Simple, Functional, Inexpensive Cell Extract for *in vitro* Prototyping of Proteins with Disulfide Bonds

Author List

Jared L. Dopp and Nigel F. Reuel*
Iowa State University Chemical and Biological Engineering
* Correspondence: reuel@iastate.edu

Abstract

In vitro expression of proteins from *E. coli* extract is a useful method for prototyping and production of cytotoxic or unnatural products. However, proteins that have multiple disulfide bonds require custom extract that, to date, requires careful addition of exogenous isomerase enzymes or the use of expensive commercial kits. This cost and complexity currently limit access to some groups who wish to rapidly prototype proteins with disulfide bonds. Herein, we present a simple solution that does not require addition of supplemental enzymes. We use a commercially available SHuffle T7 Express *lysY* strain of *E. coli* that expresses both T7 RNAP and DsbC isomerase enzymes. We experimentally determine optimal growth conditions (IPTG induction and harvest times) to balance overall productivity and efficiency of disulfide bond formation using a luciferase (from *Gaussia princeps*) that contains five disulfide bonds as our reporter protein. We also demonstrate the ability for rapid prototyping by screening the activity of four luciferase candidates against ten luciferin analogues. To display the broad applicability of the extract, three other enzymes containing ≥3 disulfide bonds (hevamine, endochitinase A, and periplasmic AppA) were also expressed from minimal genetic templates that had undergone rolling circle amplification and confirmed via activity assays.

Keywords

Disulfide bond, cell-free protein synthesis, luciferase, chitinase, T7 SHuffle

Introduction

Cell free protein synthesis (CFPS) has seen many recent applications in prototyping gene circuits (Green et al., 2017; Siegal-Gaskins et al., 2014; Sun et al., 2014), sensors (Mcnerney et al., 2019; Pardee et al., 2016a; Salehi et al., 2018; Soltani et al., 2018; Takahashi et al., 2018), and therapies (Cai et al., 2015; Pardee et al., 2016b; Salehi et al., 2016). Cell extract derived

from *E. coli* is the most widely used with simple, efficient extract protocols, multiple supplement recipes, and clear experimental protocols available in literature (Dopp et al., 2019b; Dopp et al., 2019a; Dopp and Reuel, 2018; Kwon and Jewett, 2015; Levine et al., 2019; Sun et al., 2013; Yang et al., 2012). One major limitation of *E. coli* extract for protein expression is the inability to perform post-translational modifications. There has been much work to engineer glycosylation (Jaroentomeechai et al., 2018; Schoborg et al., 2018) and disulfide bond-formation (Kim and Swartz, 2004; Matsuda et al., 2013) pathways. However, these improvements increase the complexity of extract preparation, requiring careful fermentation and skilled augmentation with purified chaperones and enzymes that are expressed in other cells. Commercial kits containing supplements to improve disulfide bond formation are now available but significantly increase the cost per reaction (GeneFrontier; NEB). It is the central goal of this work to create a simple, single cell growth method for creating inexpensive, functional extract for rapid prototyping of proteins with multiple disulfide bonds. Such simple, inexpensive approaches have been shown to have a significant impact on the accessibility and rate of scientific innovation (Whitesides, 2010; Whitesides, 2013).

Many extracellular proteins of interest contain multiple disulfide bonds (Wong et al., 2011), such as therapeutic proteins like human growth hormone (Junnila and Kopchick, 2013), proinsulin (Haataja et al., 2016), antibodies (Saeed et al., 2017), and interleukin-6 (Snouwaert et al., 1991). Disulfide bonds are important because they can improve resistance to degradation by proteases and improve thermodynamic stability (Fass, 2012). Industrial enzyme producers and the biopharmaceutical industry use this strategy to engineer and improve wild type proteins (Lobstein et al., 2012; Matsuda et al., 2013). In eukaryotes, disulfide bonds are typically formed in the endoplasmic reticulum (ER) while folding in the periplasm of *E. coli* (Inaba, 2010; Patil et al., 2015). This typically leads to low yields (Ke and Berkmen, 2014; Mamathambika and Bardwell, 2008) and can result in proteins being sequestered to inclusion bodies that require refolding (Baneyx and Mujacic, 2004; Burgess, 2009; Middelberg, 2002). Since the periplasm is compromised during cell lysis, there is no separate compartment for these oxidative processes. Figure 1 depicts the oxidation, reduction, and isomerization processes undergone by thiol groups cysteine residues.

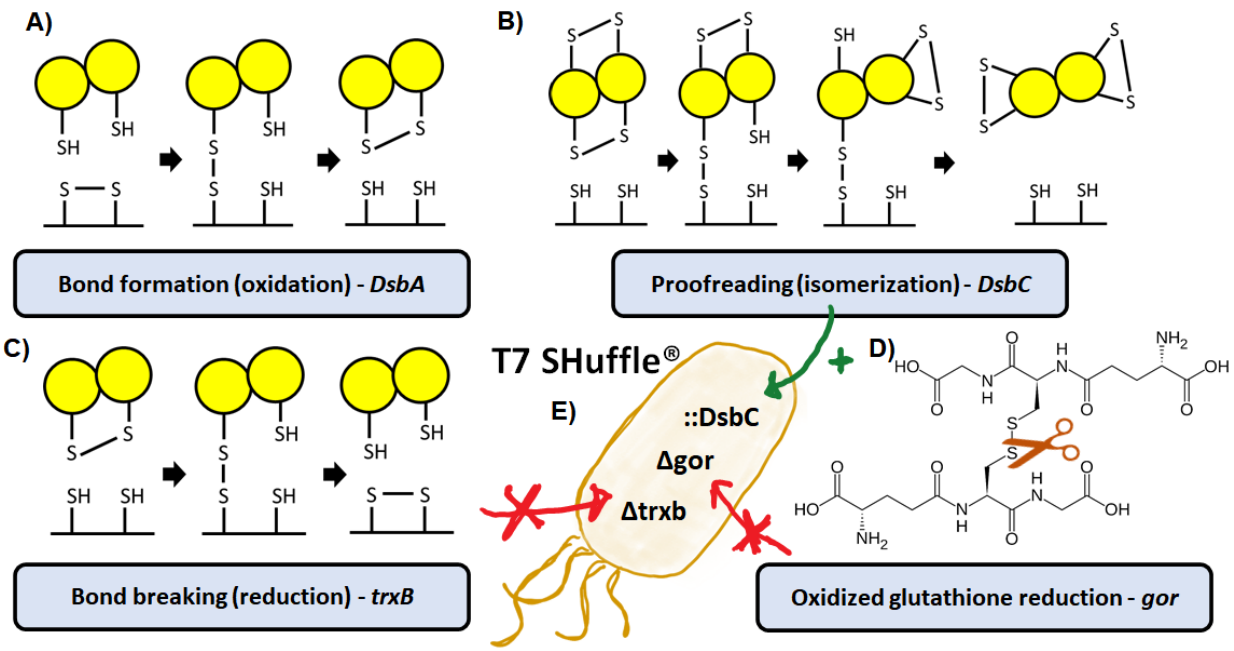


Figure 1 – Schematic of mechanisms that affect disulfide bond formation and implications in the T7 Shuffle strain. (A) Oxidation via DsbA forms bonds between thiol groups on cysteines. (B) DsbC enzymes proofread proteins and isomerize disulfide bonds. (C) Reduction can occur via trxB and (D) gor enzymes that cleave disulfide bonds on the protein. (E) The T7 SHuffle® strain is engineered to support disulfide bond formation by eliminating reducing enzymes and overexpressing the DsbC chaperone. This figure has been modified from a published schematic on disulfide bond formation (Ke and Berkmen, 2014).

With these issues in mind, researchers have taken steps to improve cytoplasmic expression (*in vivo* and *in vitro*) of proteins with disulfide bonds. In 2003, Kim and Swartz successfully expressed the protease domain of murine urokinase (6 disulfide bonds) using extract from the *E. coli* A19 strain (K12 derivative). This was achieved by manipulating the redox environment and adding disulfide bond isomerase C (DsbC) (Kim and Swartz, 2004). Soon thereafter, a truncated version of tissue type plasminogen activator (9 disulfide bonds) was successfully expressed. Decreasing the reaction temperature to 30 °C and adding the chaperone Skp improved expression (Yin and Swartz, 2004). Around the same time, the KC1 strain was developed to improve amino acid stability (arginine, tryptophan, and serine) during CFPS reactions (Michel-Reydellet et al., 2004). In 2005, an active GMCSF-sfFv fusion protein (4 disulfide bonds) was produced using the NMR2 strain, another modification from A19 (Michel-Reydellet et al., 2004; Voloshin and Swartz, 2005). More progress was made when KC1 was modified to create the KC6 mutant with enhanced cysteine stability due to the deletion of glutamate-cysteine ligase (*gshA*). This stabilization led to improved overall production but a massive increase in active murine urokinase and GM-CSF (Calhoun and Swartz, 2006). In order to support disulfide bond formation during CFPS using glucose as an energy source, the KGK10 strain was created by deleting the glutathione reductase (*gor*) gene from the KC6 strain. A hemagglutinin purification tag was also added to the C-terminus thioredoxin reductase gene (*trxB*) but its removal proved to be unnecessary. Murine urokinase and GM-CSF were produced

but with lower yields due to glucose replacing phosphoenolpyruvate (PEP) as an energy source (Knapp et al., 2007). A direct comparison of KGK10 and KC6 using the PANOx-SP system showed few decreases in production but many improvements. It was also shown to be effective at expressing multiple scFv vaccine candidates showcasing the potential for CFPS in the therapeutic space (Goerke and Swartz, 2007). Around this same time, the Buchner group was able to express a full-length IgG antibody using CFPS. 507 ng/mL of active MAK33 was produced using two commercial kits (RTS100 HY and RTS500 HY) upon the addition of DsbC, GSSG, and GSH (Frey et al., 2008). In 2012, a correctly folded mutant of the flu virus hemagglutinin stem domain was produced using the KC6 strain. The number of disulfide bonds in the monomers of the stem domain was reduced from 4 to 2. In order to stabilize the trimer that forms the stem domain, cysteines were added to form disulfide bonds between monomers (Lu et al., 2014). This was an excellent display of using CFPS for protein engineering to improve protein production and stability by manipulating disulfide bond formation.

These advancements have led to industrial-scale applications of CFPS for producing therapeutics with disulfide bonds. Sutro Biopharma achieved a large milestone by demonstrating the successful scale-up of CFPS by producing recombinant human GM-CSF, along with other therapeutic proteins. They found that protein production using a cell-free system scales linearly as the reaction volume is increased (Zawada et al., 2011). In 2012, Sutro hit another milestone by demonstrating they're Open Cell-Free System (OCFS) could be used for high throughput antibody engineering. In the process, they were able to optimize conditions for trastuzumab IgG1 (16 disulfide bonds) expression (Yin et al., 2012). Sutro then engineered a strain to improve mAb production by incorporating 2 chromosomal copies of DsbC and a plasmid with 2 copies of FkpA (Groff et al., 2014). This work has led to the production of the SBJY001 strain which requires less processing during extract production with the ability to produce tumor seeking antibody drug conjugates (Yin et al., 2017). Unfortunately, these strains require defined media, fermentation, and dialysis, and are not commercially available (Liu et al., 2005; Zawada and Swartz, 2005) which has limited their widespread impact.

To quantify impact of cell extract methods, we have quantified their use in scientific citations. The most used strain in CFPS, BL21 (DE3) Star, is a commercially available strain capable of producing functional lysate with minimal processing steps and has been widely cited (Fig 2). BL21 (DE3) Star, and similar BL21 strains, have been used to express proteins containing disulfide bonds with varying success. For example, BL21 (DE3) RIPL-Star was used to express human and murine proteins, each containing 2 disulfide bonds, without pretreating the extract. It was also shown that creating a fusion with the GB1 protein at the N terminus increases expression (Michel and Wüthrich, 2012b; Michel and Wüthrich, 2012a). More recently, BL21 (DE3) Star has been used to successfully express correctly folded hydrophobins (4 disulfide bonds) upon the addition of DsbC but no IAM pretreatment (Siddiquee et al., 2020). However, these BL21 (DE3) Star strains do not have modifications aimed at improving disulfide bond formation. The only commercial systems available are kits that provide expensive supplements for disulfide bond formation. The PURE*frex*® 2.1 system, made of purified, recombinantly expressed proteins, is one of these commercial platforms capable of reliably producing proteins with disulfide bonds (Murakami et al., 2019). This reliability comes at a steep price with a cost of

roughly US $0.57/µL of reaction and inclusion of the disulfide bond supplement raises the price to US $0.70/µL of reaction and has thus limited impact to those research groups that can afford such kits (Fig 2). On the other hand, creating CFPS components in house can reduce cost to $0.02/µL of reaction (Gregorio et al., 2019; Levine et al., 2019). The complexity and access limitations to KC6, KGK10, and SBJY001 are likewise revealed when counting citations of books, peer reviewed articles, and preprints from Google Scholar (Fig 2).

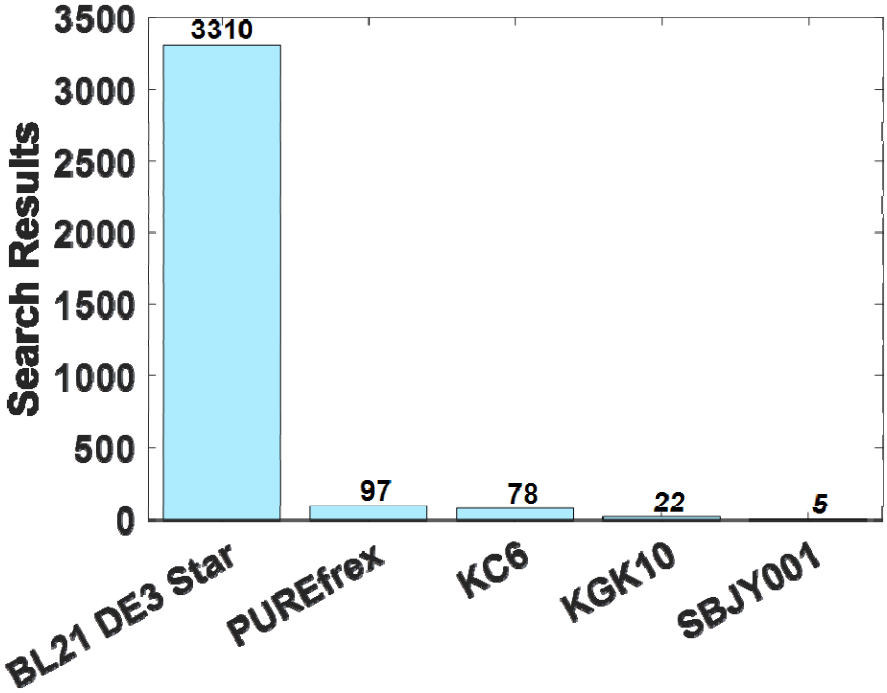


Figure 2 – The number of search results from Google Scholar when searching cell free and the strain name between the years 2000 and 2020. The results from PURExpress come from searching for PURExpress alone since it is exclusively a cell-free platform. These results include mentions in published articles, preprints, and books. These numbers represent the results from a search performed February 21, 2020.

In line with our goal to make CFPS a simple, functional, inexpensive method that is readily transferable to any research group (Dopp et al., 2019b; Dopp et al., 2019a; Dopp and Reuel, 2018), we desired to identify and optimize a strain that could be used to produce extract capable of forming disulfide bonds without the need for supplementation with exogenous enzymes. Ideally, the strain should be grown in a simple shake flask and would not require a run-off reaction or dialysis. We identified a commercially available *E. coli* strain that has been optimized to support *in vivo* disulfide bonded protein expression in the cytoplasm, T7 SHuffle® by New England BioLabs (NEB). The SHuffle strain is derived from ER2566 and has been modified to constitutively express disulfide bond isomerase C (DsbC) in the cytoplasm and overexpress T7 RNA polymerase (T7 RNAP) upon induction by isopropyl β- d-1-thiogalactopyranoside (IPTG) (Fig 1). The SHuffle strain also eliminates unwanted reduction pathways in the cytoplasm due to deletion of the *trxB* and *gor* genes (Fig 1E) (Anton et al., 2016; Lobstein et al., 2012; Robinson et al., 2015).

To understand the significance of these modifications, it is important to have an understanding of the intracellular redox environment. The indiscriminate oxidation of cysteines is carried out by disulfide bond isomerase A (DsbA) which forms disulfide bonds between successive cysteines (Fig 1A). DsbC allows for thermodynamically driven proof-reading by isomerizing disulfide bonds that have been formed between sub-optimal cysteines, thus allowing for continued folding to the proper state and disulfide linking (Fig 1B). In an unmodified strain, thioredoxin reductase will reduce disulfide bonds or prevent them from forming. Also, in unmodified strains, glutathione reductase will reduce oxidized glutathione (Fig 1D), thus removing its ability to oxidize thiol groups on cysteines (Ke and Berkmen, 2014; Lobstein et al., 2012).

In this work, we investigate the potential of T7 SHuffle® as a platform for CFPS based prototyping of proteins with multiple disulfide bonds. We optimize the extract preparation based on our scalable techniques previously used for BL21 (DE3) Star cells (Dopp and Reuel, 2018). Luciferase from *Gaussia princeps* (Gluc) is used as a disulfide bonded reporter protein for extract optimization due to its efficient luminescence and previous application in CFPS (Goerke et al., 2008; Liu et al., 2014; Yu et al., 2018). We show that T7 SHuffle® is superior to BL21 (DE3) Star when expressing Gluc (5 disulfide bonds). We then demonstrate the broader applicability of this extract by screening 4 luciferase candidates with >90% homology to Gluc against a panel of 10 luciferase substrates. We further demonstrate prototyping with this extract by expressing three additional enzymes with disulfide bonds: hevamine from the rubber tree *Hevea brasiliensis* (Bokma et al., 2002), endochitinase A (ChitA) from *Zea mays* (Hawkins et al., 2015), and the periplasmic acid phosphatase, phytase (AppA) from *Escherichia coli* (Berkmen et al., 2005) which have 3, 7, and 4 disulfide bonds respectively. Some practical applications of the two chitinases include agricultural pathogen control, contribution to asthma, and formation of chitooligosaccharides (Hamid et al., 2013). Moreover, we introduce a custom fusion protein that allows us to measure the ratio between oxidative potential and general productivity of an extract.

Materials and Methods

Extract Preparation

T7 SHuffle®

Extract from the SHuffle® T7 Express *lysY* strain (gift from the Mansell lab) was prepared according to previously established protocols (Dopp and Reuel, 2018). Cells were grown and harvested in 200 mL 2x YTPG media using a 500 mL shake flask at 37°C and 230 RMP. Cultures were harvested at different growth/induction times according to the DoE design (Supplementary Table 1). The cells were centrifuged for 15 minutes at 5,000 xg and 4°C. The cell pellets were collected and washed in 10 mL of cold buffer A and centrifuged for 15 minutes at 4,255 xg and 4°C (Dopp et al., 2019a). The supernatant was discarded, and the wet cell

mass was measured before freezing the cells at -80°C. Cells were thawed and resuspended in 1 mL of buffer A per gram of wet cell mass. Suspended cells were tip sonicated (QSonica) in 1 mL aliquots at -4°C using a 10 sec on/off pulse and 50% amplitude until the energy input reached 532 J (Kwon and Jewett, 2015). Sonicated cell suspensions were then centrifuged for 10 minutes at 4°C and 12,000 xg to create S12 extract. The supernatant (S12 extract) was collected and stored at -80°C in 45 uL aliquots. Once the optimum growth/induction times were determined for 200 mL cultures, Matlab was used to fit 200 mL and 1 L culture growth curves. The regression used was the Verhulst-Pearl equation (VPE) used in other similar studies (Supplementary Figure 1) (Wilding et al., 2018). These curve fits were used to correlate the optimum at 200 mL with a predicted optimum at 1 L. Optimized 1 L cultures were then grown in 2.5 L Tunair flasks and harvested according to the aforementioned protocol but were lysed using a French press homogenizer (Avestin EmulsiFlex C3) according to previous protocols (Dopp and Reuel, 2018). Frozen cells were resuspended in 1 mL buffer A + 1 mL ddH_2O per 1 g of wet cell mass. This double dilution reduces the viscosity and maintains buffer composition upon reconstitution. The suspension was then fed through the French press and lysed at 20,000 - 30,000 psig. The crude lysate was then centrifuged for 10 min at 12,000 xg and 4°C. The supernatant was then collected, frozen at -80°C, and lyophilized in a VirTis pilot lyophilizer (SP Scientific) overnight. The resulting powder was collected, weighed, and stored at -80°C. Concentrated extract was obtained by re-lyophilizing extract and adding less water. Dilute extract was obtained by adding more water.

KGK10

The KGK10 strain (gift from Swartz lab) was grown in a 500 mL shake flask and monitored every 15 min to obtain a growth curve. This growth curve was used to find the point at which the rate of growth was the highest. KGK10 was then grown and harvested at 240 min (Supplementary Figure 2). The cells were centrifuged for 15 minutes at 5,000 xg and 4°C. The cell pellets were collected and washed in 10 mL of buffer A and centrifuged for 15 minutes at 4,255 xg and 4°C. The supernatant was discarded, and the wet cell mass was measured before freezing the cells at -80°C. Cells were thawed and resuspended in 1 mL buffer A + 1 mL ddH_2O per 1 g of wet cell mass, identical to the method used for SHuffle extract. The suspension was then fed through the French press and lysed at 20,000 - 30,000 psig. The crude lysate was then centrifuged for 10 min at 12,000 xg and 4°C. The supernatant was then collected, frozen at -80°C, and lyophilized in a VirTis pilot lyophilizer (SP Scientific) overnight. The resulting powder was collected, weighed, and stored at -80°C.

DNA Amplification

All Gluc and sfGFP reporter protein expression experiments were carried out using plasmids. The sfGFP plasmid used was pJL1-sfGFP (gift from the Jewett lab) and the Gluc plasmid was pET24a-Gluc-6H (gift from the Swartz lab).

The production of minimal genetic templates for prototyping new enzymes has been described in previous work (Dopp et al., 2019b). In brief, the gene of interest was codon optimized using

IDT's codon optimization tool and purchased in the form of a gene fragment. The gene fragment was then amplified using OneTaq (NEB), digested to form sticky ends with HindIII (NEB), ligated with T4 ligase (NEB), and isothermally amplified with TempliPhi (GE Healthcare). This was all done in under 24 hr using a standard C1000 Touch Thermal Cycler (BioRad). The sequences for all linear and minimal genetic templates used are in the accompanying supplementary information.

Cell-Free Reaction

T7 SHuffle®

The supplement recipe used for the designed experiments and initial tests is a modified version of the PANOx-SP system that is improved to form disulfide bonds (Dopp et al., 2018). Since the SHuffle strain can be induced to produce T7 RNAP and DsbC, these proteins were omitted from the reaction mix. The cell-free reaction included: 1.2 mM ATP, 0.85 mM each of GMP, UMP, and CMP, 30 mM phosphoenolpyruvate (Roche), 130 mM potassium glutamate, 10 mM ammonium glutamate, 12 mM magnesium glutamate, 1.5 mM spermidine, 1 mM putrescine, 34 µg/mL folinic acid, 171 µg/mL *E. coli* tRNA mixture (Roche), 4 mM oxidized glutathione (GSSG), 1 mM reduced glutathione (GSH), 2 mM each of 20 unlabeled amino acids, 0.33 mM NAD, 0.27 mM Coenzyme A (CoA), 4 mM potassium oxalate, 57 mM HEPES-KOH buffer (pH 7.5), 15 ng/uL plasmid, 0.24 volumes of E. coli S12 extract. Reactions were carried out in a 384 black-walled, flat-bottom well plate and shaken at 300 rpm using a ThermoMixer® (Eppendorf). All proteins were expressed at 30°C unless otherwise stated. Gluc, hevamine, rChiA, and AppA reactions were run for 16 hrs, unless otherwise stated. The fluorescent proteins (sfGFP and rxYFP-mCherry) were expressed for 6 hrs. Reactions for the fluorescent proteins were carried out in a Synergy Neo2 HTS Multi-Mode Microplate Reader (BioTek) at 282 cpm.

KGK10

The cell-free reaction mixture for the KGK10 strain is identical to that used for the T7 SHuffle® with the addition of 100 µg/mL DsbC (GeneFrontier) and 3.33 Units/uL T7 RNAP (Roche). Previously prepared KGK10 extract fermented and processed by the Swartz lab was tested alongside the in-house KGK10 extract prepared by shake flasks.

PURE*frex*®

The components for the PURE*frex* reactions were scaled down to suit 15 uL reactions. The components from the PURE*frex*® 2.1 (GeneFrontier) were supplemented with DsbC and GSSG from the DS Supplement kit (GeneFrontier) according to manufacturer's protocols.

Assays

Bioluminescence

The Gluc assay is based on previous work in the Swartz lab (Goerke et al., 2008). 3 µL of CFPS reaction was diluted in 100 µL nickel affinity EB buffer containing 1% w/v bovine serum albumin (BSA). The activity was measured by adding 2 µL of diluted sample to 100 µL assay buffer with 1 µL of 0.5 µg/µL coelenterazine (NanoLight Technology) dissolved in ethanol and immediately mixing before reading the luminosity. The sample buffer consisted of PBS pH 7.4 with 0.01% v/v Tween 20. Readings taken in a U-bottom 96 well plate. Luminescence was measured every 10 sec for 60 sec using a Synergy Neo2 HTS Multi-Mode Microplate Reader (BioTek). The plate reader settings used a 1536 filter, top optics position, and a gain of 100.

Coelenterazine analogues for the luciferase screening experiment were ordered from BioTium and suspended at a concentration of 0.5 µg/µL in the suggested solvent (ethanol or methanol). A water-soluble analogue was ordered from NanoLight Technology and dissolved at 0.5 µg/µL in ddH_2O.

Fluorescence

Fluorescence measurements were also taken using a Synergy Neo2 HTS Multi-Mode Microplate Reader (BioTek). Readings were taken every 5 min. The fluorescence of sfGFP was measured at an excitation of 485 nm and emission of 528 nm using a ±20 bandpass window and 61 gain setting. The reaction was also stirred in orbital motion at 237 cpm. For rxYFP, readings were taken at an excitation of 512 and emission of 523 using a ±5 bandpass window. Measurements for mCherry were taken at an excitation of 587 ±11 and an emission of 610 ±10.

Enzyme Activity

Hevamine activity was determined by adding 10 µL CFPS reaction to 90 µL McIlvaine buffer pH 6.0 with 10 µL of 1 mM 4-methylumbelliferyl β-D-N,N′,N″-triacetylchitotrioside (4-MUF-TriNAG). rChiA activity was determined by adding 10 µL CFPS reaction to McIlvaine buffer pH 4.0 with 10 µL 4-methylumbelliferyl N-acetyl-β-D-glucosaminide (4-MUF-NAG). Fluorescence was measured at an excitation of 360 and emission of 445 with a bandpass window of ±20. Measurements were taken every 60 sec for 2 hrs in a black walled, flat bottom 384 well plate. AppA activity was determined by adding 10 µL CFPS reaction to 100 µL glycine buffer (250 mM glycine + 25 mM p-nitrophenyl phosphate) titrated to pH 3.0 with HCl. Absorbance readings were taken at 410 nm every 20 sec for 5 min in a 96 well plate at 37°C.

Results and Discussion

Designed Experiments

To optimize the growth conditions of the SHuffle strain, we performed two, face-centered-cubic (FCC) response surface design of experiments (DoE). This is an established statistical method to optimize an experimental outcome, with minimal experimental runs while accounting for cross-effects of input variables (Ferreira et al., 2007). In this case, the input variables under study are time of induction with ITPG to produce T7 RNAP and DsbC proteins as well as the overall time of harvest. The outcome to optimize is the amount of protein: in the first DoE, we measure the yield of sfGFP to compare to other strains (namely our work with BL21 DE3 Star cells); in the second, we measure the level of correctly folded Gluc to observe the level of disulfide-bonded protein yield. Gluc is a luciferase with five disulfide bonds and a strong luminescent signal making it a convenient reporter for extract optimization (Wu et al., 2015; Yu et al., 2018). In both cases, we constrained the growth times to encompass a single workday (< 7 hr growth time, as pelleting, washing, and storage must be completed that same day). Each response surface design yielded 13 runs, of which 5 were replicated center points to measure experimental drift. The measured values were then fit to quadratic models and the response surfaces are presented as contour plots (Fig 3 A-B).

The response surface for sfGFP shows that longer growth times would improve the expression (Fig 3A). This indicates that insufficient T7 RNAP is being expressed to maximize overall yield; this is also apparent in the level of sfGFP fluorescence observed which is less than we typically observe with BL21 DE3 Star extract when assayed on the same plate. Since sfGFP contains no disulfide bonds, this response surface is not capable of predicting optimum DsbC production for chaperoning disulfide bond formation. This was resolved using our second reporter protein, Gluc. The response surface for Gluc luminescence output has a clear maximum within our experimental bounds (Fig 3B), indicating the optimum balance of chaperone protein and T7 RNAP to produce disulfide proteins. An optimum growth time of 5.6 hr and an optimum induction point at 3.7 hr are found for these smaller growth vessels (200 mL media in 500 mL Tunair flasks), with greater sensitivity in growth time (Supplementary Figure 1A). These optimum induction and harvest time points were then translated to percentages of the total growth curve (fit by a Verhulst-Pearl logistics function) to allow for application to other growth curves occurring in larger vessels or other orbital shakers (50.49% and 13.12% respectively), as we have done before (Dopp and Reuel, 2018). The optimum time points for the 2.5 L shake flasks we use with 1 L culture are then found from these percentages (5.6 and 3.9 hr for harvest and induction respectively). The statistics for both fits are listed in Supplementary Table 2. It should be noted that while Gluc is a convenient reporter protein, optimal extract processing conditions may be protein dependent. This is also evident when using Shuffle for *in vivo* expression conditions where temperature, time of induction, and IPTG concentration must be optimized for each new protein to see best results (Lobstein et al., 2012). However, for the prototyping phase of new proteins we have found sufficient level of expression to quantify protein and measure desired properties using a single general extract; the quest for custom,

optimal extracts for each protein is better suited in a later stage such as scale up for production of a few desired proteins.

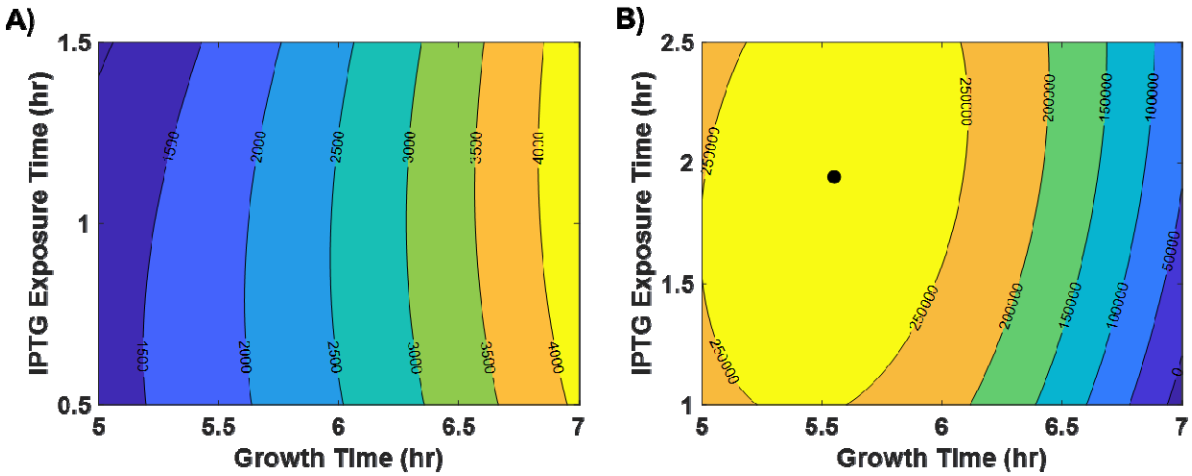


Figure 3 – Experiments to optimize expression from T7 Shuffle extract. (A) Response surface fit to determine optimal growth conditions for sfGFP expression; z axis shows fluorescence (B) Response surface fit to determine optimal growth conditions for Gluc expression; z axis shows luminescence.

Extract and Iodoacetamide Concentration

To improve expression, we explored the impact of extract concentration and adding a common reagent used to inactivate thiol containing reductases, iodoacetamide (IAM) (Kim and Swartz, 2004; Knapp et al., 2007). Previous literature has shown that increasing extract concentration can lead to increased expression rates at the cost of a reduction in total protein produced so we decided to explore the effect of dilution (Didovyk et al., 2017). We assumed that IAM may inhibit the efficacy of the constitutively produced DsbC since IAM is a small molecule that covalently bonds with thiol groups on reductases to reduce their ability to impair disulfide bond formation. Therefore, we reduced the IAM concentration from 50 μM to 30 μM for the experiments shown in Fig 4A. Surprisingly, subsequent dilutions, except 1/5x, outperformed the standard, non-diluted extract. We then tested possible benefits to manipulating the IAM concentration. The previous assumption that IAM would interfere with the thiol groups of endogenous DsbC was observed to be false. Figure 4B shows that of the IAM concentrations screened, 50 μM performed the best. Henceforth, all experiments were conducted with 1/2x, diluted SHuffle extract.

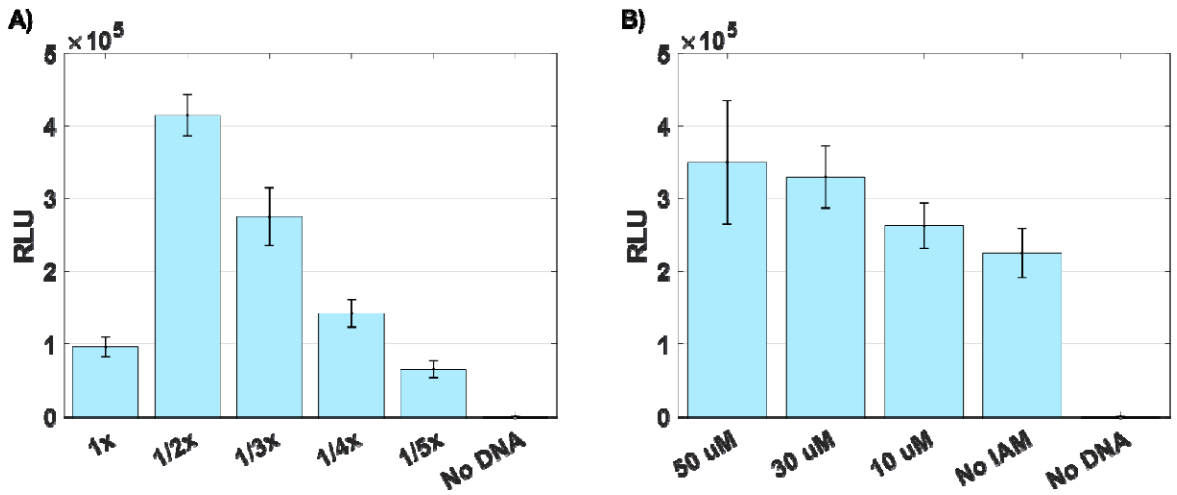


Figure 4 – The effects of extract concentration (A) when IAM is held at a constant concentration (30 µM) and the effect of IAM concentration (B) when extract is held at a constant concentration (1/2x) (n=5).

Production Comparison

To investigate the effect of strain selection on disulfide bond formation, we expressed Gluc using 5 different extracts: SHuffle, KGK10 extract prepared by the Swartz group, KGK10 prepared in our lab, BL21 (DE3) Star, and PURE*frex* 2.1. Figure 5A shows the results of all extracts except PURE*frex* because the signal saturated the detector at gain = 100; by reducing the gain to 80, we can see the superior performance of the commercial kit (Fig 5B). Comparative analysis of the non-commercial extracts is done at gain = 100. The data show that while BL21 (DE3) Star, upon supplementation with DsbC, can support the production of Gluc, the signal is 2.17x lower than that of SHuffle. In contrast, KGK10 extract from the Swartz lab (KGK10 #1) is 1.71x more productive than Shuffle, however, recall that the KGK10 extract requires supplementation of exogenous proteins (T7 RNAP and DsbC) while the Shuffle extract requires neither. Additionally, when we attempt to produce KGK10 using inexpensive shake flasks (Dopp et al., 2019a; Dopp and Reuel, 2018), rather than fermentation in a controlled bioreactor (KGK10 #2), we see very little expression. This suggests that KGK10 requires very strict growth and processing conditions (Zawada and Swartz, 2005) that are inaccessible to labs that do not specialize in cell fermentation (intended ‘users’ of this technology). The SHuffle extract is the best choice for inexpensive, functional *in vitro* production of this reporter protein.

Since many proteins of interest do not have disulfide bonds, relative expression levels of proteins without disulfide bonds can also be important information when selecting a strain. To investigate this, we expressed sfGFP using BL21 DE3 Star, KGK10, and T7 SHuffle. We found that BL21 DE3 Star has the most productivity with KGK10 closely following (Fig 5C). This corroborates the finding of our designed experiment which identified the optimal extract for disulfide bond formation to be suboptimal in overall level of protein expression (as measured by sfGFP). These data suggest that the BL21 DE3 star strain is still the better option for proteins with no or few disulfide bonds.

To explore the use of BL21 DE3 Star for proteins with few, proximal disulfide bonds (due to native DsbA present in the extract), we designed a method to compare the ratio between oxidative potential and general productivity. We did this by designing and expressing a new fusion protein made of a yellow fluorescent protein variant (rxYFP) sensitive to the redox environment, a flexible linker, and mCherry to measure overall production. To our knowledge, this is the first fusion protein designed to measure redox/productive potential, especially in cell-free extract. The rxYFP variant has been modified by the addition of two non-native cysteine residues. The fluorescent signal from rxYFP is significantly reduced in the presence of an oxidative environment due to disulfide bond formation between these non-native cysteines (Østergaard et al., 2001). The linker selected was previously used to successfully fuse a nanobody to EGFP and still support proper folding (Li et al., 2012). In order to keep consistent molarity of the DNA template between the reactions, the rxYFP-mCherry fusion template was introduced as linear DNA (see methods). To quantify oxidative potential, the fluorescence of rxYFP was divided by that of mCherry (the lower the ratio, the better the environment for disulfide bond protein expression). Surprisingly, the Shuffle and KGK10 strains have higher ratios than the BL21 DE3, suggesting, again, that the Star strain is suitable to express proteins with low number (in this case 1) of disulfide bonds (Fig 5D).

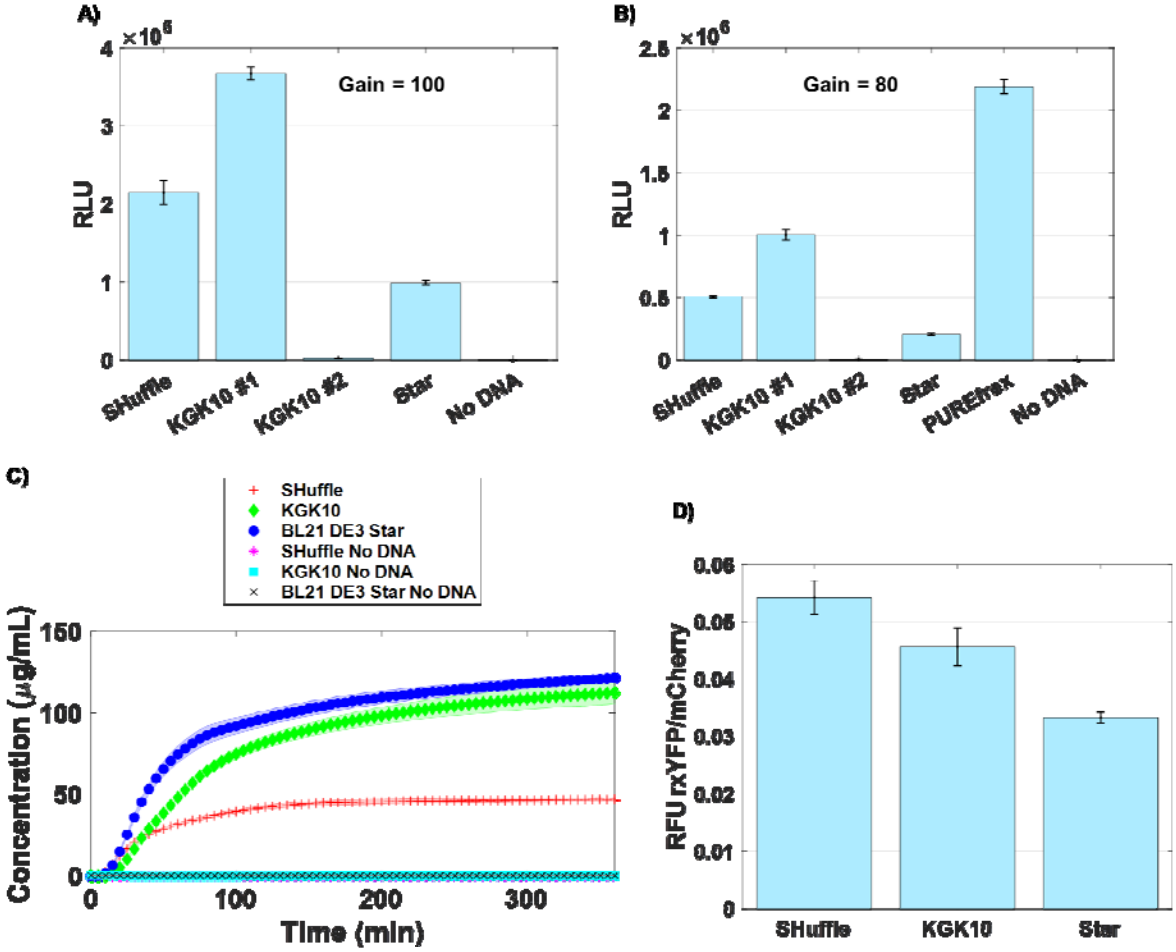


Figure 5 – Experiments to compare the yield of T7 SHuffle®, KGK10, BL21 DE3 Star, and the PURE*frex* 2.1 kit. KGK10 #1 is carefully fermented KGK10 extract that was provided by the Swartz lab. KGK10 #2 is grown in-house using a shake flask. (A) Gluc signal comparisons at a gain of 100 without PURE*frex* data (n=3). (B) Gluc signal comparisons with an instrument gain of 80 so PURE*frex* data does not saturate the detector. (C) Expression of sfGFP to compare yield of proteins without disulfide bonds over 16 hr (n=3). (D) Is a ratio of oxidation potential/productivity using a YFP-mCherry fusion where a S-S bond has been introduced into a YFP variant (n=3).

Exogenous Supplementation

To explore the impact of added enzymes, exogenous DsbC and T7RNAP were added to the SHuffle reactions. We also investigated the effect of removing these supplements from the fermented KGK10 reactions. Interestingly, the performance of the SHuffle extract decreased when exogenous proteins were added (Fig 6). With the KGK10 extract, we observed that it can still effectively express Gluc without the addition of DsbC, but that T7RNAP is essential for protein production. The tolerance to no chaperone protein may be isolated to this reporter protein as the majority of Gluc's disulfide bonds occur between consecutive cysteines, thus relying mostly on the native periplasmic isomerase DsbA to form bonds. Previous work with DsbC shows that the bond pattern, or the order in which the cysteines are bonded to one another (consecutive *vs* nonconsecutive), determines how important DsbC is for functional protein production (Berkmen et al., 2005; Ke and Berkmen, 2014; Lobstein et al., 2012). While the KGK10 strain is entirely reliant on exogenous T7RNAP and DsbC for expression, the Shuffle extract's best expression occurs without the addition of either enzyme. This correlates well with the DoE showing that diluting the extract increases expression in SHuffle. These data again highlight that SHuffle has the advantage of being simpler to use than KGK10. The only requirement for efficient production using SHuffle extract is the optimized induction time of IPTG and the optimized harvest time in a shake flask reactor. KGK10 extract is more complicated to produce because the strain doesn't produce T7RNAP or DsbC and requires fermentation with a defined media (Knapp et al., 2007).

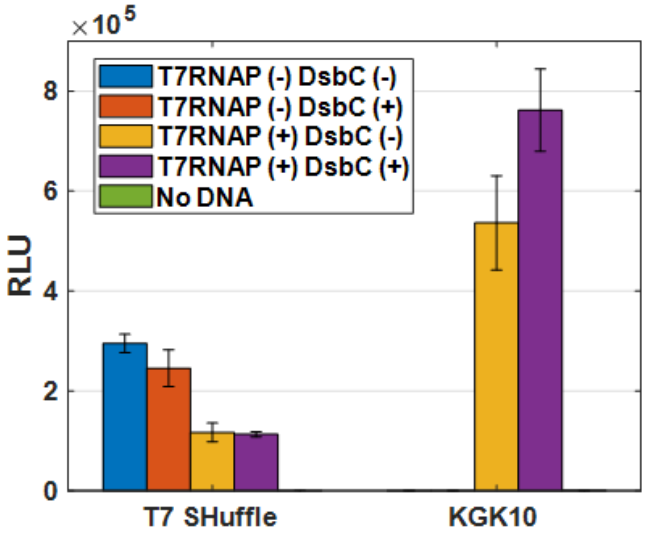


Figure 6 – Exploring the effects of including/excluding exogenous T7 RNAP and DsbC to/from cell-free reactions. Gluc activity was measured with luminescence (RLU) (n=3).

Prototyping Proteins

Optimizing Process Conditions

Before prototyping proteins with the Shuffle extract, we determined best process conditions. While typical cell-free reactions are conducted at 37°C for 4 hr, reactions producing proteins with disulfide bonds are commonly conducted at 30°C and conducted overnight (Dopp et al., 2018; Gregorio et al., 2019). The lower temperature has been shown to enhance eukaryotic protein folding efficiency in *E. coli* and longer reaction times can lead to improved titers (Siller et al., 2010). Since we noticed our sfGFP expression flattened off relatively quickly, we decided to test Gluc expression at a shorter time and higher temperature to improve throughput (Fig 7A). The resulting signal is well above what is needed for detection but significantly lower than a 16 hr expression at 30°C. For our work, we have decided to prototype proteins at the slower, but more productive conditions of 16 hr at 30°C. Using these conditions, we then tested the efficacy of different genetic templates at equimolar concentrations (Fig 7B). We've previously shown that rolling circle amplification (RCA) can be applied to minimal linear template to avoid time-consuming cloning steps (Dopp et al., 2019b). While the plasmid was much more effective, the signal from the RCA products is still well above noise. We also determined that the SHuffle extract can produce suitable Gluc signal without the 30 min IAM pretreatment in a 4 hr reaction (Supplementary Fig 3). These process conditions were then used for a few prototyping demonstrations, though extract was still pretreated with IAM.

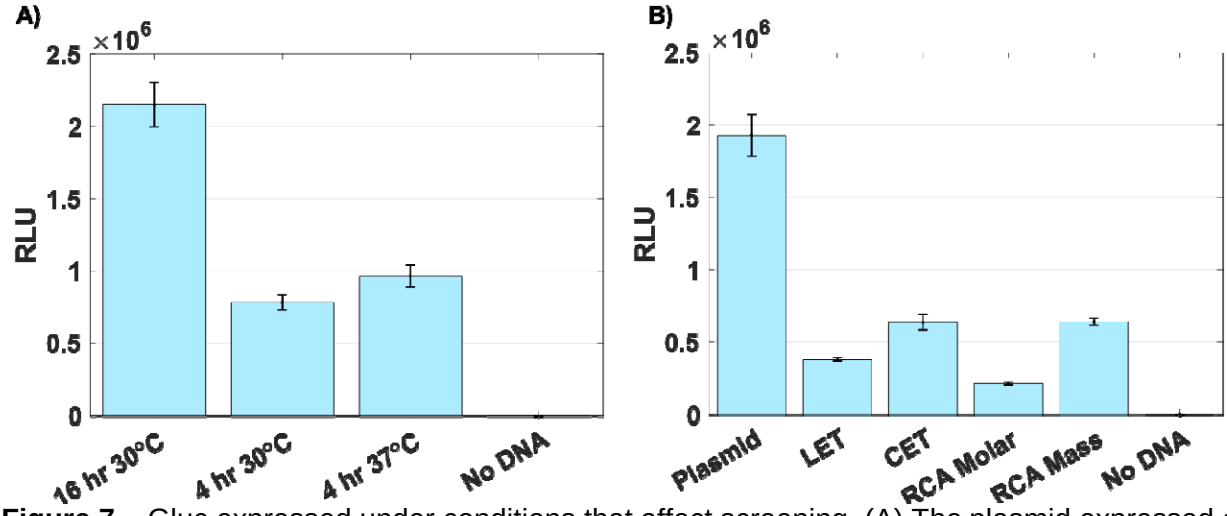


Figure 7 – Gluc expressed under conditions that affect screening. (A) The plasmid expressed at different times and temperatures (n=3). (B) Different genetic templates expressed at 30°C for 16 hr (n=3). 'RCA Molar' assumes RCA and CET share the same molecular weight so all templates should have an equimolar concentration. 'RCA Mass' has the same mass concentration as the plasmid. 'No DNA' control shows level of background signal from extract in luminescence assay.

Prototyping Luciferases

To demonstrate the prototyping utility of the SHuffle extract, we assembled a panel of 5 luciferase sequences from the Uniprot database; one is the native Gluc sequence while the other four have at least 90% homology with native Gluc but were annotated as 'uncharacterized' (never expressed or tested, but found via genome sequencing). A high percentage of homology was chosen to ensure the four, putative luciferases contained multiple disulfide bonds and acted on the substrate coelenterazine. To further demonstrate the power of cell free prototyping, we tested the activity of these putative luciferase enzymes on various coelenterazine analogues to determine if slight differences in amino acid sequence would have noticeable effects on substrate preference (Dikici et al., 2009). As seen in Fig 8, the most effective combination in this panel was the native luciferase from *Gaussia princeps* and the water-soluble analogue of coelenterazine. These results are consistent with current literature (Morse and Tannous, 2012). We can also readily see the substrate compatibility and effect of amino acid sequence changes. Such information could be used in protein engineering studies to improve substrate or protein sequence. All luciferase candidates were expressed from minimal templates that had undergone RCA. The genetic sequences for the luciferase candidates are listed in the Supplementary material (Zhao et al., 2004). Images of the coelenterazine analogues are also provided in the supplement. We were unable to provide a structure for the water soluble (sol) analogue and believe the structure to be proprietary.

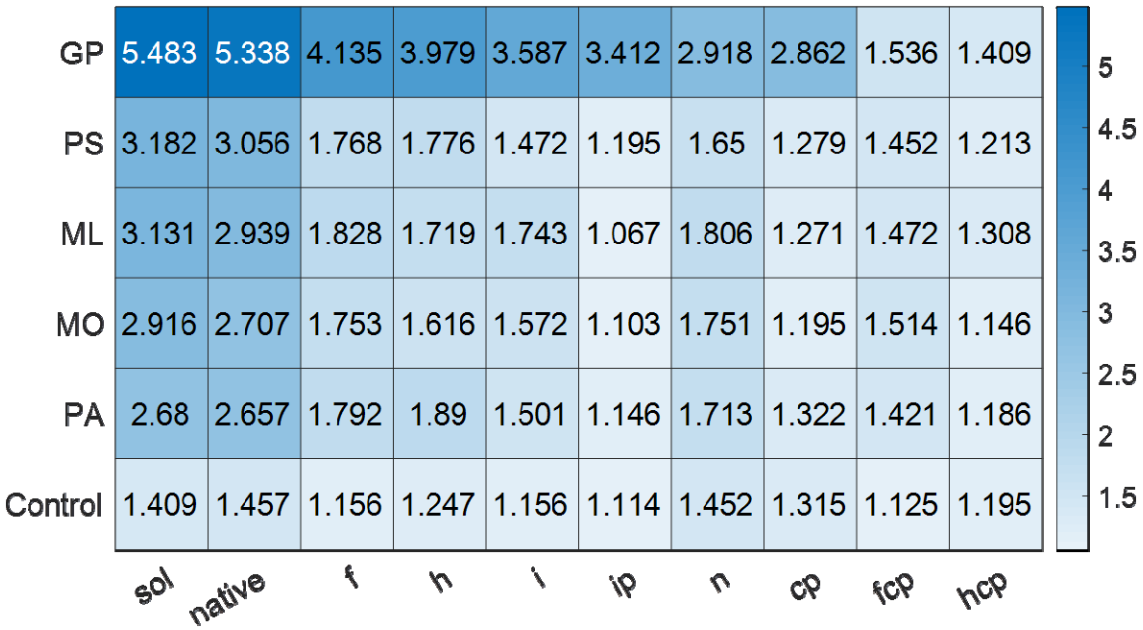


Figure 8 – A heat map representing the activity of luciferase from *Gaussia princeps* as well as four other putative luciferases from various organisms against a panel of 10 coelenterazine analogues as the luciferase substrate. The values in the heat map are on a log scale of the recorded luminescence (n=3). GP – *Gaussia princeps* (Uniprot Q9BLZ2), PS - *Pleuromamma scutullata* (Uniprot U3TIM6), ML - *Metridia longa* (Uniprot A0A0B4UFT5), MO - *Metridia okhotensis* (Uniprot H3JS09), PA - *Pleuromamma abdominalis* (Uniprot H3JS12).

Prototyping Enzymes

To further investigate the SHuffle strain's ability to quickly prototype proteins, we chose to express three enzymes (hevamine, endochitinase A (ChitA), and periplasmic AppA (which have 3, 7, and 4 disulfide bonds) also using our previously established minimal, rapid-order genetic template method (Dopp et al., 2019b). Hevamine is a class IIIa chitinase from the rubber tree *Hevea brasiliensis* and exhibits exochitinase activity (Kitaoku et al., 2015). ChitA is a class IV chitinase found primarily in developing maize kernels of *Zea mays* that exhibits endochitinase activity (Volpicella et al., 2017). Both chitinases are responsible for protecting the plants against fungal pathogens (Bokma et al., 2000; Bokma et al., 2002; Hawkins et al., 2015; Volpicella et al., 2017). AppA is a periplasmic protein native to *E. coli* that acts as a phytase (Berkmen et al., 2005). Phytases are an interesting class of enzymes because they break down phytate which is considered an antinutrient and reduces nutrient absorption from seeds (grains, beans, and nuts) in monogastric animals (Bitar and Reinhold, 1972; Zinin et al., 2004). Undigested phytates from monogastric livestock are also a large source of environmental phosphorous pollution in agriculture (Golovan et al., 2001; Zinin et al., 2004). To our knowledge, none of these enzymes have been expressed in an *E. coli*-based cell-free system.

The resulting products can be screened directly in the lysate without purification. Hevamine samples demonstrated 1.3x greater activity than no DNA controls over the length of the entire reaction (Fig 9A). The substrate chose for this reaction was 4-Methylumbelliferyl β-D-N,N′,N″-triacetylchitotrioside which is a fluorogenic used to characterize endochitinases. Hevamine's kinetics and preferred substrate have been published in previous literature (Bokma et al., 2000). In contrast, ChitA samples performed 2.4x better than the no DNA controls when comparing initial reaction rate (change in fluorescence over time during the first 60 sec, (Fig 9B). The substrate used for ChitA was 4-methylumbelliferyl N-acetyl-β-D-glucosaminide which is fluorogenic substrate used to characterize exochitinases. When screening the expressed AppA, we found the sample with enzyme reaches a kinetic rate of 0.041±.005 abs/min over the span of 5 minutes while the control (no DNA) only reaches 0.031±.005 abs/min (Fig 9C). The substrate used was p-nitrophenyl phosphate which is used to characterize alkaline and acid phosphatase activity.

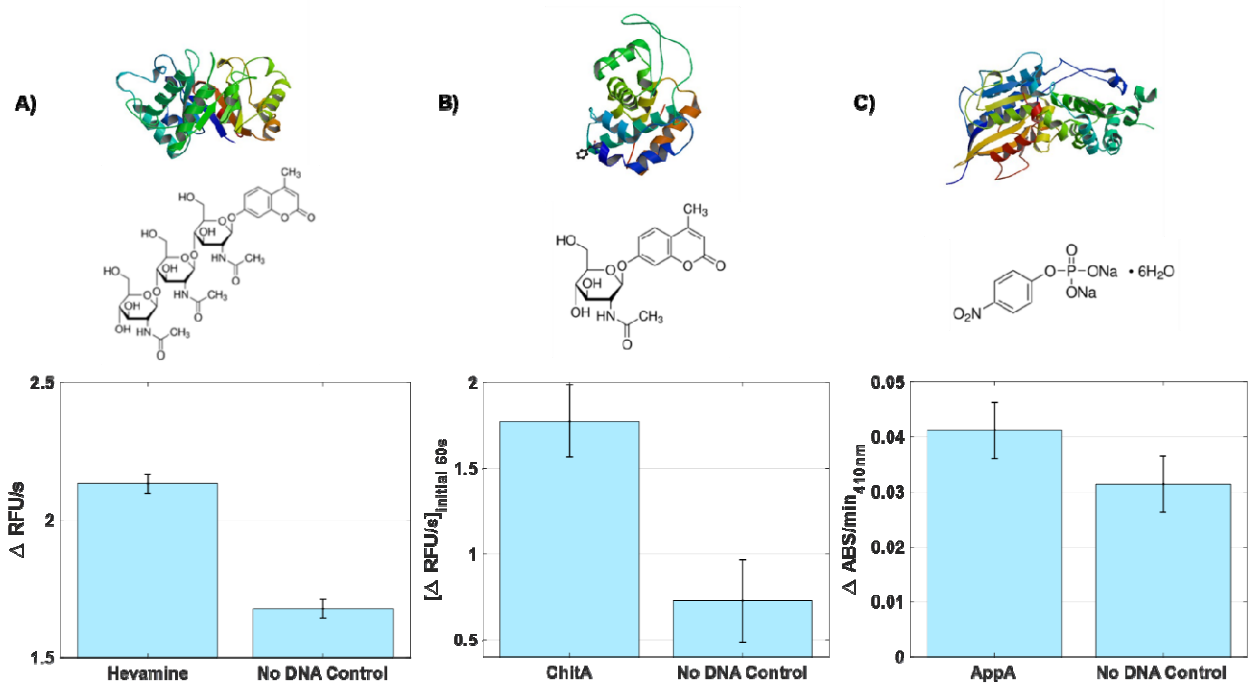


Figure 9 – The crystal structure, corresponding substrate, and activity of enzymes used in this work. (A) Total hevamine (RCSB PDB # 2HVM) activity on 4-Methylumbelliferyl β-D-N,N′,N″-triacetylchitotrioside (p-value = 1.10×10^{-5}) (n=3). (B) Initial rate (60 sec) of ChitA (RCSB PDB # 4MCK) activity on 4-Methylumbelliferyl N-acetyl-β-D-glucosaminide (p-value = 5.92×10^{-4}) (n=4). (C) The activity of AppA (RCSB PDB # 1DKL) on p-nitrophenyl phosphate, disodium salt hexahydrate (p-value = 0.034) (n=4).

Future Work

We have identified a few strategies to further improve the development and performance of this extract. One source of variability was the use of a reporter protein (Gluc) that requires a secondary assay which has careful, time-based steps to measure level of production (via luminescence). This is much less convenient (and accurate) than measuring the direct signal from a protein rapidly folding (like sfGFP). This could be addressed if a reporter protein is found/designed that can be detected while the reaction is running (such as a fluorescent protein which requires multiple disulfide bonds to produce its fluorescence). Additionally, further genome edits could be made on the SHuffle strain, such as a mutated RNaseE gene (rne131) that makes the Star strain uniquely qualified for expression from LET or RCA templates. Another way to improve expression from linear templates is to add GamS protein to the reaction which can improve expression to near plasmid levels (Ahn et al., 2005; Sitaraman et al., 2004; Sun et al., 2014; Yin et al., 2012). However, this work focuses on the advantage of not adding exogenous proteins. Another strategy of including material, would be the addition of short DNA

oligos containing six χ (chi) sites which has been shown to drastically improve yield when using linear DNA (Marshall et al., 2017). This could be included in supplement mix.

Conclusion

From the experiments performed, we conclude that the commercial cell line T7 SHuffle® can be used to create simple, inexpensive, functional extract for cell free expression of proteins with many disulfide bonds. In processing, this extract is lyophilized to drive off extra water used for the continuous homogenizer which increases the flexibility of extract storage and distribution. This extract is also more economical due to the scalable production method, simple steps and no need for exogenous protein supplementation). This extract can be used for rapid prototyping and design of new proteins from genomic data using minimal genetic templates that are amplified with isothermal rolling circle amplification. In the case of prototyping, where enough protein is needed to measure activity or desired properties, we have demonstrated sufficient level of protein expression. In cases where increased yield is needed, there are some general approaches to further improve yield that we have yet to apply to this strain, such as: increasing the surface area to volume ratio (Gregorio et al., 2019; Hong et al., 2015; Voloshin and Swartz, 2005), implementing machine learning to improve reaction composition (Borkowski et al., 2020), and conducting metabolomic studies to identify and optimize key metabolites (Miguez et al., 2019). However, even with these added improvements, the SHuffle strain may yet express less protein than other strains that are supplemented with exogenous T7RNAP and DsbC, simply due to exhausted machinery. Yet, the simple approach presented in this work can make significant impact at the prototyping phase where a large amount of extract is needed to test many sequence candidates and thus the cost of extract is imperative. Once desired enzymes, therapies, or protein-based materials are discovered, they can be scaled up in more efficient extract or using traditional cell-based methods. Finally, this simple approach which requires less specialized machinery, skill, and time, will be more readily accessible to groups wanting to prototype proteins with disulfide bonds.

Acknowledgements

The authors acknowledge Prof. Thomas Mansell for the investigative tip to look into T7 SHuffle® capabilities. NFR acknowledges the Black & Veatch Building a World of Difference Faculty Fellow in Engineering for partial funding and Iowa State University Startup Funds. We would also like to thank the Swartz group at Stanford for sharing their knowledge, Gluc plasmid, and KGK10 extract.

References

Ahn JH, Chu HS, Kim TW, Oh IS, Choi CY, Hahn GH, Park CG, Kim DM. 2005. Cell-free synthesis of recombinant proteins from PCR-amplified genes at a comparable productivity to that of plasmid-based reactions. *Biochem. Biophys. Res. Commun.* **338**:1346–1352. www.elsevier.com/locate/ybbrc.

Anton BP, Fomenkov A, Raleigh EA, Berkmen M. 2016. Complete Genome Sequence of the Engineered Escherichia coli SHuffle Strains and Their Wild-Type Parents. *Genome Announc* **4**:230–246. http://dx.doi.org/10.1126/science.1063143.

Baneyx F, Mujacic M. 2004. Recombinant protein folding and misfolding in Escherichia coli. *Nat. Biotechnol.*

Berkmen M, Boyd D, Beckwith J. 2005. The nonconsecutive disulfide bond of Escherichia coli phytase (AppA) renders it dependent on the protein-disulfide isomerase, DsbC. *J. Biol. Chem.* **280**:11387–11394. http://www.jbc.org/.

Bitar K, Reinhold JG. 1972. Phytase and alkaline phosphatase activities in intestinal mucosae of rat, chicken, calf, and man. *BBA - Enzymol.* **268**:442–452.

Bokma E, Barends T, Terwisscha Van Scheltinga AC, Dijkstra BW, Beintema JJ. 2000. Enzyme kinetics of hevamine, a chitinase from the rubber tree Hevea brasiliensis. *FEBS Lett.* **478**:119–122. http://doi.wiley.com/10.1016/S0014-5793%2800%2901833-0.

Bokma E, Rozeboom HJ, Sibbald M, Dijkstra BW, Beintema JJ. 2002. Expression and characterization of active site mutants of hevamine, a chitinase from the rubber tree Hevea brasiliensis. *Eur. J. Biochem.* **269**:893–901.

Borkowski O, Koch M, Zettor A, Pandi A, Batista AC, Soudier P, Faulon J-L. 2020. Large scale active-learning-guided exploration for in vitro protein production optimization. *Nat. Commun.* **11**:1872. http://www.nature.com/articles/s41467-020-15798-5.

Burgess RR. 2009. Refolding Solubilized Inclusion Body Proteins. In: . *Methods Enzymol.* Academic Press Inc., Vol. 463, pp. 259–282.

Cai Q, Hanson JA, Steiner AR, Tran C, Masikat MR, Chen R, Zawada JF, Sato AK, Hallam TJ, Yin G. 2015. A simplified and robust protocol for immunoglobulin expression in Escherichia coli cell-free protein synthesis systems. *Biotechnol. Prog.* **31**:823–831.

Calhoun KA, Swartz JR. 2006. Total amino acid stabilization during cell-free protein synthesis reactions. *J. Biotechnol.* **123**:193–203. https://ac.els-cdn.com/S0168165605007479/1-s2.0-S0168165605007479-main.pdf?_tid=18ab9d82-a1a9-4b12-8964-d8f41f8ddc6a&acdnat=1531942690_b2d1171c10f1dbf1a5fb5afd85e1c352.

Didovyk A, Tonooka T, Tsimring L, Hasty J. 2017. Rapid and Scalable Preparation of Bacterial Lysates for Cell-Free Gene Expression. *ACS Synth. Biol.* **6**:2198–2208. https://pubs-acs-org.proxy.lib.iastate.edu/doi/pdfplus/10.1021/acssynbio.7b00253.

Dikici E, Qu X, Rowe L, Millner L, Logue C, Deo SK, Ensor M, Daunert S. 2009. Aequorin variants with improved bioluminescence properties. *Protein Eng. Des. Sel.* **22**:243–248.

Dopp JL, Jo YR, Reuel NF. 2019a. Methods to reduce variability in E. Coli-based cell-free protein expression experiments. *Synth. Syst. Biotechnol.* **4**:204–211.

Dopp JL, Reuel NF. 2018. Process optimization for scalable E. coli extract preparation for cell-free protein synthesis. *Biochem. Eng. J.* **138**:21–28. https://linkinghub.elsevier.com/retrieve/pii/S1369703X18302171.

Dopp JL, Tamiev DD, Reuel NF. 2018. Cell-free supplement mixtures: Elucidating the history and biochemical utility of additives used to support in vitro protein synthesis in E. coli extract. *Biotechnol. Adv.* **37**:246–258. https://linkinghub.elsevier.com/retrieve/pii/S0734975018302040.

Dopp JL, Rothstein SM, Mansell TJ, Reuel NF. 2019b. Rapid prototyping of proteins: Mail order gene fragments to assayable proteins within 24 hours. *Biotechnol. Bioeng.* **116**:667–676. http://doi.wiley.com/10.1002/bit.26912.

Fass D. 2012. Disulfide Bonding in Protein Biophysics. *Annu. Rev. Biophys.* **41**:63–79. www.annualreviews.org.

Ferreira SLC, Bruns RE, Ferreira HS, Matos GD, David JM, Brandão GC, da Silva EGP, Portugal LA, dos Reis PS, Souza AS, dos Santos WNL. 2007. Box-Behnken design: An alternative for the optimization of analytical methods. *Anal. Chim. Acta*. https://ac.els-cdn.com/S0003267007011671/1-s2.0-S0003267007011671-main.pdf?_tid=b8c02cd4-

606b-4c5b-a467-754fcf968ed9&acdnat=1526077589_d4ff727240d806a9d55df53ac1ab8dab.
Frey S, Haslbeck M, Hainzl O, Buchner J. 2008. Synthesis and characterization of a functional intact IgG in a prokaryotic cell-free expression system. *Biol. Chem.* **389**:37–45.
GeneFrontier. PURExpress® DsbC Set. https://www.cosmobiousa.com/products/ds-supplement.
Goerke AR, Loening AM, Gambhir SS, Swartz JR. 2008. Cell-free metabolic engineering promotes high-level production of bioactive Gaussia princeps luciferase. *Metab. Eng.* **10**:187–200.
Goerke AR, Swartz JR. 2007. Development of cell-free protein synthesis platforms for disulfide bonded proteins. *Biotechnol. Bioeng.* **99**:351–367.
Golovan SP, Hayes MA, Phillips JP, Forsberg CW. 2001. Transgenic mice expressing bacterial phytase as a mode for phosphorus pollution control. *Nat. Biotechnol.* **19**:429–433.
Green AA, Kim J, Ma D, Silver PA, Collins JJ, Yin P. 2017. Complex cellular logic computation using ribocomputing devices. *Nature* **548**:117–121. http://www.nature.com.proxy.lib.iastate.edu/articles/nature23271.pdf.
Gregorio NE, Levine MZ, Oza JP. 2019. A User's Guide to Cell-Free Protein Synthesis. *Methods Protoc.* **2**:24. https://www.mdpi.com/2409-9279/2/1/24.
Groff D, Armstrong S, Rivers PJ, Zhang J, Yang J, Green E, Rozzelle J, Liang S, Kittle JD, Steiner AR, Baliga R, Thanos CD, Hallam TJ, Sato AK, Yam AY. 2014. Engineering toward a bacterial "endoplasmic reticulum" for the rapid expression of immunoglobulin proteins. *MAbs* **6**:671–678. https://www.tandfonline.com/action/journalInformation?journalCode=kmab20.
Haataja L, Manickam N, Soliman A, Tsai B, Liu M, Arvan P. 2016. Disulfide mispairing during proinsulin folding in the endoplasmic reticulum. *Diabetes* **65**:1050–1060.
Hamid R, Khan MA, Ahmad M, Ahmad MM, Abdin MZ, Musarrat J, Javed S. 2013. Chitinases: An update. *J. Pharm. Bioallied Sci.*
Hawkins LK, Mylroie JE, Oliveira DA, Smith JS, Ozkan S, Windham GL, Williams WP, Warburton ML. 2015. Characterization of the maize chitinase genes and their effect on Aspergillus flavus and aflatoxin accumulation resistance. *PLoS One* **10**. www.usaid.gov.
Hong SH, Kwon Y-C, Martin RW, Des Soye BJ, de Paz AM, Swonger KN, Ntai I, Kelleher NL, Jewett MC. 2015. Improving cell-free protein synthesis through genome engineering of Escherichia coli lacking release factor 1. *ChemBioChem* **16**:844–53. http://dx.doi.org/10.1002/cbic.201402708.
Inaba K. 2010. Structural basis of protein disulfide bond generation in the cell. *Genes to Cells*.
Jaroentomeechai T, Stark JC, Natarajan A, Glasscock CJ, Yates LE, Hsu KJ, Mrksich M, Jewett MC, DeLisa MP. 2018. Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nat. Commun.* **9**:2686. www.nature.com/naturecommunications.
Junnila RK, Kopchick JJ. 2013. Significance of the disulphide bonds of human growth hormone. *Endokrynol. Pol.*
Ke N, Berkmen M. 2014. Production of disulfide-bonded proteins in Escherichia coli. *Curr. Protoc. Mol. Biol.* **2014**:16.1B.21-16.1B.21.
Kim DM, Swartz JR. 2004. Efficient Production of a Bioactive, Multiple Disulfide-Bonded Protein Using Modified Extracts of Escherichia coli. *Biotechnol. Bioeng.* **85**:122–129.
Kitaoku Y, Umemoto N, Ohnuma T, Numata T, Taira T, Sakuda S, Fukamizo T. 2015. A class III chitinase without disulfide bonds from the fern, Pteris ryukyuensis: crystal structure and ligand-binding studies. *Planta* **242**:895–907. https://link-springer-com.proxy.lib.iastate.edu/content/pdf/10.1007%2Fs00425-015-2330-4.pdf.
Knapp KG, Goerke AR, Swartz JR. 2007. Cell-free synthesis of proteins that require disulfide bonds using glucose as an energy source. *Biotechnol. Bioeng.* **97**:901–908.
Kwon YC, Jewett MC. 2015. High-throughput preparation methods of crude extract for robust

cell-free protein synthesis. *Sci. Rep.* **5**.
Levine MZ, Gregorio NE, Jewett MC, Watts KR, Oza JP. 2019. Escherichia coli-Based Cell-Free Protein Synthesis: Protocols for a robust, flexible, and accessible platform technology. *J. Vis. Exp.*:e58882. https://www.jove.com/video/58882/escherichia-coli-based-cell-free-protein-synthesis-protocols-for.
Li T, Bourgeois JP, Celli S, Glacial F, Le Sourd AM, Mecheri S, Weksler B, Romero I, Couraud PO, Rougeon F, Lafaye P. 2012. Cell-penetrating anti-GFAP VHH and corresponding fluorescent fusion protein VHH-GFP spontaneously cross the blood-brain barrier and specifically recognize astrocytes: Application to brain imaging. *FASEB J.* **26**:3969–3979. http://www.fasebj.org/doi/pdf/10.1096/fj.11-201384.
Liu D V., Zawada JF, Swartz JR. 2005. Streamlining Escherichia Coli S30 extract preparation for economical cell-free protein synthesis. *Biotechnol. Prog.* **21**:460–465.
Liu M, Blinn C, McLeod SM, Wiseman JW, Newman J V., Fisher SL, Walkup GK. 2014. Secreted Gaussia princeps Luciferase as a Reporter of Escherichia coli Replication in a Mouse Tissue Cage Model of Infection. Ed. Patrick M. Schlievert. *PLoS One* **9**:e90382. https://dx.plos.org/10.1371/journal.pone.0090382.
Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M. 2012. SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microb. Cell Fact.* **11**:753. http://microbialcellfactories.biomedcentral.com/articles/10.1186/1475-2859-11-56.
Lu Y, Welsh JP, Swartz JR. 2014. Production and stabilization of the trimeric influenza hemagglutinin stem domain for potentially broadly protective influenza vaccines. *Proc. Natl. Acad. Sci.* **111**:125–130. http://www.pnas.org/cgi/doi/10.1073/pnas.1308701110.
Mamathambika BS, Bardwell JC. 2008. Disulfide-Linked Protein Folding Pathways. *Annu. Rev. Cell Dev. Biol.* **24**:211–235. http://www.annualreviews.org/doi/10.1146/annurev.cellbio.24.110707.175333.
Marshall R, Maxwell CS, Collins SP, Beisel CL, Noireaux V. 2017. Short DNA containing χ sites enhances DNA stability and gene expression in E. coli cell-free transcription-translation systems. *Biotechnol. Bioeng.*
Matsuda T, Watanabe S, Kigawa T. 2013. Cell-free synthesis system suitable for disulfide-containing proteins. *Biochem. Biophys. Res. Commun.* **431**:296–301. http://dx.doi.org/10.1016/j.bbrc.2012.12.107.
Mcnerney MP, Zhang Y, Steppe P, Silverman AD, Jewett MC, Styczynski MP. 2019. Point-of-care biomarker quantification enabled by sample-specific calibration. *Sci. Adv.* **5**:eaax4473. http://advances.sciencemag.org/.
Michel-Reydellet N, Calhoun K, Swartz J. 2004. Amino acid stabilization for cell-free protein synthesis by modification of the Escherichia coli genome. *Metab. Eng.* **6**:197–203.
Michel E, Wüthrich K. 2012a. High-yield Escherichia coli-based cell-free expression of human proteins. *J. Biomol. NMR* **53**:43–51.
Michel E, Wüthrich K. 2012b. Cell-free expression of disulfide-containing eukaryotic proteins for structural biology. *FEBS J.* **279**:3176–3184.
Middelberg APJ. 2002. Preparative protein refolding. *Trends Biotechnol.* Elsevier Current Trends.
Miguez AM, McNerney MP, Styczynski MP. 2019. Metabolic Profiling of Escherichia coli-Based Cell-Free Expression Systems for Process Optimization. *Ind. Eng. Chem. Res.* **58**:22472–22482.
Morse D, Tannous BA. 2012. A water-soluble coelenterazine for sensitive in vivo imaging of coelenterate luciferases. *Mol. Ther.* Elsevier.
Murakami S, Matsumoto R, Kanamori T. 2019. Constructive approach for synthesis of a functional IgG using a reconstituted cell-free protein synthesis system. *Sci. Rep.* **9**:671. http://www.nature.com/articles/s41598-018-36691-8.

NEB. PURExpress® Disulfide Bond Enhancer. https://www.neb.com/products/e6820-purexpress-disulfide-bond-enhancer#Product Information.

Østergaard H, Henriksen A, Hansen FG, Winther JR. 2001. Shedding light on disulfide bond formation: Engineering a redox switch in green fluorescent protein. *EMBO J.* **20**:5853–5862.

Pardee K, Green AA, Takahashi MK, Braff D, Lambert G, Lee JW, Ferrante T, Ma D, Donghia N, Fan M, Daringer NM, Bosch I, Dudley DM, O'Connor DH, Gehrke L, Collins JJ. 2016a. Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* **165**:1255–1266.

Pardee K, Slomovic S, Nguyen PQ, Lee JW, Donghia N, Burrill D, Ferrante T, McSorley FR, Furuta Y, Vernet A, Lewandowski M, Boddy CN, Joshi NS, Collins JJ. 2016b. Portable, On-Demand Biomolecular Manufacturing. *Cell* **167**:248-259.e12.

Patil NA, Tailhades J, Hughes RA, Separovic F, Wade JD, Hossain MA. 2015. Cellular disulfide bond formation in bioactive peptides and proteins. *Int. J. Mol. Sci.* www.mdpi.com/journal/ijms.

Robinson MP, Ke N, Lobstein J, Peterson C, Szkodny A, Mansell TJ, Tuckey C, Riggs PD, Colussi PA, Noren CJ, Taron CH, Delisa MP, Berkmen M. 2015. Efficient expression of full-length antibodies in the cytoplasm of engineered bacteria. *Nat. Commun.* **6**.

Saeed AFUH, Wang R, Ling S, Wang S. 2017. Antibody engineering for pursuing a healthier future. *Front. Microbiol.* https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5368232/pdf/fmicb-08-00495.pdf.

Salehi ASM, Smith MT, Bennett AM, Williams JB, Pitt WG, Bundy BC. 2016. Cell-free protein synthesis of a cytotoxic cancer therapeutic: Onconase production and a just-add-water cell-free system. *Biotechnol. J.* **11**:274–281.

Salehi ASM, Yang SO, Earl CC, Shakalli Tang MJ, Porter Hunt J, Smith MT, Wood DW, Bundy BC. 2018. Biosensing estrogenic endocrine disruptors in human blood and urine: A RAPID cell-free protein synthesis approach. *Toxicol. Appl. Pharmacol.* **345**:19–25. https://linkinghub.elsevier.com/retrieve/pii/S0041008X18300620.

Schoborg JA, Hershewe JM, Stark JC, Kightlinger W, Kath JE, Jaroentomeechai T, Natarajan A, DeLisa MP, Jewett MC. 2018. A cell-free platform for rapid synthesis and testing of active oligosaccharyltransferases. *Biotechnol. Bioeng.* **115**:739–750. http://doi.wiley.com/10.1002/bit.26502.

Siddiquee R, Choi S, Lam S, Wang P, Qi R, Otting G, Sunde M, Kwan A. 2020. Cell-free expression of natively folded hydrophobins. *Protein Expr. Purif.* **170**:105591. https://www.sciencedirect.com/science/article/pii/S1046592819305613?dgcid=rss_sd_all&utm_source=researcher_app&utm_medium=referral&utm_campaign=RESR_MRKT_Researcher_inbound.

Siegal-Gaskins D, Tuza ZA, Kim J, Noireaux V, Murray RM. 2014. Gene circuit performance characterization and resource usage in a cell-free "breadboard." *ACS Synth. Biol.* **3**:416–425.

Siller E, DeZwaan DC, Anderson JF, Freeman BC, Barral JM. 2010. Slowing Bacterial Translation Speed Enhances Eukaryotic Protein Folding Efficiency. *J. Mol. Biol.* **396**:1310–1318. www.sciencedirect.com.

Sitaraman K, Esposito D, Klarmann G, Le Grice SF, Hartley JL, Chatterjee DK. 2004. A novel cell-free protein synthesis system. *J. Biotechnol.* **110**:257–263.

Snouwaert JN, Leebeek FWG, Fowlkes DM. 1991. Role of disulfide bonds in biologic activity of human interleukin-6. *J. Biol. Chem.* **266**:23097–23102.

Soltani M, Davis BR, Ford H, Andrew J, Nelson D, Bundy BC. 2018. Reengineering cell-free protein synthesis as a biosensor: Biosensing with transcription, translation, and protein-folding. *Biochem. Eng. J.* **138**:165–171. https://doi.org/10.1016/j.bej.2018.06.014.

Sun ZZ, Hayes CA, Shin J, Caschera F, Murray RM, Noireaux V. 2013. Protocols for

Implementing an Escherichia coli Based TX-TL Cell-Free Expression System for Synthetic Biology. *J. Vis. Exp.* http://www.jove.com/video/50762/protocols-for-implementing-an-escherichia-coli-based-tx-tl-cell-free.

Sun ZZ, Yeung E, Hayes CA, Noireaux V, Murray RM. 2014. Linear DNA for rapid prototyping of synthetic biological circuits in an escherichia coli based TX-TL cell-free system. *ACS Synth. Biol.* **3**:387–397. https://pubs-acs-org.proxy.lib.iastate.edu/doi/pdfplus/10.1021/sb400131a.

Takahashi MK, Tan X, Dy AJ, Braff D, Akana RT, Furuta Y, Donghia N, Ananthakrishnan A, Collins JJ. 2018. A low-cost paper-based synthetic biology platform for analyzing gut microbiota and host biomarkers. *Nat. Commun.* **9**:3347. http://www.nature.com/articles/s41467-018-05864-4.

Voloshin AM, Swartz JR. 2005. Efficient and scalable method for scaling up cell free protein synthesis in batch mode. *Biotechnol. Bioeng.* **91**:516–521.

Volpicella M, Leoni C, Fanizza I, Distaso M, Leoni G, Farioli L, Naumann T, Pastorello E, Ceci LR. 2017. Characterization of maize chitinase-A, a tough allergenic molecule. *Allergy Eur. J. Allergy Clin. Immunol.* **72**:1423–1429. http://doi.wiley.com/10.1111/all.13164.

Whitesides GM. 2010. Toward a science of simplicity. *TEDEd*. https://ed.ted.com/lessons/toward-a-science-of-simplicity-george-whitesides.

Whitesides GM. 2013. Cool, or simple and cheap? Why not both? *Lab Chip*. Royal Society of Chemistry.

Wilding KM, Hunt JP, Wilkerson JW, Funk PJ, Swensen RL, Carver WC, Christian ML, Bundy BC. 2018. Endotoxin-Free E. coli-Based Cell-Free Protein Synthesis: Pre-Expression Endotoxin Removal Approaches for on-Demand Cancer Therapeutic Production. *Biotechnol. J.*, September 20. http://doi.wiley.com/10.1002/biot.201800271.

Wong JWH, Ho SYW, Hogg PJ. 2011. Disulfide bond acquisition through eukaryotic protein evolution. *Mol. Biol. Evol.* **28**:327–334. https://academic.oup.com/mbe/article-abstract/28/1/327/980888.

Wu N, Rathnayaka T, Kuroda Y. 2015. Bacterial expression and re-engineering of Gaussia princeps luciferase and its use as a reporter protein. *BBA - Proteins Proteomics* **1854**:1392–1399. http://dx.doi.org/10.1016/j.bbapap.2015.05.008.

Yang WC, Patel KG, Wong HE, Swartz JR. 2012. Simplifying and streamlining Escherichia coli-based cell-free protein synthesis. *Biotechnol. Prog.* **28**:413–420.

Yin G, Garces ED, Yang J, Zhang J, Tran C, Steiner AR, Roos C, Bajad S, Hudak S, Penta K, Zawada J, Pollitt S, Murray CJ. 2012. Aglycosylated antibodies and antibody fragments produced in a scalable in vitro transcription-translation system. *MAbs* **4**:217–225.

Yin G, Stephenson HT, Yang J, Li X, Armstrong SM, Heibeck TH, Tran C, Masikat MR, Zhou S, Stafford RL, Yam AY, Lee J, Steiner AR, Gill A, Penta K, Pollitt S, Baliga R, Murray CJ, Thanos CD, McEvoy LM, Sato AK, Hallam TJ. 2017. RF1 attenuation enables efficient non-natural amino acid incorporation for production of homogeneous antibody drug conjugates. *Sci. Rep.* **7**.

Yin G, Swartz JR. 2004. Enhancing Multiple Disulfide Bonded Protein Folding in a Cell-Free System. *Biotechnol. Bioeng.* **86**:188–195. http://doi.wiley.com/10.1002/bit.10827.

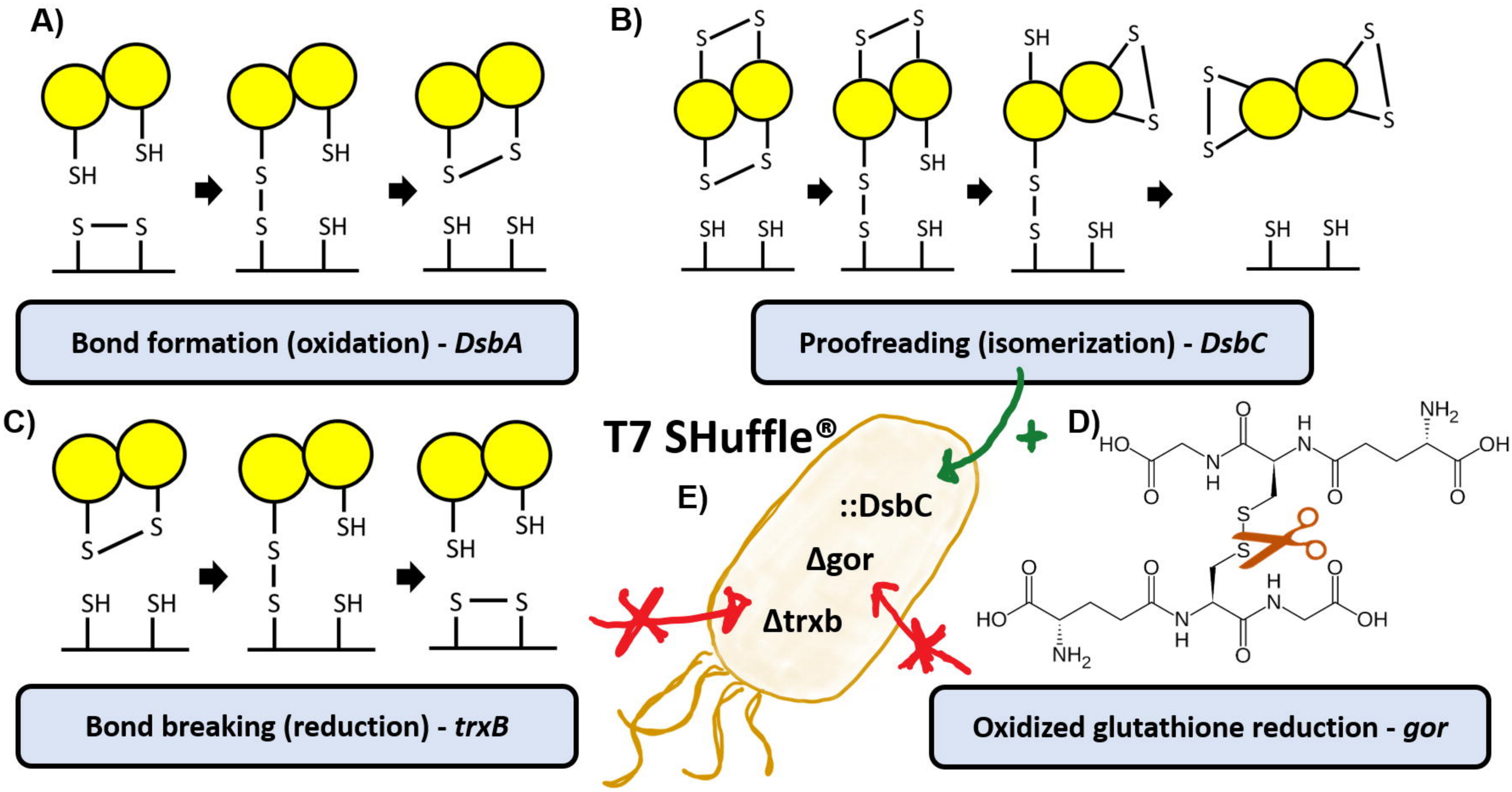
Yu T, Laird JR, Prescher JA, Thorpe C. 2018. Gaussia princeps luciferase: a bioluminescent substrate for oxidative protein folding. *Protein Sci.* **27**:1509–1517. http://doi.wiley.com/10.1002/pro.3433.

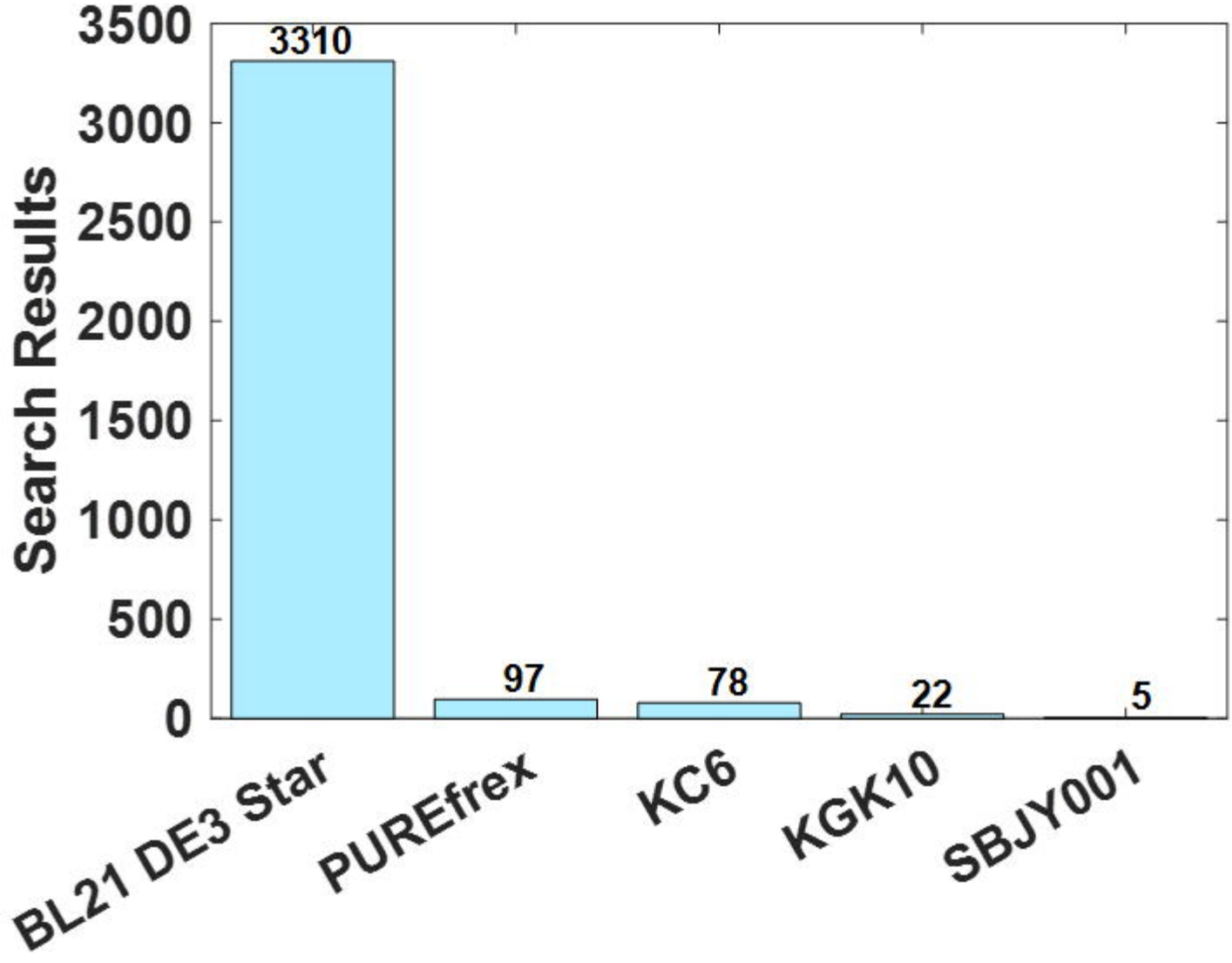
Zawada JF, Yin G, Steiner AR, Yang J, Naresh A, Roy SM, Gold DS, Heinsohn HG, Murray CJ. 2011. Microscale to manufacturing scale-up of cell-free cytokine production-a new approach for shortening protein production development timelines. *Biotechnol. Bioeng.* **108**:1570–1578.

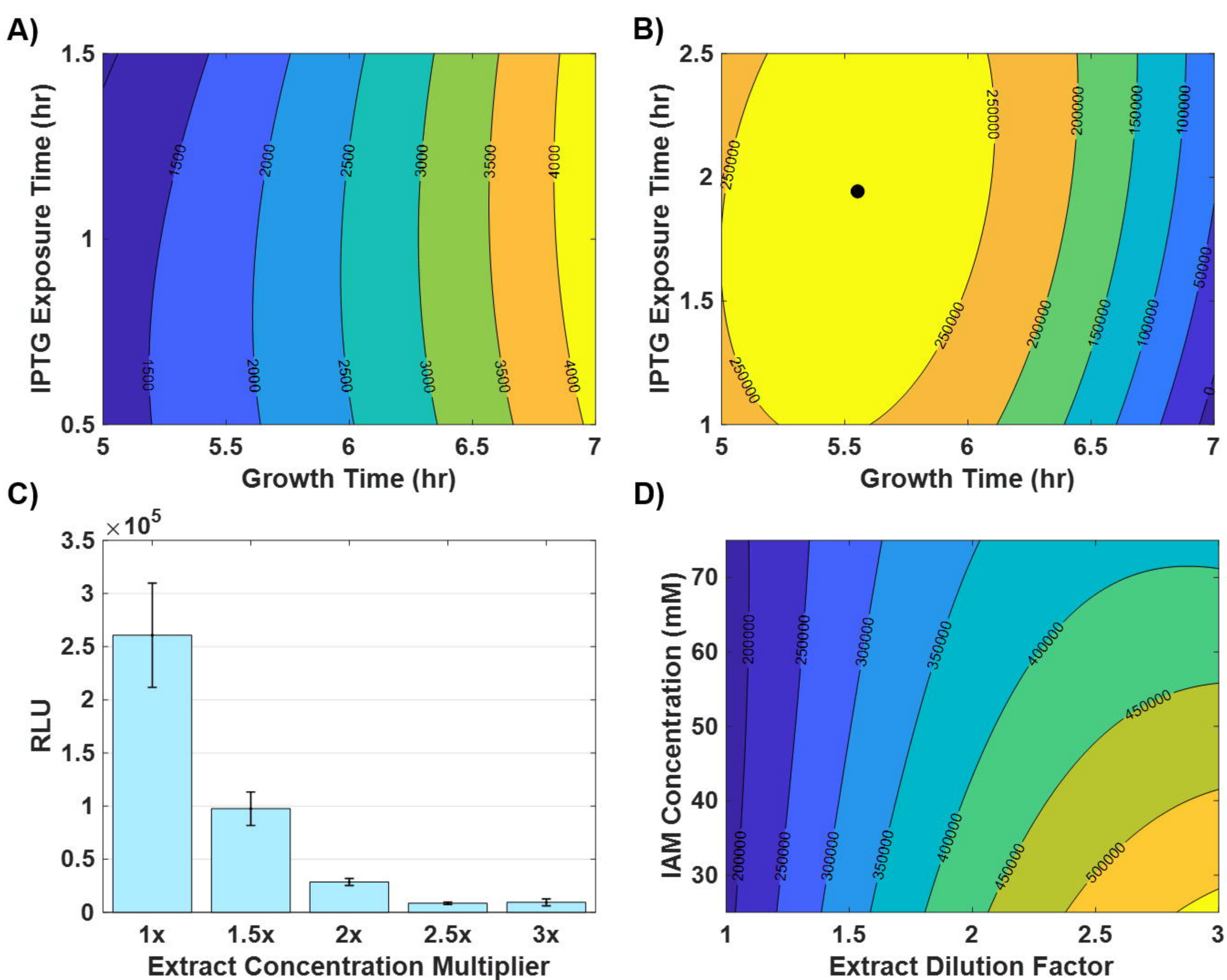
Zawada J, Swartz J. 2005. Maintaining rapid growth in moderate-density Escherichia coli fermentations. *Biotechnol. Bioeng.* **89**:407–415.

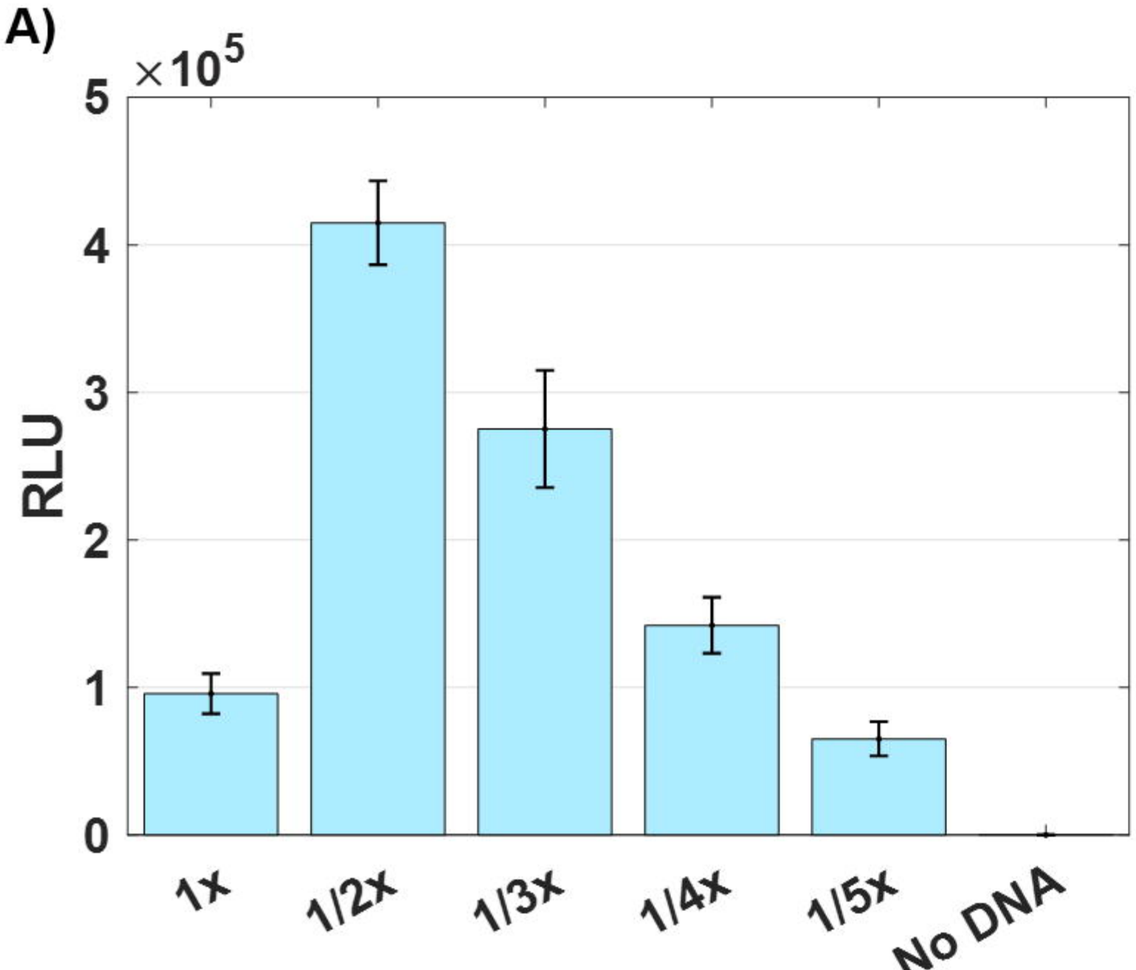
Zhao H, Doyle TC, Wong RJ, Cao Y, Stevenson DK, Piwnica-Worms D, Contag CH. 2004. Characterization of coelenterazine analogs for measurements of renilla luciferase activity in live cells and living animals. *Mol. Imaging* **3**:43–54. http://www.ncbi.nlm.nih.gov/pubmed/15142411.

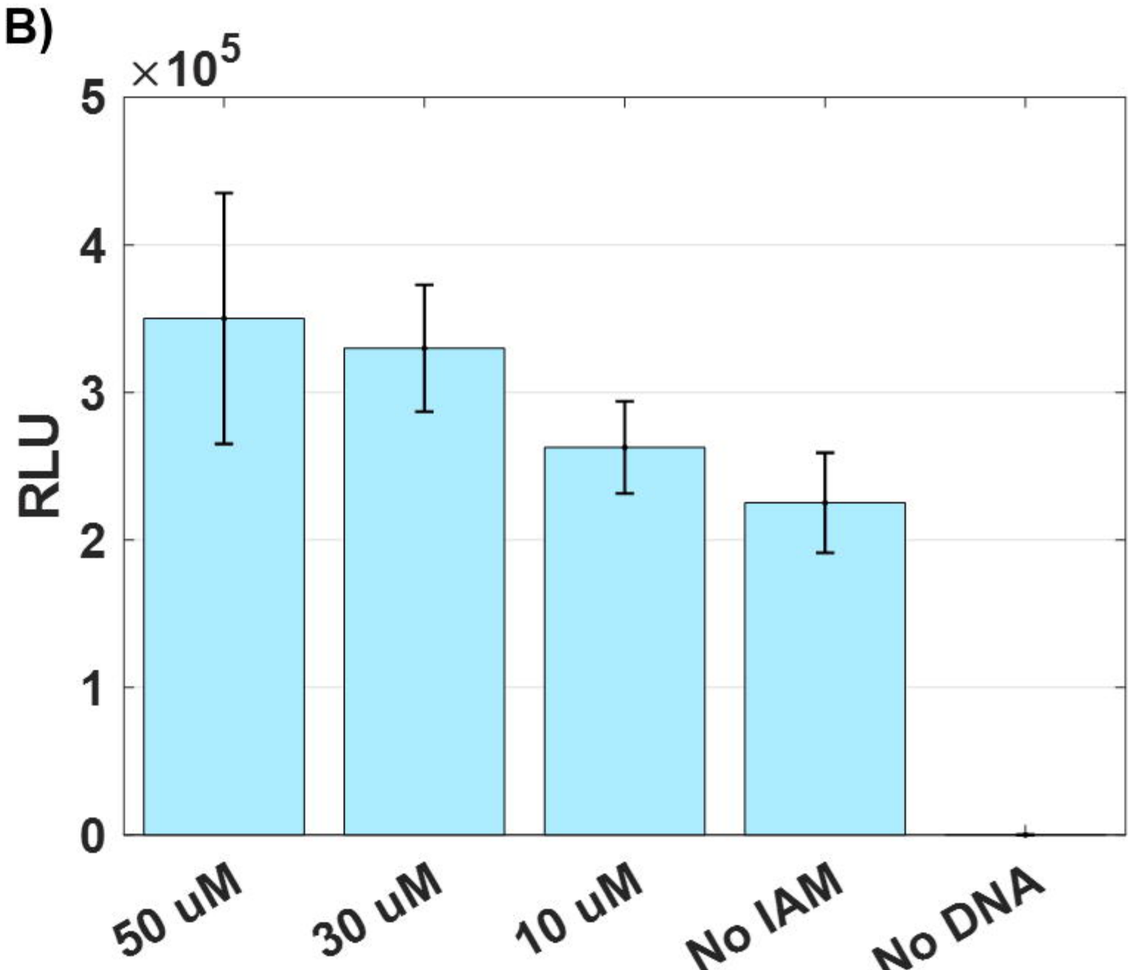
Zinin N V, Serkina A V, Gelfand MS, Shevelev AB, Sineoky SP. 2004. Gene cloning, expression and characterization of novel phytase from Obesumbacterium proteus. *FEMS Microbiol. Lett.* **236**:283–290. https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.2004.tb09659.x.

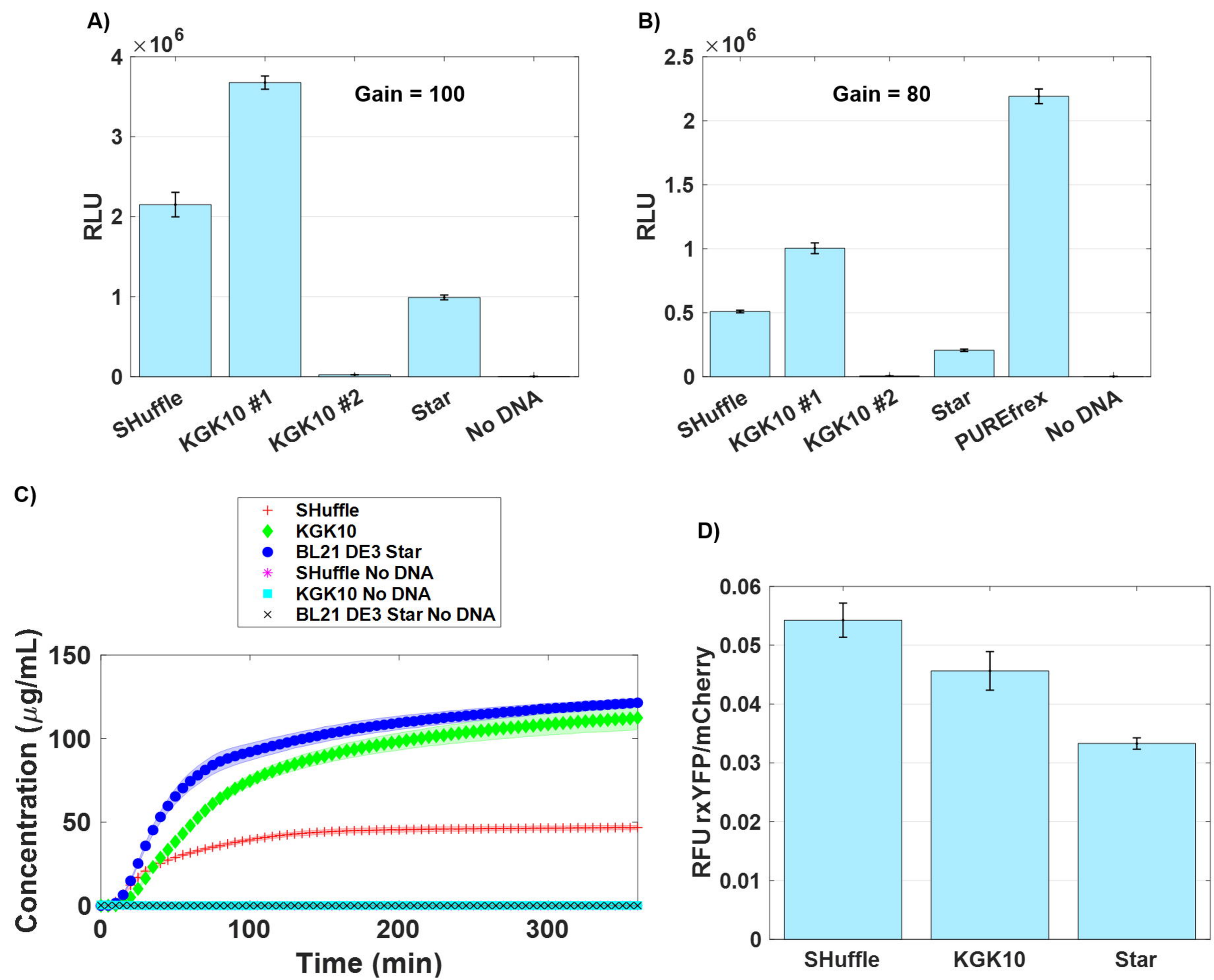

A)
Bond formation (oxidation) - DsbA
B)
Proofreading (isomerization) - DsbC
C)
Bond breaking (reduction) - trxB
T7 SHuffle®
E)
::DsbC
Δgor
Δtrxb
D)
Oxidized glutathione reduction - gor

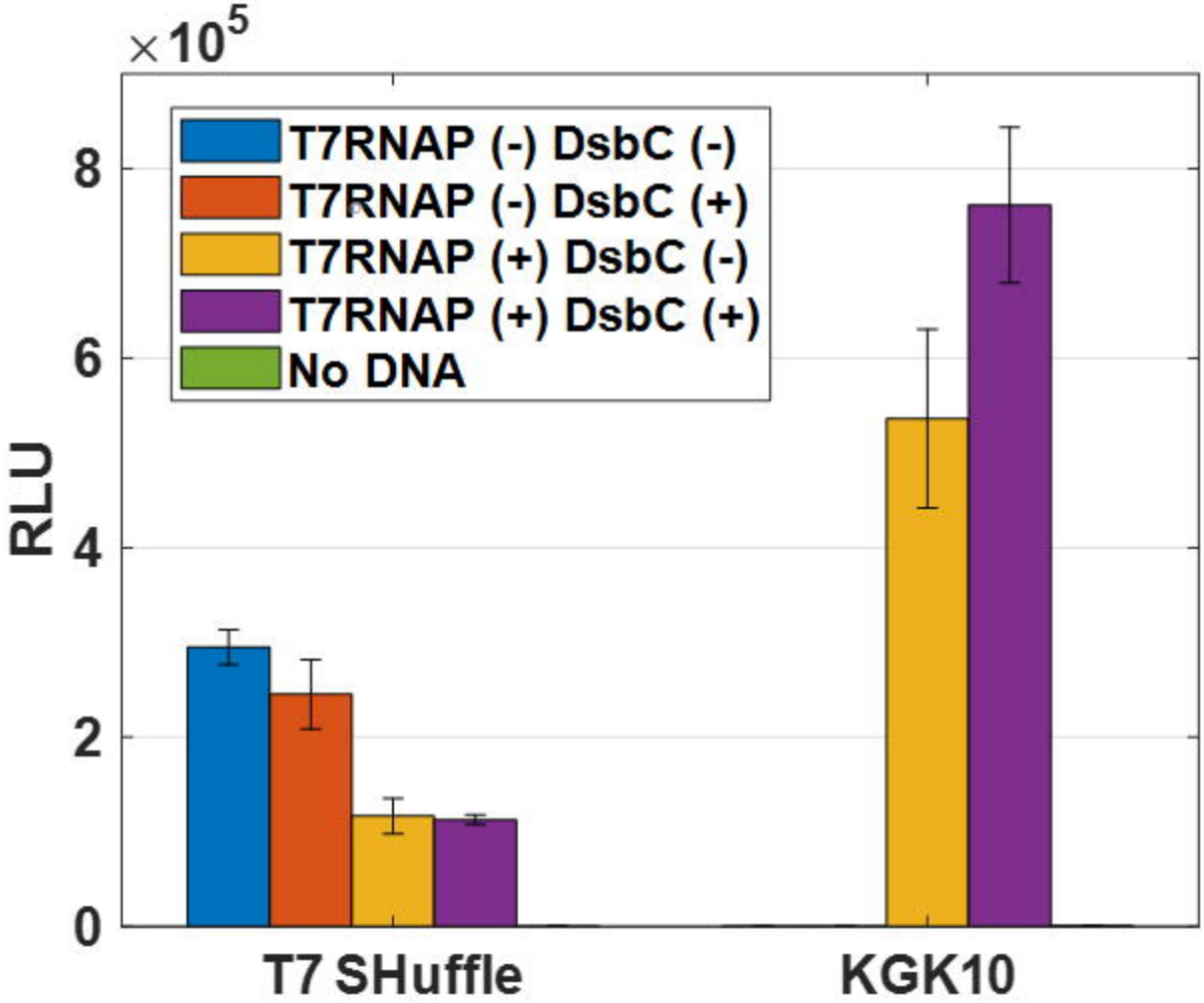

Search Results
3500
3000
2500
2000
1500
1000
500
0
3310
97
78
22
5
BL21 DE3 Star
PURExpress
KC6
KGK10
SBJY001

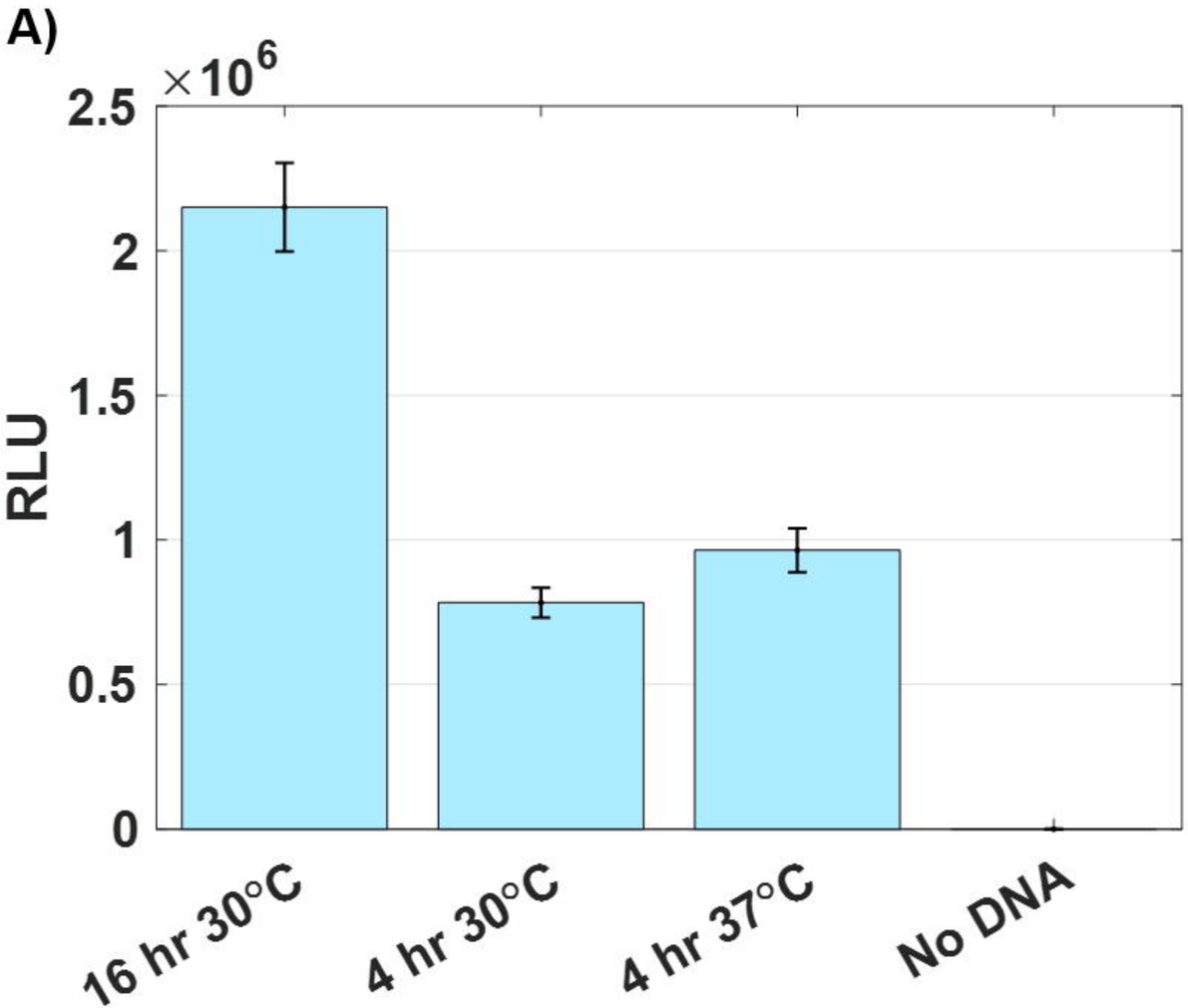

A)
IPTG Exposure Time (hr)
Growth Time (hr)
1500
2000
2500
3000
3500
4000
B)
IPTG Exposure Time (hr)
Growth Time (hr)
250000
200000
150000
100000
50000
0
C)
×10^5
RLU
1x
1.5x
2x
2.5x
3x
Extract Concentration Multiplier
D)
IAM Concentration (mM)
Extract Dilution Factor
200000
250000
300000
350000
400000
450000
500000

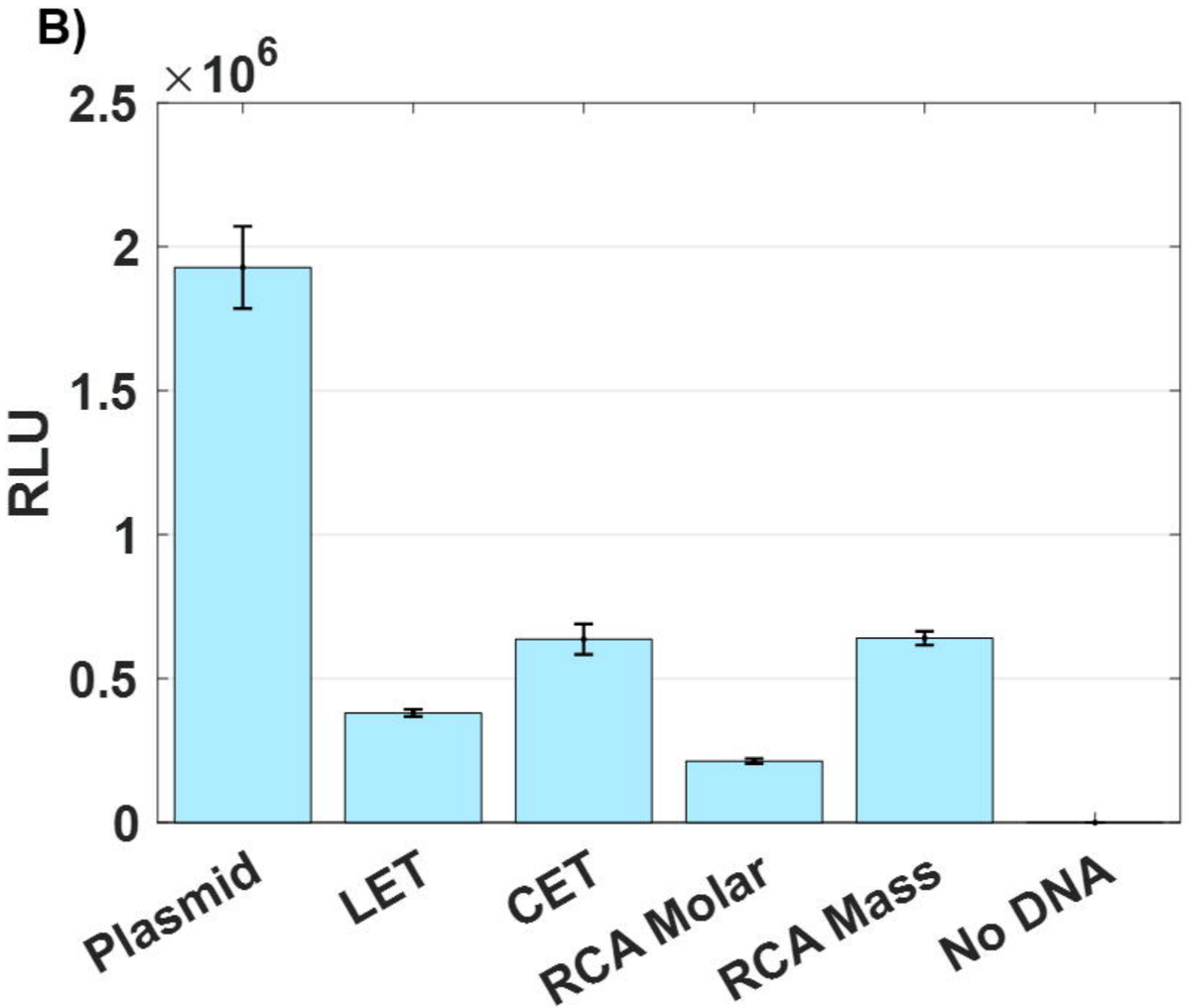

A)
×10^5
5
4
3
2
1
0
RLU
1x
1/2x
1/3x
1/4x
1/5x
No DNA


B)
×10^5
5
4
3
2
1
0
RLU
50 uM
30 uM
10 uM
No IAM
No DNA


A)
×10^6
Gain = 100
RLU
SHuffle
KGK10 #1
KGK10 #2
Star
No DNA
B)
×10^6
Gain = 80
RLU
SHuffle
KGK10 #1
KGK10 #2
Star
PURExpress
No DNA
C)
SHuffle
KGK10
BL21 DE3 Star
SHuffle No DNA
KGK10 No DNA
BL21 DE3 Star No DNA
Concentration (μg/mL)
Time (min)
D)
RFU rxYFP/mCherry
SHuffle
KGK10
Star

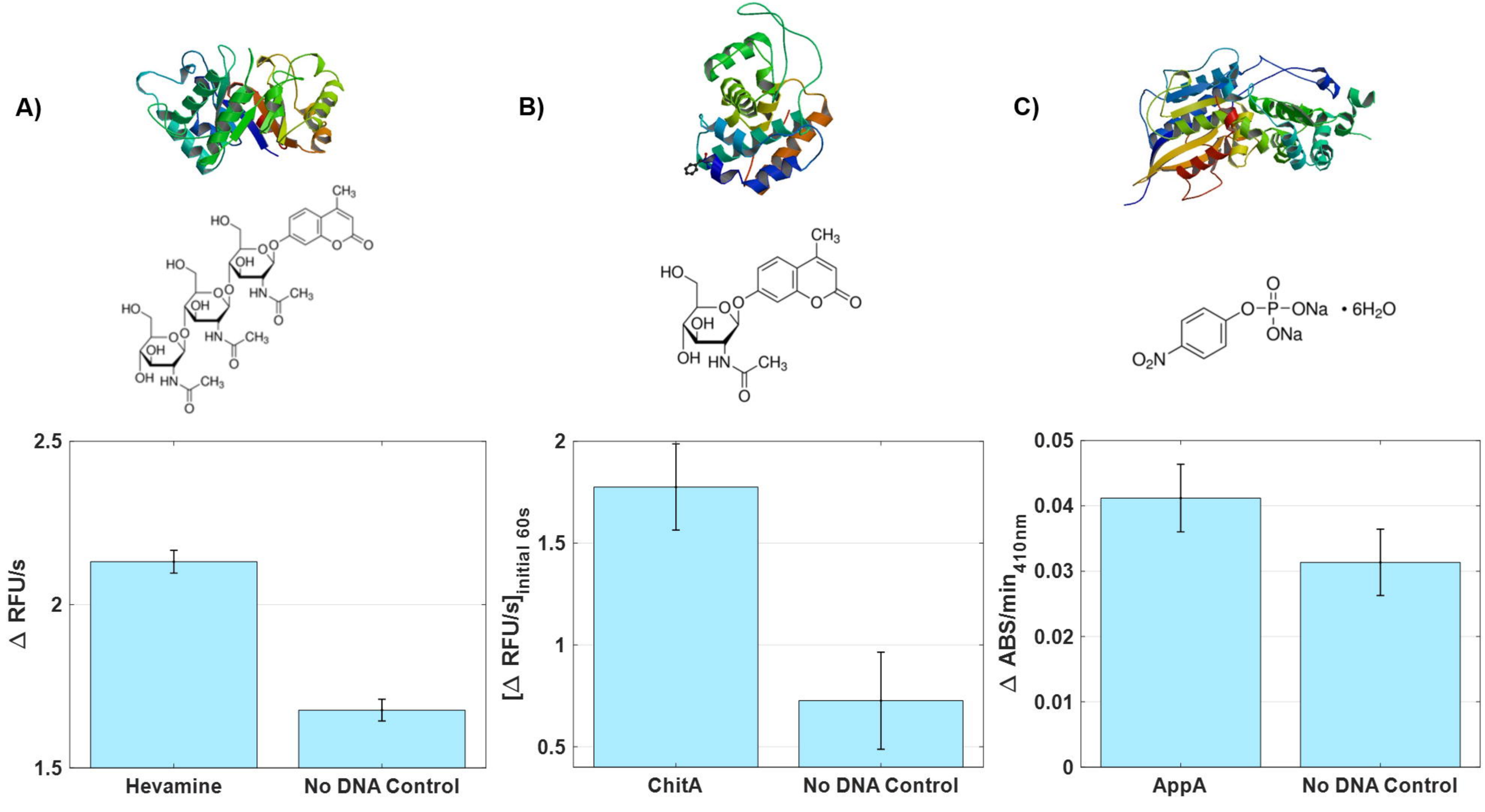

×10^5
8
6
4
2
0
RLU
T7RNAP (-) DsbC (-)
T7RNAP (-) DsbC (+)
T7RNAP (+) DsbC (-)
T7RNAP (+) DsbC (+)
No DNA
T7 SHuffle
KGK10


A)
×10^6
2.5
2
1.5
1
0.5
0
RLU
16 hr 30°C
4 hr 30°C
4 hr 37°C
No DNA


B)
×10^6
2.5
2
1.5
1
0.5
0
RLU
Plasmid
LET
CET
RCA Molar
RCA Mass
No DNA

	sol	native	f	h	i	ip	n	cp	fcp	hcp
GP	5.483	5.338	4.135	3.979	3.587	3.412	2.918	2.862	1.536	1.409
PS	3.182	3.056	1.768	1.776	1.472	1.195	1.65	1.279	1.452	1.213
ML	3.131	2.939	1.828	1.719	1.743	1.067	1.806	1.271	1.472	1.308
MO	2.916	2.707	1.753	1.616	1.572	1.103	1.751	1.195	1.514	1.146
PA	2.68	2.657	1.792	1.89	1.501	1.146	1.713	1.322	1.421	1.186
Control	1.409	1.457	1.156	1.247	1.156	1.114	1.452	1.315	1.125	1.195

5
4.5
4
3.5
3
2.5
2
1.5


A)
B)
C)
Δ RFU/s
2.5
2
1.5
Hevamine
No DNA Control
[Δ RFU/s]initial 60s
2
1.5
1
0.5
ChitA
No DNA Control
Δ ABS/min410nm
0.05
0.04
0.03
0.02
0.01
0
AppA
No DNA Control