of-fit for a model in which both workflows shared the same biological variance. To accomplish this, a random-effects model was fit to the complete collection of P^*_{ijkl} and $P^*_{i'jkl}$ (across all values of j, k and l) values from both workflows subject to the constraint that $\sigma^2_{\alpha i}$ and $\sigma^2_{\alpha i'}$ must be the same (say σ^2_{α}). Within this combined model, all random effects were treated as mutually independent, and no effect was shared across workflows. The model was fit by optimizing the likelihood across the model parameters, namely σ^2_{α}, μ_i and $\mu_{i'}$, $\sigma^2_{\beta i}$ and $\sigma^2_{\beta i'}$ and $\sigma^2_{\epsilon i}$ and $\sigma^2_{\epsilon i'}$. $l_{i,i'}$ represents the maximized log-likelihood obtained from this fitting process. l_i and $l_{i'}$ denote the respective log-likelihoods produced by the fits of the random-effects model to the normalized feature values from workflows i and i', respectively.

The constrained space in which $\sigma^2_{\alpha i}$ and $\sigma^2_{\alpha i'}$ are equal is a subspace of the full parameter space of μ_i and $\mu_{i'}$, $\sigma^2_{\alpha i}$ and $\sigma^2_{\alpha i'}$, $\sigma^2_{\beta i}$ and $\sigma^2_{\beta i'}$ and $\sigma^2_{\epsilon i}$ and $\sigma^2_{\epsilon i'}$. We therefore calculated a drop-in-deviance test statistic of $2(l_i + l_{i'} - l_{i,i'})$ and found the proportion of a chi-squared distribution with one degree of freedom that lies above this value to construct a P-value. With eight different workflows, there are 28 possible pairwise comparisons between workflows. We therefore applied the Bonferroni test to control for multiple testing (58). The observed difference between the parameter estimates was therefore declared statistically significant if this proportion was <0.05/28. We also applied this approach to assess the statistical significance of observed differences across workflows in estimates for the technical ($\sigma^2_{\beta i}$) and measurement variances ($\sigma^2_{\epsilon i}$).

To demonstrate a practical outcome of the observed differences in the variance between the DNA preparation workflows, we estimated the number of technical replicates needed for each DNA preparation workflow that when pooled physically match the variability achieved for only a single technical replicate from the workflow with the lowest overall variance, MFP-. In this analysis, we assumed that the variance in P values for a pooled sample corresponds to the variance of the mean P value among

the replicates being pooled. For example, if P values for Replicates 1 and 2 have respective variances σ_1^2 and σ_2^2, we assumed that P for a pooled sample formed by combining Replicates 1 and 2 is $\frac{\sigma_1^2+\sigma_2^2}{4}$. We considered hypothetical scenarios, where each technical replicate from a biological replicate was used for 1, 2, 3, 5 or 10 measurement replicates. The reproducibility for DNA preparation workflow i was quantified as the inverse variance of the mean calculated over the measurement replicates. That is, when n measurement replicates are prepared using workflow i for each technical replicate, the reproducibility is calculated as $r_{i,n} = \left(\sigma_{\alpha i}^2 + \frac{\sigma_{\epsilon i}^2}{n}\right)^{-1}$ [Equation 1]. The ratio $R(i,i') = \frac{r_{i,n}}{r_{i',n}}$ conveys the number of technical replicates prepared by DNA preparation workflow i, each with n measurement replicates, needed to achieve the same expected variability in P for a given biological replicate as one technical replicate prepared using DNA preparation workflow i'. Because $R(i,i') = \frac{R(i,MFP\_)}{R(i',MFP\_)}$, we summarize all pairwise ratios by only reporting $R(i,MFP\_)$ for each DNA preparation workflow. Uncertainty bounds for $R(i,MFP\_)$ were evaluated with 1000 iterations of a bootstrapping process. For each DNA preparation workflow, each bootstrapping iteration consisted of randomly sampling five numbers from one to five with replacement, which represents the indices of j resulting from a random resampling of the biological replicates. Then, for each of the sampled j indices, five numbers from one to five were sampled randomly with replacement, representing the indices of k from a random resampling of technical replicates within the randomly sampled biological replicate. Finally, for each of the five sampled technical replicate indices k within each sampled biological replicate, three numbers from one to three were randomly sampled with replacement, representing the indices of l resulting from random resampling of measurement replicates within the selected combination of biological and technical replicates. This bootstrapping iteration produces 5*5*3 = 75 index sets of (j,k,l). These index sets corresponded to median-normalized feature values P^*_{ijkl} with the same indices of (j,k,l), to produce a set of resampled values that mimic the potential correlation structure of the experimental measurements. The resampled data produced in bootstrap iteration m was used to refit the random-effects model, and the resulting variance estimates then gave the corresponding reproducibility $r_{i,n}^{(m)}$ according to Equation 1. Having obtained $r_{i,n}^{(m)}$ for each DNA preparation workflow, we calculated the corresponding values of $R(i,MFP\_)^{(m)}$. Upper and lower bounds for the confidence intervals of $R(i,MFP\_)$ were calculated as the 0.975 and 0.025 quantiles, respectively, from the set $\left\{R(i,MFP\_)^{(m)}\right\}_{m=1}^{1000}$. To assess the statistical significance on the observed difference between $r_{i,n}$ and $r_{i',n}$, we calculated the test statistic $z = \frac{log\left(r_{i,n}\right) - log\left(r_{i',n}\right)}{\sqrt{var\left(log\left(r_{i,n}\right)\right) + var\left(log\left(r_{i',n}\right)\right)}}$, where we estimated the values of $var\left(log\left(r_{i,n}\right)\right)$ and $var\left(log\left(r_{i',n}\right)\right)$ using the sample variances of $\left\{log\left(r_{i,n}^{(m)}\right)\right\}_{m=1}^{1000}$ and $\left\{log\left(r_{i',n}^{(m)}\right)\right\}_{m=1}^{1000}$, respectively. We considered the observed difference between $r_{i,n}$ and $r_{i',n}$ statistically significant if the proportion of a standard normal distribution that lies farther from 0 than z was <0.05/28, using the Bonferroni correction for multiple testing.

For each CFE reaction, eGFP calibration wells were used to calculate a range of molar concentrations of eGFP produced from fluorescence intensity measurements in arbitrary units. Calibration curves relating fluorescence intensity to eGFP concentration were estimated using eGFP calibration samples of 0, 5, 10 and 30 μmol/l. Each CFE plate contained nine series of eGFP calibration samples across four nominal concentrations of eGFP (SI Figure S11), for a total of 27 calibration series across all three CFE plates. A least-squares linear regression fit to the mean calibration measurements did not adequately represent the fluorescence intensity measurements for calibration samples with the same nominal eGFP concentration. Therefore, we considered all linear regression fits at each time point to estimate a range of eGFP concentration for each measured fluorescence intensity using the following approach. Let the estimated intercept and slope resulting from least-squares linear fit to the measured fluorescence intensity versus eGFP concentration data for calibration series c at time t be given by I_{tc} andS_{tc}, respectively. Let F_{ijklt} denote the measured fluorescence intensity P at time t for measurement replicate l (l = 1, 2, 3) from technical replicate k (k = 1, 2, 3, 4, 5) and biological replicate j (j = 1, 2, 3, 4, 5), prepared using workflow i (i = 1,..., 8). The range of eGFP concentration associated with F_{ijklt} is given by $\min_c\left(\frac{F_{ijklt}-I_{tc}}{S_{tc}}\right)$ to $\max_c\left(\frac{F_{ijklt}-I_{tc}}{S_{tc}}\right)$, where c ranges from 1 to 27. Because the variability in the eGFP calibration samples did not exhibit a clear relationship with either region of CFE plate or CFE plate number, we assumed that the appropriate conversion from fluorescence intensity to molar concentration was similar across region and plate.

3. Results and discussion

Toward improving reproducibility in cell-free protein production, we characterized the effects of two common DNA extraction methodologies, a postprocessing step and manual versus automated preparation on eGFP titer and rate of eGFP produced by the myTXTL CFE system.

To span the range of common preparation methods, we investigated eight DNA preparation workflows (SI Table S2): manual filter extracted non-purified (MFP-), manual filter extracted purified (MFP_+), manual magnetic bead extracted non-purified (MBP-), manual magnetic bead extracted purified (MBP_+), automated filter extracted non-purified (AFP-), automated filter extracted purified (AFP_+), automated magnetic bead extracted non-purified (ABP-) and automated magnetic bead extracted purified (ABP_+). We examined the concentration of DNA in solution, the physical damage to the DNA template and the purity of the DNA resulting from each DNA preparation workflow.

Regarding DNA concentration, manual DNA preparation workflows yielded more concentrated DNA solutions than automated workflows (Figure 2). The mean concentrations of DNA solutions prepared with MFP-, MFP_+, MBP- and MBP_+ workflows and measured with fluorometry were 84.2 nmol/l, 89.6 nmol/l, 59.9 nmol/l and 63.9 nmol/l, respectively, as compared to 42.5 nmol/l, 34.5 nmol/l, 4.43 nmol/l and 4.64 nmol/l with AFP-, AFP_+, ABP and ABP_+ workflows, respectively, in ∼50 μl of solution. The myTXTL protocol calls for at least 5 nmol/l of DNA in 12 μl of reaction, where the volume of DNA solution accounts for 25% of the reaction volume. Therefore, we considered each workflow successful, if it produced a minimum of 20 nmol/l intact and biologically functional DNA in 50 μl of solution.

Differences in DNA concentration resulting from manual and automated workflows may have arisen from the use of centrifugation in manual workflows versus a positive pressure manifold in automated workflows for clearing the lysate and eluting the DNA. Centrifuging the lysate to clear it of cell debris resulted in a tightly formed pellet of debris at the bottom of the microcentrifuge tube, separate from the surrounding lysate. In contrast, pushing the lysate through a filter plate using positive pressure led to the retention of cell debris by the filter in each well, along

with ~50–75 μl of cleared lysate that was effectively lost. Eluting DNA with the nuclease-free water from the spin column with centrifugation allowed the complete recovery of the entire volume added. However, using positive pressure to push nuclease-free water through a binding plate led to the retention of ~50–75 μl of nuclease-free water by the filter in each well, along with any DNA still bound to membrane. Because a minimum of 50 μl of DNA solution was required from each DNA preparation workflow, a greater volume of nuclease-free water was used to elute DNA from automated workflows than from manual workflows. This diluted the DNA solutions prepared with automated workflows as compared to those prepared with manual workflows.

Workflows using filter-based extraction (MFP-, MFP_+, AFP- and AFP_+) yielded more concentrated DNA solutions than workflows using bead-based extraction (MBP-, MBP_+, ABP- and ABP_+), presumably due to differences in the number of DNA manipulation steps (Figure 2) between these workflows. Specifically, MBP-, MBP_+, ABP- and ABP_+ workflows had four times as many DNA wash steps as MFP-, MFP_+, AFP- and AFP_+ workflows. Each additional manipulation increased the chances that some portion of the DNA would be lost.

Regarding DNA quality as related to contaminants in the DNA solution, none of the DNA preparation workflows consistently yielded DNA solutions with A_{260}/A_{230} and A_{260}/A_{280} ratios within community-accepted values (Figure 3 and SI Figure S12), even those workflows that included a postprocessing step in the form of a DNA purification kit. Characterizing and reducing contaminants in the DNA solution mitigates their potentially confounding effects on eGFP production. First, solutions containing <8.60 nmol/l of DNA as measured with UV spectrophotometry have been found to give unreliable A_{260}/A_{230} and A_{260}/A_{280} results (42), which includes DNA solutions prepared with ABP- and ABP_+ workflows. However, it is important to note that significant carry-over between DNA wash and elution steps occurred in ABP- and ABP_+ workflows, due to the limitations of the labware compatible with the magnetic separation modules used for these workflows. We found that in plates with round-bottom wells, magnetic beads formed a wide ring in each well leaving sufficient space for aspirating supernatant with pipette tips without disturbing the beads. The round shape of the wells also prevented the pipette tips from aspirating all the liquid from each well, leaving a few drops of wash buffer or DNA solution in each well. This likely led to cross-contamination between DNA wash steps, where traces of supernatant containing genomic DNA, proteins and endotoxins eluted with the plasmid DNA (SI Figure S12). Second, in workflows that used bead-based extraction (MBP-, MBP_+, ABP- and ABP_+) as opposed to filter-based extraction (MFP-, MFP_+, AFP- and AFP_+), incomplete neutralization of the cell lysate likely led to the presence of denatured proteins, genomic DNA, polysaccharides and cellular debris in the eluted DNA solution. To neutralize 500 μl of lysed cells, MBP-, MBP_+, ABP- and ABP_+ workflows used 125 μl of neutralization buffer, as opposed to 350 μl of buffer used in MFP-, MFP_+, AFP- and AFP_+ workflows. As a result, cell lysates in MBP-, MBP_+, ABP- and ABP_+ workflows were cloudy, indicating at least some denatured proteins, genomic DNA and cellular debris from the lysis step did not precipitate in salt–detergent complexes, as is typical for completely neutralized lysates. Although we could have adjusted the volume of neutralization buffer to be identical in magnetic-bead-based and filter-based extraction methods, we chose to follow the protocol specific to each kit manual and report our findings accordingly. Third, additional sample handling associated with including a postprocessing step following DNA extraction in MFP_+, MBP_+, AFP_+ and ABP_+ workflows generally did not improve the purity of the resulting DNA solution to community-accepted values. Only the MBP_+ workflow showed notable improvement from postprocessing.

Regarding the quality of the DNA, as related to physical damage to the plasmid itself, DNA templates prepared with automated

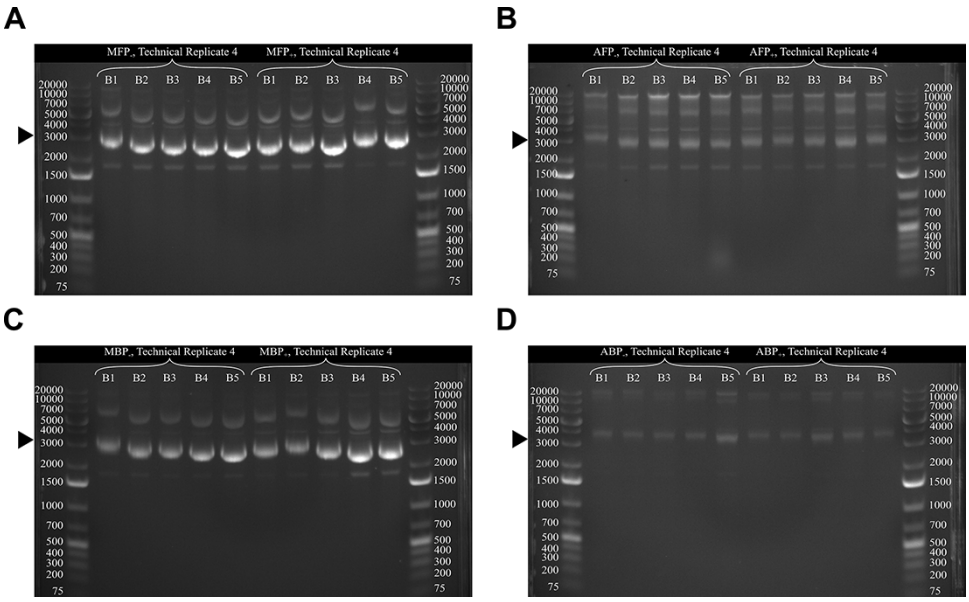

Figure 4. Representative gels of physical damage to DNA templates. DNA solutions from each DNA preparation workflow were separated by size using 1.2% pre-poured agarose gels. Lanes 1 and 12 contain 2 μg of 0.25–10.0 kb premixed DNA ladder. The molecular sizes in the DNA ladder are noted in white in number of base pairs. The black triangle indicates the expected location of 3000-bp DNA, the approximate size of intact pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 plasmid DNA. Lanes 2 through 6 contain DNA solutions from Technical Replicate 4 of each biological replicate prepared with (A) MFP-, (B) AFP-, (C) MBP- and (D) ABP- workflows. Lanes 7 through 11 contain DNA solutions from Technical Replicate 4 of each biological replicate prepared with (A) MFP_+, (B) AFP_+, (C) MBP_+ and (D) ABP_+ workflows. Significant physical damage is present in all DNA solutions prepared with automated workflows.

workflows (AFP-, AFP_+, ABP- and ABP_+) showed clear evidence of damage and fragmentation, while those prepared with manual workflows (MFP-, MFP_+, MBP- and MBP_+) remained largely intact (Figure 4). Bands in the agarose gel of the same apparent size as 3000 bp in the DNA ladder indicated intact pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 plasmid DNA. Bands of the same apparent size as 15 000, 4000 and 2000 bp in the DNA ladder indicated a distribution of DNA lengths in the prepared DNA solutions, ranging from co-extracted genomic DNA to short fragments of DNA damaged by excessive agitation. Assessing the fraction of intact DNA template in the CFE reaction can help ensure that all DNA added to the reaction is biologically active. Differences in physical damage between DNA templates prepared with automated and manual workflows may have arisen from how cells were mixed after the addition of lysis and neutralization buffers in these workflows. In AFP-, AFP_+, ABP- and ABP_+ workflows, cells were mixed by shaking after the addition of lysis buffer and by shaking and pipetting after the addition of neutralization buffer. In MFP-, MFP_+, MBP- and MBP_+ workflows, cells were mixed by slowly inverting the tubes after the addition of lysis and neutralization buffers. While the automated method used widebore pipette tips and reduced speeds for both aspiration and dispensing, we hypothesize that mixing by pipetting was aggressive enough to physically damage the DNA. Differences in physical damage between DNA templates prepared with automated bead-based (ABP- and ABP_+) workflows and manual bead-based (MBP- and MBP_+) workflows may have arisen from how beads were resuspended during wash and DNA elution steps. In ABP- and ABP_+ workflows, a combination of pipetting and shaking was used to resuspend the beads after the addition of ETR wash, VHB wash, SPM buffers and nuclease-free water. Aspiration and dispensing speeds were increased to as much as 450 µl/s, for up to 10 cycles to ensure complete resuspension. In MBP- and MBP_+ workflows, beads were resuspended after the addition of ETR wash, VHB wash, SPM buffers and nuclease-free water by briefly vortexing each tube for up to 6 s. The aggressive agitation of the samples during wash and DNA elution steps in ABP- and ABP_+ workflows likely led to damaged and fragmented DNA templates, while the templates prepared with MBP- and MBP_+ workflows remained largely intact.

In general, CFE reactions were composed of 27 µl Sigma 70 Master Mix and 5 nmol/l DNA per 36 µl of total sample volume. CFE reactions containing DNA templates prepared with ABP- and ABP_+ workflows contained a variable concentration of DNA <5 nmol/L, due to the insufficient DNA concentrations yielded by these workflows. As a result, these CFE reactions contained a variable amount of DNA, <5 nmol/L, and the fluorescence intensity measurements from eGFP produced by these reactions were scaled by the DNA concentration added to each reaction to facilitate comparison across all CFE reactions.

In each CFE reaction, we analyzed eGFP titer (P_{titer}) and the rate of eGFP production (P_{rate}) as two common metrics for a biomanufacturing process. Titer refers to the concentration of protein produced and is a benchmark for the upstream efficiency of a biomanufacturing process (59). Rate corresponds to the concentration of protein produced per unit time and is a measure of productivity of a biomanufacturing process (60).

Because of substantial and unexpected variability between different series of eGFP calibration samples, we chose to analyze the effects of eight DNA preparation workflows on P_{titer}, P_{rate}, variance in P_{titer} and variance in P_{rate} in units of arbitrary fluorescence intensity instead of units of molar concentration. Each CFE plate contained nine series of eGFP calibration samples across four nominal concentrations of eGFP (SI Figure S11) measured every 5 min for 24 h. For each time point, a linear regression model was fit to the observed intensity versus nominal concentration for each of the calibration samples. The slopes and intercepts resulting from these fits displayed substantial variability between calibration samples with the same nominal eGFP concentration and clear trends across time (SI Figure S14–S15). This variability would not be adequately represented by choosing a single slope and intercept for mapping arbitrary intensity units to molar protein concentration. The source of this variability is unknown. Additionally, the residuals at each eGFP calibration concentration strongly indicated a lack of linearity between eGFP concentration and fluorescence intensity (SI Figure S15). These residuals mostly suggested a concave rather than linear relationship, but the nonlinearity was not reproduced even within the same region and plate. To facilitate comparison with other studies, we report ranges of estimated molar concentration of eGFP produced by each CFE reaction (SI Figures S16–S17).

To compare P_{titer}, P_{rate}, variance in P_{titer} and variance in P_{rate} across all CFE reactions, we performed instrument calibration (SI Figures S2–S4) ahead of each CFE measurement and measured three types of control samples on each CFE plate. Daily instrument calibration with internal standards (see Materials and Methods) ruled out systematic bias in the fluorescence measurements across each CFE plate (SI Figures S8–S10). The fluorescence intensity from all wells of the blank 384-well plate, the nuclease-free water plate and fluorescein plate across all measurements and days varied <10% of standard deviation of the mean (SI Table S4). For each CFE plate, visual analysis of the fluorescence of the negative control CFE reactions, nuclease-free water blank samples and PBS blank samples (SI Figure S18) indicated that the variability due to difference across plates or position within each plate was not substantial. For nuclease-free water and PBS blank samples, the total observed variability was less negligible compared to the fluorescence intensities of the CFE reactions (SI Figure S18). The fluorescence intensity from wells containing the negative control CFE reactions tended to increase slightly with time (SI Figure S18). However, the observed differences did not appear to be strongly associated with either plate number or region within the plate.

For each DNA preparation workflow, we calculated mean P_{titer} and mean P_{rate} from five biological replicates. SI Tables S5 and S6 provide the ratio of mean P_{titer} and ratio of mean P_{rate}, respectively, for each pair of workflows from Nemenyi's all-pairs comparison tests. P values from a Friedman test were used to identify statistically significant differences in mean P_{titer} and mean P_{rate} between workflows. We observed significant and substantial differences in the magnitudes of P_{titer} and P_{rate} across DNA preparation workflows (SI Tables S5 and S6). However, because the primary focus of this study lay in measuring the effects of DNA template preparation on variability in CFE, all subsequent analyses characterized variability in P_{titer} and P_{rate}, rather than magnitude, for each DNA preparation workflow.

To characterize variability in P_{titer} and P_{rate} for each DNA preparation workflow, we used a random-effects model to assess relative variance associated with 5 biological replicates of DNA, 25 technical replicates of DNA prepared with the same DNA preparation workflow and 75 measurement replicates of nominally identical CFE reactions containing DNA prepared with the same DNA preparation workflow (see Materials and Methods). From drop-in-deviance tests, we determined that relative variance associated with biological replicate did not exhibit statistically significant differences among the workflows for either P_{titer} or P_{rate} (see Materials and Methods, Figure 5, and SI Figure S19).

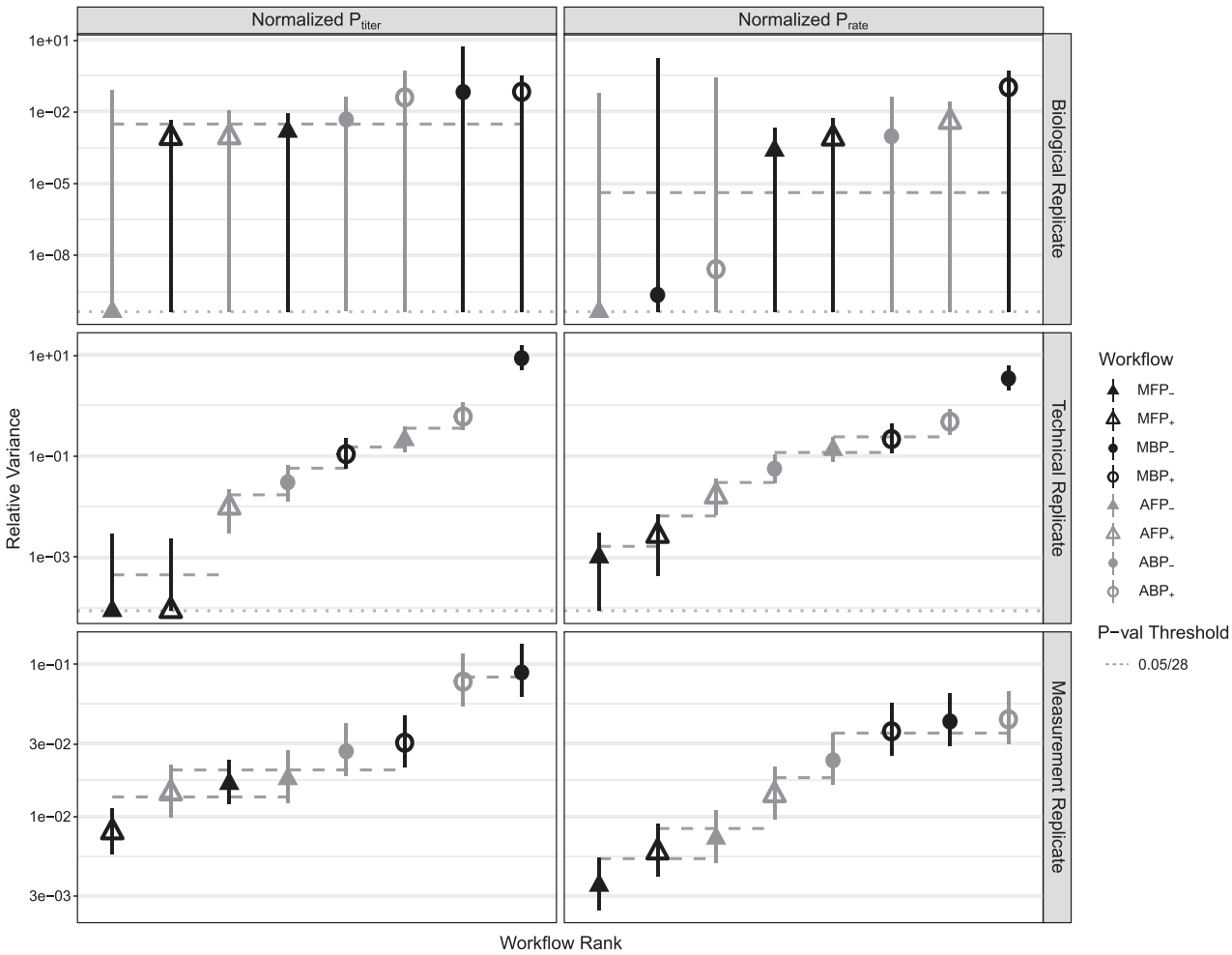

Figure 5. Ranked variance analysis in P_{titer} and P_{rate}. Relative variance estimates associated with the biological replicate of the DNA template (top panel), technical replicate of the DNA prepared with the same DNA preparation workflow (middle panel) and measurement replicate of nominally identical CFE reactions containing DNA prepared with the same DNA preparation workflow (bottom panel) calculated using a random-effects model. Estimates connected by a horizontal dashed line are not statistically significantly different according to the Bonferroni corrected *P*-value threshold of 0.05/28. Error bars provide 95% confidence intervals evaluated using profile likelihoods. Large error bars in the top panel result from the small sample size (five) of biological replicates for each experimental workflow. We show a horizontal dotted line near the bottom of each top and middle panels to represent a relative variance of zero. Technical replicates were the most significant source of variance, while the relative variance associated with biological replicate was not statistically different from zero for all DNA preparation workflows.

Additionally, for all DNA preparation workflows, the relative variance associated with biological replicate in both P_{titer} and P_{rate} was not statistically different from zero. The large error bars in these variance estimates are a consequence of the small sample size (five) of biological replicates for each DNA preparation workflow. A more rigorous study of the effects of biological replicate associated with the DNA template on protein production could be achieved by including more biological replicates for each DNA preparation workflow. Because the preparation of biological replicates was separate from the DNA preparation workflows, we excluded biological variance from further consideration in comparisons of variability in P_{titer} and P_{rate} resulting from different DNA workflows. Variance estimates corresponding to technical replicates differed from zero with statistical significance for each of the DNA preparation workflows for P_{rate} and for six of the eight DNA preparation workflows for P_{titer} (Figure 5). The variance estimates associated with technical replicates were non-zero for all DNA preparation workflows; the 25 technical replicates in our study were sufficient to detect differences in the variance of P_{titer} and P_{rate} between DNA preparation workflows, whereas five biological replicates were not. We encourage researchers to choose a sufficient number of each type of replicate to characterize variance estimates with enough confidence for the needs of their target application. CFE reactions that exhibited low relative variance associated with technical replicates also exhibited low relative variance associated with measurement replicates, for example, CFE reactions containing DNA prepared with manual filter-based (MFP- and MFP_+) workflows as compared to manual bead-based (MBP- and MBP_+) workflows. A large variance estimate among measurement replicates indicates an increased benefit to taking more replicate measurements of a CFE reaction containing DNA prepared with the same DNA preparation workflow, for example, CFE reactions containing DNA prepared with bead-based (MBP-, MBP_+, ABP- and ABP_+) workflows.

CFE reactions containing DNA prepared with automated (AFP-, AFP_+, ABP- and ABP_+) workflows generally had larger variance in P_{titer} and P_{rate} than CFE reactions containing DNA prepared with manual (MFP-, MFP_+, MBP- and MBP_+) workflows (Figure 5). For filter-based workflows, although CFE reactions containing DNA prepared with automated (AFP- and AFP_+) workflows appear more

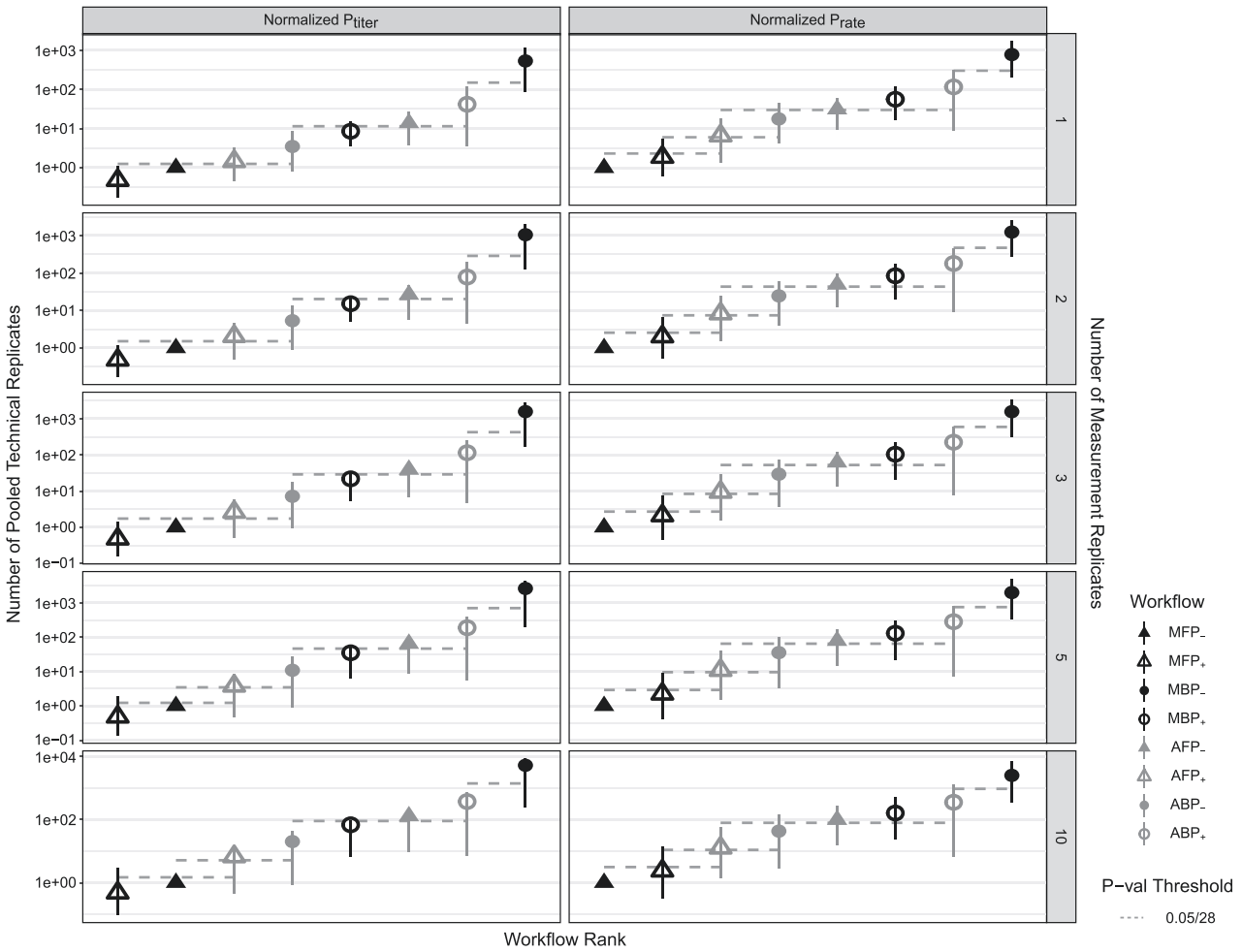

Figure 6. Experimental design needed for each DNA preparation workflow to achieve the estimated relative variance in P_{titer} and P_{rate} from the least variable workflow. DNA workflows are compared based on how many technical replicates from one workflow would need to be pooled to achieve the same variance in P_{titer} and P_{rate} as one technical replicate from another workflow. The number of replicate measurements taken for each pool of technical replicates is given at the right of each panel row. Estimates connected by a horizontal dashed line are not statistically significantly different according to the Bonferroni corrected *P*-value threshold of 0.05/28. Error bars provide 95% confidence intervals evaluated using nonparametric bootstrapping. This analysis provides practical guidance on the number of pooled technical replicates required to achieve the same variance in P_{titer} and P_{rate} as CFE reactions containing DNA prepared with the least variable (MFP-) workflow.

variable in P_{titer} and P_{rate} than CFE reactions containing DNA prepared with manual (MFP- and MFP_+) workflows, this difference was not statistically significant. The performance of the bead-based (MBP-, MBP_+, ABP- and ABP_+) workflows was more difficult to evaluate, because fluorescence measurements from CFE reactions containing DNA prepared with ABP- and ABP_+ workflows were scaled to account for the variable DNA concentration in those reactions (see Materials and Methods). Bead-based workflows generally had larger variance from both technical and measurement replicates.

To interpret practical implications of the observed differences in variance in P_{titer} and P_{rate} between DNA preparation workflows, we compared the workflows based on how many technical replicates from one workflow would need to be pooled to achieve the same variance as one technical replicate from another workflow (see Materials and Methods). The MFP- workflow was chosen as a baseline for comparison due to its low variability for both P_{titer} and P_{rate} (Figure 6). We determined that the MBP- workflow was the most variable overall in terms of P_{titer} and P_{rate} and was estimated to require the most technical replicates compared to other DNA preparation workflows. In general, bead-based workflows required more technical replicates than filter-based workflows to achieve the same variance as MFP-. The number of technical replicates required for the AFP- and AFP_+ workflows was not statistically different, indicating that the performance of these workflows is comparable in terms of variance in P_{titer} and P_{rate} and the postprocessing for these workflows did not confer an obvious benefit. The automated DNA preparation workflows often required more technical replicates than their manual counterparts, indicating that a skilled technician may perform higher-quality work than a high-throughput automated workflow. However, the greater throughput offered by automation may outweigh disadvantages associated with performing additional technical replicates to reduce the overall variance.

Numerous factors may be considered to further improve the automated DNA preparation workflows for CFE applications. The automated workflows in this study combine individual sub-operations performed in order, such as (i) aliquoting of large liquid cultures into technical replicates, (ii) removal of the supernatant from spun-down cell aliquots, (iii) resuspension of the cells in buffer, (iv) DNA extraction, (v) DNA solution postprocessing and (vi) DNA quantitation. Each sub-operation and its accompanying

equipment need to be optimized individually for a given automated workflow. For all automated DNA preparation workflows, sample mixing upon addition of the lysis and neutralization buffers could be better characterized, for example, by using a colorimetric dye to visualize sample mixing. For lysate clearing, we found using a centrifuge to be superior to using a positive pressure manifold in terms of the volume and clarity of the lysate achieved. If using a positive pressure manifold, further optimization of the pressure and duration parameters as well as a more deliberate sampling of filter plates including a variety of filter materials, pore sizes and plate nozzle dimensions could further improve the lysate clearing step. For magnetic-bead-based extraction workflows, optimizing the volumetric ratio of the neutralization buffer and the lysis buffer to cell volume could improve the lysate quality. Optimizing the resuspension of the magnetic beads after the addition of each wash buffer to reduce physical damage to the DNA template due to vigorous pipetting could improve the quality of the resulting DNA solution. Similarly, testing a variety of magnetic separation racks and complementary labware to reduce sample carryover during DNA wash and elution steps could reduce contamination in the DNA solution. Pipetting volumes <10 μl using traditional positive pressure displacement liquid handlers should be avoided, because the accuracy and precision of each pipetting step can vary up to 150% (40).

4. Conclusion

To understand and reduce variability in the performance of CFE, the concentration of the DNA template, the quality of the DNA template in terms of physical damage and the quality of the DNA solution in terms of purity should be characterized and controlled. Beyond the concentration of DNA in solution, further characterization of the quality—and by extension the biological activity—of that DNA would give insight into the expected performance of CFE using that prepared DNA template. For example, size and topology determination using established separation techniques, such as capillary electrophoresis, atomic force microscopy or single molecule methods, would improve on typical measurements at the state of the art. This will likely become increasingly important for applications of CFE that rely on measurements of single molecules, such as may be used to characterize CFE reactions or for very-small-volume or fluidic devices (61).

Based on our results, we encourage researchers to test the effects of DNA template preparation workflows within the context and needs of their target application, for example, using the analysis of variability presented here. Our study could be repeated to characterize other DNA preparation workflows or CFE reaction systems. We caution that different technicians may perform manual work of varying quality that may also change with specific equipment and training. Automation removes the human operator but may introduce other potentially larger sources of variability that must be similarly characterized and understood. When selecting a DNA preparation workflow fit for a particular CFE application, the speed of DNA extraction, sample throughput, and reproducibility requirements in P_{titer} and P_{rate} should be considered. Once the workflow to prepare the DNA template is isolated as a source of variability, other aspects of the reaction may be explored to ensure that the performance of the CFE reaction meets specifications for the application at hand.

Supplementary data

Supplementary data are available at SYNBIO online.

Materials and data availability

Automation methods are available at protocols.io (dx.doi.org/10.17504/protocols.io.j8nlkk8rdl5r/v1). Analysis code is available in the Supplemental Information.

Acknowledgments

We would like to thank Matthew Lux, William Poole and Vincent Noireaux for thoughtful discussion during the planning of this work. We would like to thank Doug Densmore for insights into the automation workflow.

Author contribution

E.F.R. and E.A.S. conceived the study and developed the experimental workflow. E.F.R. and N.A. performed DNA preparation and quantitation experiments. E.F.R. and D.R. programmed automated workflows. E.F.R. performed cell-free experiments. S.L and E.A.S. performed data analysis. All authors wrote the manuscript.

Conflict of interest statement. The authors declare that they have no conflict of interest.

Funding

This work was supported by the National Institute of Standards and Technology.

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.

References

1. Pardee,K., Slomovic,S., Nguyen,P.Q., Lee,J.W., Donghia,N., Burrill,D., Ferrante,T., McSorley,F.R., Furuta,Y., Vernet,A. *et al.* (2016) Portable, on-demand biomolecular manufacturing. *Cell*, **167**, 248–259 e212. 10.1016/j.cell.2016.09.013.
2. Hunt,J.P., Yang,S.O., Wilding,K.M. and Bundy,B.C. (2017) The growing impact of lyophilized cell-free protein expression systems. *Bioengineered*, **8**, 325–330. 10.1080/21655979.2016.1241925.
3. Adiga,R., Al-Adhami,M., Andar,A., Borhani,S., Brown,S., Burgenson,D., Cooper,M.A., Deldari,S., Frey,D.D., Ge,X. *et al.* (2018) Point-of-care production of therapeutic proteins of good-manufacturing-practice quality. *Nat. Biomed. Eng.*, **2**, 675–686. 10.1038/s41551-018-0259-1.
4. Karig,D.K., Bessling,S., Thielen,P., Zhang,S. and Wolfe,J. (2017) Preservation of protein expression systems at elevated temperatures for portable therapeutic production. *J. R. Soc. Interface*, **14**, 129. 10.1098/rsif.2016.1039.
5. Timm,A.C., Shankles,P.G., Foster,C.M., Doktycz,M.J. and Retterer,S.T. (2016) Toward microfluidic reactors for cell-free protein synthesis at the point-of-care. *Small*, **12**, 810–817. 10.1002/smll.201502764.
6. Martin,R.W., Majewska,N.I., Chen,C.X., Albanetti,T.E., Jimenez,R.B.C., Schmelzer,A.E., Jewett,M.C. and Roy,V. (2017) Development of a CHO-based cell-free platform for synthesis of active monoclonal antibodies. *ACS Synth. Biol.*, **6**, 1370–1379. 10.1021/acssynbio.7b00001.

7. Min,S.E., Lee,K.H., Park,S.W., Yoo,T.H., Oh,C.H., Park,J.H., Yang,S.Y., Kim,Y.S. and Kim,D.M. (2016) Cell-free production and streamlined assay of cytosol-penetrating antibodies. *Biotechnol. Bioeng.*, **113**, 2107–2112. 10.1002/bit.25985.
8. Stech,M. and Kubick,S. (2015) Cell-free synthesis meets antibody production: a review. *Antibodies*, **4**, 12–33. 10.3390/antib4010012.
9. Thangavelu,U., Arumugam,D.I., Takashima,E., Tachibana,M., Ishino,T., Torii,M. and Tsuboi,T. (2014) Application of wheat germ cell-free protein expression system for novel malaria vaccine candidate discovery. *Expert Rev. Vaccines*, **13**, 75–85. 10.1586/14760584.2014.861747.
10. Welsh,J.P., Lu,Y., He,X.S., Greenberg,H.B. and Swartz,J.R. (2012) Cell-free production of trimeric influenza hemagglutinin head domain proteins as vaccine antigens. *Biotechnol. Bioeng.*, **109**, 2962–2969. 10.1002/bit.24581.
11. Kanter,G., Yang,J., Voloshin,A., Levy,S., Swartz,J.R. and Levy,R. (2007) Cell-free production of scFv fusion proteins: an efficient approach for personalized lymphoma vaccines. *Blood*, **109**, 3393–3399. 10.1182/blood-2006-07-030593.
12. Chappell,J., Jensen,K. and Freemont,P.S. (2013) Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res.*, **41**, 3471–3481. 10.1093/nar/gkt052.
13. Chappell,J., Westbrook,A., Verosloff,M. and Lucks,J.B. (2017) Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. *Nat. Commun.*, **8**, 1051. 10.1038/s41467-017-01082-6.
14. Hu,C.Y., Varner,J.D. and Lucks,J.B. (2015) Generating effective models and parameters for RNA genetic circuits. *ACS Synth. Biol.*, **4**, 914–926. 10.1021/acssynbio.5b00077.
15. Karim,A.S. and Jewett,M.C. (2016) A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. *Metab. Eng.*, **36**, 116–126. 10.1016/j.ymben.2016.03.002.
16. Niederholtmeyer,H., Sun,Z.Z., Hori,Y., Yeung,E., Verpoorte,A., Murray,R.M. and Maerkl,S.J. (2015) Rapid cell-free forward engineering of novel genetic ring oscillators. *Elife*, **4**, e09771. 10.7554/eLife.09771.
17. Takahashi,M.K., Chappell,J., Hayes,C.A., Sun,Z.Z., Kim,J., Singhal,V., Spring,K.J., Al-Khabouri,S., Fall,C.P., Noireaux,V. *et al.* (2015) Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. *ACS Synth. Biol.*, **4**, 503–515. 10.1021/sb400206c.
18. Chemla,Y., Ozer,E., Schlesinger,O., Noireaux,V. and Alfonta,L. (2015) Genetically expanded cell-free protein synthesis using endogenous pyrrolysyl orthogonal translation system. *Biotechnol. Bioeng.*, **112**, 1663–1672. 10.1002/bit.25587.
19. Iwane,Y., Hitomi,A., Murakami,H., Katoh,T., Goto,Y. and Suga,H. (2016) Expanding the amino acid repertoire of ribosomal polypeptide synthesis via the artificial division of codon boxes. *Nat. Chem.*, **8**, 317–325. 10.1038/nchem.2446.
20. Marshall,R., Maxwell,C.S., Collins,S.P., Jacobsen,T., Luo,M.L., Begemann,M.B., Gray,B.N., January,E., Singer,A., He,Y. *et al.* (2018) Rapid and scalable characterization of CRISPR technologies using an *E. coli* cell-free transcription-translation system. *Mol. Cell*, **69**, 146–157 e143. 10.1016/j.molcel.2017.12.007.
21. Maxwell,C.S., Jacobsen,T., Marshall,R., Noireaux,V. and Beisel,C.L. (2018) A detailed cell-free transcription-translation-based assay to decipher CRISPR protospacer-adjacent motifs. *Methods*, **143**, 48–57. 10.1016/j.ymeth.2018.02.016.
22. Hong,S.H., Ntai,I., Haimovich,A.D., Kelleher,N.L., Isaacs,F.J. and Jewett,M.C. (2014) Cell-free protein synthesis from a release factor 1 deficient *Escherichia coli* activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS Synth. Biol.*, **3**, 398–409. 10.1021/sb400140t.
23. Martin,R.W., Des Soye,B.J., Kwon,Y.C., Kay,J., Davis,R.G., Thomas,P.M., Majewska,N.I., Chen,C.X., Marcum,R.D., Weiss,M.G. *et al.* (2018) Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nat. Commun.*, **9**, 1203. 10.1038/s41467-018-03469-5.
24. Pardee,K., Green,A.A., Takahashi,M.K., Braff,D., Lambert,G., Lee,J.W., Ferrante,T., Ma,D., Donghia,N., Fan,M. *et al.* (2016) Rapid, low-cost detection of Zika virus using programmable biomolecular components. *Cell*, **165**, 1255–1266. 10.1016/j.cell.2016.04.059.
25. Sen,S., Apurva,D., Satija,R., Siegal,D. and Murray,R.M. (2017) Design of a toolbox of RNA thermometers. *ACS Synth. Biol.*, **6**, 1461–1470. 10.1021/acssynbio.6b00301.
26. Silverman,A.D., Akova,U., Alam,K.K., Jewett,M.C. and Lucks,J.B. (2020) Design and optimization of a cell-free atrazine biosensor. *ACS Synth. Biol.*, **9**, 671–677. 10.1021/acssynbio.9b00388.
27. Slomovic,S., Pardee,K. and Collins,J.J. (2015) Synthetic biology devices for in vitro and in vivo diagnostics. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 14429–14435. 10.1073/pnas.1508521112.
28. Jackson,K., Kanamori,T., Ueda,T. and Fan,Z.H. (2014) Protein synthesis yield increased 72 times in the cell-free PURE system. *Integr. Biol.*, **6**, 781–788. 10.1039/c4ib00088a.
29. Kuruma,Y. and Ueda,T. (2015) The PURE system for the cell-free synthesis of membrane proteins. *Nat. Protoc.*, **10**, 1328–1344. 10.1038/nprot.2015.082.
30. Vilkhovoy,M., Adhikari,A., Vadhin,S. and Varner,J.D. (2020) The evolution of cell free biomanufacturing. *Processes*, **8**, 675. 10.3390/pr8060675.
31. Garenne,D. and Noireaux,V. (2018) Cell-free transcription-translation: engineering biology from the nanometer to the millimeter scale. *Curr. Opin. Biotechnol.*, **58**, 19–27. 10.1016/j.copbio.2018.10.007.
32. Garenne,D., Haines,M.C., Romantseva,E.F., Freemont,P., Strychalski,E.A. and Noireaux,V. (2021) Cell-free gene expression. *Nat. Rev. Methods Primers*, **1**, 1–18. 10.1038/s43586-021-00046-x.
33. Romantseva,E. and Strychalski,E.A. (2020) CELL-FREE (Comparable Engineered Living Lysates for Research Education and Entrepreneurship) workshop report. NIST Special Publication (SP). National Institute of Standards and Technology. 10.6028/nist.Sp.1500-13.
34. Silverman,A.D., Kelley-Loughnane,N., Lucks,J.B. and Jewett,M.C. (2019) Deconstructing cell-free extract preparation for in vitro activation of transcriptional genetic circuitry. *ACS Synth. Biol.*, **8**, 403–414. 10.1021/acssynbio.8b00430.
35. Takahashi,M.K., Hayes,C.A., Chappell,J., Sun,Z.Z., Murray,R.M., Noireaux,V. and Lucks,J.B. (2015) Characterizing and prototyping genetic networks with cell-free transcription-translation reactions. *Methods*, **86**, 60–72. 10.1016/j.ymeth.2015.05.020.
36. Chizzolini,F., Forlin,M., Yeh Martin,N., Berloffa,G., Cecchi,D. and Mansy,S.S. (2017) Cell-free translation is more variable than transcription. *ACS Synth. Biol.*, **6**, 638–647. 10.1021/acssynbio.6b00250.
37. Dopp,J.L., Jo,Y.R. and Reuel,N.F. (2019) Methods to reduce variability in *E. coli*-based cell-free protein expression experiments. *Synth. Syst. Biotechnol.*, **4**, 204–211. 10.1016/j.synbio.2019.10.003.
38. Cole,S.D., Beabout,K., Turner,K.B., Smith,Z.K., Funk,V.L., Harbaugh,S.V., Liem,A.T., Roth,P.A., Geier,B.A., Emanuel,P.A. *et al.* (2019) Quantification of interlaboratory cell-free protein synthesis variability. *ACS Synth. Biol.*, **8**, 2080–2091. 10.1021/acssynbio.9b00178.

39. Marshall,R. and Noireaux,V. (2018) Synthetic biology with an all *E. coli* TXTL system: quantitative characterization of regulatory elements and gene circuits. In: Braman J (ed). *Synthetic Biology. Methods in Molecular Biology*, vol. 1772. Humana Press, New York. 10.1007/978-1-4939-7795-6_4.
40. Romantseva,E.F., Tack,D.S., Alperovich,N., Ross,D. and Strychalski,E.A. (2022) Best practices for DNA template preparation towards improved reproducibility in cell-free protein production. In: Karim, A.S., Jewett, M.C (eds). *Cell-Free Gene Expression. Methods in Molecular Biology*, vol. 2433. Humana Press, New York. 10.1007/978-1-0716-1998-8_1.
41. Green,M.R. and Sambrook,J. (2018) Isolation and quantification of DNA. *Cold Spring Harb. Protoc.*, **2018**, pdbtop093336. 10.1101/pdb.top093336.
42. Morrow,J. and Olson,N. (2012) DNA extract characterization process for microbial detection methods development and validation. *BMC Res. Notes*, **5**, 668–681. 10.1186/1756-0500-5-668.
43. Yue,K., Jiang,J., Zhang,P. and Kai,L. (2020) Functional analysis of aquaporin water permeability using an *Escherichia coli*-based cell-free protein synthesis system. *Front. Bioeng. Biotechnol.*, **8**, 1000. 10.3389/fbioe.2020.01000.
44. Holland,I. and Davies,J.A. (2020) Automation in the life science research laboratory. *Front. Bioeng. Biotechnol.*, **8**, 571777. 10.3389/fbioe.2020.571777.
45. Kopniczky,M.B., Canavan,C., McClymont,D.W., Crone,M.A., Suckling,L., Goetzmann,B., Siciliano,V., MacDonald,J.T., Jensen,K. and Freemont,P.S. (2020) Cell-free protein synthesis as a prototyping platform for mammalian synthetic biology. *ACS Synth. Biol.*, **9**, 144–156. 10.1021/acssynbio.9b00437.
46. Kelwick,R.J.R., Webb,A.J. and Freemont,P.S. (2020) Biological materials: the next frontier for cell-free synthetic biology. *Front. Bioeng. Biotechnol.*, **8**, 399. 10.3389/fbioe.2020.00399.
47. Chao,R., Mishra,S., Si,T. and Zhao,H. (2017) Engineering biological systems using automated biofoundries. *Metab. Eng.*, **42**, 98–108. 10.1016/j.ymben.2017.06.003.
48. Hillson,N., Caddick,M., Cai,Y., Carrasco,J.A., Chang,M.W., Curach,N.C., Bell,D.J., Le Feuvre,R., Friedman,D.C., Fu,X. *et al.* (2019) Building a global alliance of biofoundries. *Nat. Commun.*, **10**, 2040. 10.1038/s41467-019-10079-2.
49. Moore,S.J., MacDonald,J.T., Wienecke,S., Ishwarbhai,A., Tsipa,A., Aw,R., Kylilis,N., Bell,D.J., McClymont,D.W., Jensen,K. *et al.* (2018) Rapid acquisition and model-based analysis of cell-free transcription–translation reactions from nonmodel bacteria. *PNAS*, **115**, E4340–E4349. 10.1073/pnas.1715806115.
50. Shin,J. and Noireaux,V. (2010) Efficient cell-free expression with the endogenous *E. coli* RNA polymerase and sigma factor 70. *J. Biol. Eng.*, **4**, 1–9. 10.1186/1754-1611-4-8.
51. ArborBiosciences. (2019) *myTXTL Cell-Free Expression Handbook*. https://www.google.com/url?sa=#x003D;t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwix-P6tw935AhX7F1kFHeq1DtoQFnoECAwQAQ&url=https%3A%2F%2Farborbiosci.com%2Fwp-content%2Fuploads%2FmyTXTL-Handbook-v01.pdf&usg=AOvVaw0yXQO6RnlGI3jdKGZnkvrD.
52. Blainey,P., Krzywinski,M. and Altman,N. (2014) Points of significance: replication. *Nat. Methods*, **11**, 879–880. 10.1038/nmeth.3091.
53. Koetsier,G. and Cantor,E. (2019) A practical guide to analyzing nucleic acid concentration and purity with microvolume spectrophotometers. New England BioLabs.
54. Nemenyi,P. (1963) Distribution-free multiple comparisons. Doctoral Thesis. Princeton University, Princeton, New Jersey.
55. Pohlert,T. (2013) PMCMRplus: calculate pairwise multiple comparisons of mean rank sums extended.
56. Team RC. (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.
57. Bates,D., Mächler,M., Bolker,B. and Walker,S. (2015) Fitting linear mixed-effects models using lme4. *J. Stat. Softw.*, **67**, 1–48. 0.18637/jss.v067.i01.
58. Armstrong,R.A. (2014) When to use the Bonferroni correction. *Ophthalmic Physiol. Opt.*, **34**, 502–508. 10.1111/opo.12131.
59. Rader,R.A. and Langer,E.S. (2015) 30 years of upstream productivity improvements. *Bioprocess Int.* https://bioprocessintl.com/upstream-processing/expression-platforms/30-years-upstream-productivity-improvements/.
60. Wehrs,M., Tanjore,D., Eng,T., Lievense,J., Pray,T.R. and Mukhopadhyay,A. (2019) Engineering robust production microbes for large-scale cultivation. *Trends Microbiol.*, **27**, 524–537. 10.1016/j.tim.2019.01.006.
61. Karzbrun,E., Tayar,A.M., Noireaux,V. and Bar-Ziv,R.H. (2014) Synthetic biology. Programmable on-chip DNA compartments as artificial cells. *Science*, **345**, 829–832. 10.1126/science.1255550.