

System-level studies of a cell-free transcription-translation platform for metabolic engineering

Yong Y. Wu^{*}, Hirokazu Sato[‡], Hongjun Huang[‡], Stephanie J. Culler[‡], Julia Khandurina[‡], Harish Nagarajan[‡], Tae Hoon Yang[‡], Stephen Van Dien[‡], and Richard M. Murray^{*}

^{*}: California Institute of Technology, 1200 E. California Blvd., Pasadena CA 91125

[‡]: Genomatica Inc., 4757 Nexus Center Dr., San Diego, CA 92121

Keywords: cell-free systems, metabolic engineering, design space exploration, and 1,4-butanediol

Abstract

Current methods for assembling biosynthetic pathways in microorganisms require a process of repeated trial and error and have long design-build-test cycles. We describe the use of a cell-free transcription-translation (TX-TL) system as a biomolecular breadboard for the rapid engineering of the 1,4-butanediol (BDO) pathway. We demonstrate the reliability of TX-TL as a platform for engineering biological systems by undertaking a careful characterization of its transcription and translation capabilities and provide a detailed analysis of its metabolic output. Using TX-TL to survey the design space of the BDO pathway enables rapid tuning of pathway enzyme expression levels for improved product yield. Leveraging TX-TL to screen enzyme variants for improved catalytic activity accelerates design iterations that can be directly applied to *in vivo* strain development.

Introduction

Utilizing fast-growing microorganisms to produce molecules of industrial relevance has the potential to rapidly advance the progress of green chemistry. Processes of traditional chemical synthesis require heavy metal catalysts, toxic solvents, and fossil fuels as feedstocks. The biosynthetic approach, which uses naturally occurring enzymes, less energy, and renewable feedstocks, is becoming an attractive alternative.¹⁻² However, biosynthetic approaches are challenged by long design-build-test cycles. Microbial pathway engineering often has about one-week cycle time.³ The performance of the pathway is frequently far from design, requiring many iterations.⁴ For example, it took DuPont and Genencor more than 100 person-years of work to develop the commercialization of biobased 1,3-propanediol.⁵ Recent advances in cell-free systems offer an alternative to this costly approach. Cell-free systems have been used to reduce the cycle time of pathway construction. The design-build-test cycle in a cell-free system using linear DNA takes less than one day.⁶

Cell-free systems simulate a controlled cellular environment that delivers repeatable results. Recent research has explored the application of cell-free systems for biological circuits, renewable energy, and medicine. The cell-free transcription-translation (TX-TL) system was first developed as a biomolecular breadboard to test genetic circuits, and many have been demonstrated since.⁷⁻¹⁰ The synthesis of hydrogen and the development of enzymatic fuel cells in cell-free systems has charted new paths for renewable energy.¹¹⁻¹² The high yield of therapeutic proteins in *E. coli*-based cell-free synthesis system also offered new methods for medicinal synthesis.¹³⁻¹⁵ Using cell-free systems for prototyping metabolic pathways is an attractive alternative platform for the engineering of biosynthesis in microbial hosts. Lysate of engineered

E. coli has been used to support high-level conversion of valuable chemicals,¹⁶ and a cell-free protein synthesis system has been used to screen enzyme variants.¹⁷ Systematic studies of cell-free systems have monitored the change of pH, measured the change of metabolites over time, and evaluated protein synthesis.¹⁸⁻²⁰ Studies have shown that the depletion of ATP is limiting near the beginning of a cell-free reaction, and the consumption of glutamate is critical in regenerating cofactors.²¹⁻²³ A correlation between cell-free and *in vivo* systems has not been demonstrated. As such, a systematic side-by-side analysis between *in vivo* and cell-free systems is required for cell-free systems to be considered widely as a prototyping platform for metabolic engineering.

Metabolic pathways consist of multiple parts and factors that are combined in precise combinations to achieve desired functions. Enzyme expression levels and activity affect target product yield. Tuning enzyme expression requires engineering the level of transcription, translation, and enzyme activity. Furthermore, balancing expression of multiple genes in parallel remains a challenge. Often, enzymes are overexpressed when they are identified to be essential for improving target metabolite production. Protein overexpression can affect cell growth²⁴ and possibly reduce metabolite production.²⁵ Studies have previously demonstrated the feasibility of tuning protein expression levels *in vivo* for improving metabolite productions.²⁶⁻²⁸ Cell-free TX-TL system also provides a platform for investigating the correlation between protein expression levels and metabolite production. Cell-free TX-TL system allows simultaneous protein expression from multiple pieces of DNA, including linear DNA. Such properties can be used to verify pathway enzyme expression and activity. Sun *et al.*'s work has connected protein expression level in TX-TL to *in vivo* systems by comparing the expression strength of different promoters.⁶ This work takes a further step to compare the dynamics of a metabolic pathway in TX-TL to *in vivo* systems by correlating protein expression levels to metabolite production in both systems.

This work aims to demonstrate the reduction of traditional metabolic engineering design-build-test cycles using the 1,4-butanediol (BDO) pathway as a TX-TL prototype. BDO and its derivatives are widely used for producing automotive plastics, electronic chemicals, and elastic fibers. BDO has a projected global market of \$8.96 billion by 2019.²⁹ Historically, BDO has been produced from petrochemical feedstocks, but recently a bio-based process was commercialized.³⁰ The BDO pathway used for this bioprocess is shown in **Figure 1**. The pathway converts a tricarboxylic acid (TCA) cycle intermediate succinyl-coA to BDO. From top to bottom of the pathway schematic, pathway intermediates include succinyl semialdehyde, 4-hydroxybutyrate (4HB), 4-hydroxybutyryl-coA (4HB-coA), and 4-hydroxybutyraldehyde (4HB-aldehyde). CoA-dependent succinate semialdehyde dehydrogenase (*sucD*) catalyzes the conversion of succinyl-coA to succinyl semialdehyde. 4-hydroxybutyrate dehydrogenase (*hbd*) catalyzes succinyl-coA to 4HB. 4-hydroxybutyryl-coA transferase (*cat2*) catalyzes 4HB to 4HB-coA. 4-hydroxybutyryl-coA reductase (*ald*) catalyzes 4HB-coA to 4HB-aldehyde. Alcohol dehydrogenase (*adh*) catalyzes 4HB-aldehyde to BDO.

The main goal of this work is to demonstrate TX-TL as a research tool for metabolic engineering and to establish the feasibility of design space exploration (shown in **Figure 1**). Using the BDO pathway previously developed by Genomatica Inc.,³¹ we add linear DNA encoding pathway enzymes to TX-TL. We measure the resulting transcriptional and translational outputs and

pathway related metabolites. We also use TX-TL to rapidly tune pathway enzyme expression levels for design space exploration of the BDO pathway. Ribosome-binding site (RBS) elements of varying strengths are chosen from the bicistronic design (BCD) library to adequately explore the design space of the BDO pathway in TX-TL. The two Shine-Dalgarno motifs from BCD allows the first one to make a leader peptide to open the second one, which delivers precise and reliable translation initiation. The translational coupling architecture BCD ensures protein expression at expected levels independent of downstream sequence.³² Through exploring the design space of the BDO pathway in TX-TL and *in vivo*, we systematically compare the metabolic output and enzyme expression levels. To show the industrial relevance of TX-TL, we demonstrate that the use of linear DNA in TX-TL has the capabilities to serve as a biomolecular breadboard to speed up design iterations, and results from TX-TL can be applied to *in vivo* strain development.

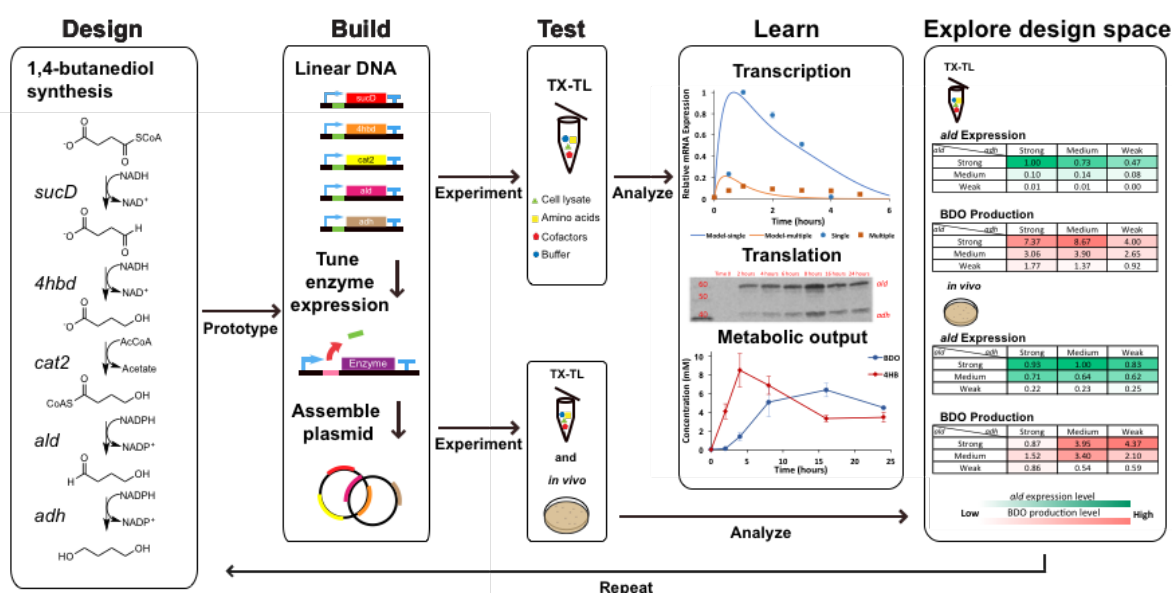


Figure 1: TX-TL as a platform for engineering biosynthetic pathways. A design-build-test-learn cycle using a BDO pathway previously designed by Genomatica Inc. is shown. Linear DNA that encodes pathway enzymes is added to TX-TL. The transcription and translation machinery and metabolic output of TX-TL are analyzed. The design space of the BDO pathway is explored in TX-TL and *in vivo* for comparison. Design: Genomatica Inc. previously designed the BDO synthesis pathway. Build: Generate linear DNA with sequences of pathway enzymes under the same promoters, UTRs, and terminators, subsequently tune enzyme expression levels by varying UTR strength, and assemble into plasmids for design space exploration of the BDO pathway. Test: The TX-TL system is analyzed systematically and compared to the *in vivo* system in parallel. Learn: Systematic analysis of TX-TL through reverse transcription-PCR (RT-PCR), Western blot, and metabolomics. Explore design space: Compare the expression of bottleneck enzyme and BDO production in TX-TL and *in vivo*.

Results and Discussion

Pathway Verification in TX-TL

Pathway enzymes are expressed in TX-TL reactions by adding linear DNA. Linear pieces of DNA are generated by amplifying regions encoding a promoter, a 5'-untranslated region (UTR),

a coding sequence of an individual pathway enzyme, and a terminator. End products of TX-TL reactions are directly used for polyacrylamide gel electrophoresis with sodium dodecyl sulfate (SDS-PAGE) preparation, and all enzymes of the BDO pathway show up on the gel at expected sizes (as shown in **Figure S1**). By adding linear DNA encoding enzymes from the BDO pathway *sucD* (035), *4hbd* (036), *cat2* (034), *ald* (025B), and *adh* (012), we identify that the conversion from 4HB-coA to downstream metabolites is limiting the production of BDO. After 16-hour reactions, more than 10 mM of 4HB is detected, and only 0.5 (\pm 0.1) mM of BDO is detected. 3.1 (\pm 0.1) mM of *gamma*-butyrolactone (GBL) is also detected. GBL is the lactonized form of 4HB, and it is hypothesized to be produced spontaneously from 4HB-coA via *cat2*. The production of 4HB and GBL suggests that *sucD* (035), *4hbd* (036), and *cat2* (034) are not rate-limiting enzymes for the pathway, and *ald* (025B) and *adh* (012) can be rate-limiting. Previous results from Yim *et al.*³³ and Barton *et al.*³⁴ also agree with our observation of such pathway dynamics. We, therefore, hypothesize *ald* and *adh* as the bottleneck enzymes for the production of BDO in TX-TL. To better understand TX-TL as a platform for metabolic engineering, we pick a more advanced *cat2*, *ald* and *adh* combination, *cat2* (C), *ald* (C), and *adh* (C), for studying a wider range of system dynamics. During an initial experiment with the advanced enzymes, 3.4 (\pm 0.9) mM of BDO is detected, and about 3.2 (\pm 0.5) mM of GBL is detected.

The conversion from 4HB to BDO requires the TX-TL system to supplement for electron transfer via cofactors. We subsequently test cofactor concentrations for directing metabolic flux into the production of BDO. We focus on cofactors directly related to the last three steps of the BDO pathway: NADP, NADPH, acetyl-coA, and coA. The absence of acetyl-coA or coA added into the TX-TL system results in less BDO, less GBL, but more 4HB. Also, the addition of acetyl-coA instead of coA helps produce more GBL, which translates to more 4HB-coA synthesized. We hypothesize that the TX-TL system needs more acetyl-coA for the BDO pathway. 1 mM of acetyl-coA is added for the rest of the work. The absence of NADP or NADPH does not drastically affect target metabolite production, but the addition of NADPH improves the production of BDO. 4 mM of NADPH is added for rest of the work, and details can be found in **Figure S2**. These results suggest that the availability of NADPH in TX-TL is limiting the synthesis of downstream products. Such indicates that more strain engineering around the TCA cycle or the pentose phosphate pathway (PPP) may help resolve cofactor imbalance in TX-TL.

System-level studies in TX-TL

Results from system-level studies of TX-TL system are shown in **Figure 2**. For the analysis of transcription and translation, we focus on the bottleneck enzyme *ald* (C). The characterization of the transcription and translation of *ald* (C) are carried out by adding linear DNA encoding individual pathway enzymes. We also perform metabolomics on TX-TL reactions to grasp an understanding of the metabolic network.

Transcription

We predict and observe resource limitations of TX-TL by studying the change in mRNA expression dynamics based on how much DNA is added to the TX-TL reaction. **Figure 2a** shows the mRNA expression of *ald* (C) from TX-TL reaction added with just linear DNA encoding *ald* (C) in blue circles. The orange squares show expression of *ald* (C) from TX-TL reaction added with linear DNA encoding individual pathway enzymes. The blue and orange line are the

respective predicted mRNA dynamics generated by the TX-TL modeling toolbox.³⁵ The linear DNA encoding other pathway enzymes competes with linear DNA encoding *ald* (C) for transcription. The mRNA expression difference of the peak mRNA value is greater than 80%. Note that 30 nM of linear DNA encoding *ald* (C) is added, and a total of 60 nM of linear DNA encoding other pathway enzymes is added. The transcription of *ald* (C) is analyzed using RT-PCR. The background signal from DNA is subtracted. The mRNA level peaks within the first hour and then drops to zero by the end of the first five hours. The rapid degradation of mRNA in TX-TL reflects resource limitations of the system. The dynamics of mRNA in TX-TL is similar to previously reported.³⁶

Translation

The expression of *ald* (C) and *adh* (C) in TX-TL is analyzed using Western blots and can be normalized by total protein intensities measured from SDS-PAGE.³⁷ Complete SDS-PAGE and Western blots for the time-course *ald* (C) expression is shown in **Figure S4**. Protein degradation is observed, which is most likely due to the presence of protease in the extract. Extract used in this work is prepared with S138 (MG 1655 $\Delta adhE \Delta ldhA \Delta pflB + lpdA^*$), an engineered *E. coli* strain previously reported.³³ Most cell-free systems are either developed with purified reagents or with cell lysates from strains with protease deletion. Protein degradation is rarely captured. However, protein degradation is expected here because of the lack of protease gene knockout. The total and soluble protein expression of *ald* (C) and *adh* (C) in TX-TL are shown in **Figure 2b**. Their respective sizes are 60 kDa and 40 kDa. Most of the enzymes remain in the soluble fraction of the TX-TL reactions. The amount of each protein band on the ladder is approximately 200 ng,³⁸ and the concentration of *ald* (C) should be on the order of 10 ng/ul.

Metabolomics

To understand the metabolism in TX-TL, we carry out experiments to collect time-course data of metabolites. The production of pathway intermediate 4HB and target metabolite BDO is shown in **Figure 2c**. The production of 4HB peaks around 4 hours into the reaction. The level of 4HB drops as the compound is converted to downstream metabolites. The concentration of BDO increases as the concentration of 4HB decreases. The concentration of 4HB-coA peaks between 6 and 8 hours, when BDO concentration starts to plateau, and byproducts GBL starts to accumulate (shown in **Figure S3d**). We hypothesize that the consumption of glutamate links to cofactor regeneration. The production of BDO depends on two steps of NADPH-dependent electron transfer. The concentration of NADPH drops to the detection limit by the end of the fourth hour into the TX-TL reaction. Since NADPH is a critical cofactor for the synthesis of BDO, the rate of NADPH being reduced should be the same magnitude as the rate of NADPH being oxidized. The ratio of NADPH/NADP is shown in **Figure S3b**. Carbon flux through the TCA cycle is very weak, but the consumption of glutamate is very prominent. In **Figure 2d**, the conversion of glutamate reaches 80% by the end of TX-TL reactions. A preliminary ¹³C analysis also confirms that glutamate consumption is the primary energy source (data not shown). Furthermore, the conversion from glutamate to α -ketoglutarate reduces NADP to NADPH. The amount of glutamate consumed equals to roughly the sum of 4HB produced, GBL produced, and the BDO produced. The consumption of glutamate is the primary metabolic flux in the TX-TL system.

We learn from the system-level studies that the first few hours are valuable for comparison between pathways or enzyme variants. TX-TL is a resource-limited system: mRNA degradation starts after the first hour, and protein degradation and cofactor imbalance happens. Metabolic flux in TX-TL mainly comes from glutamate, which is also the energy source for regenerating cofactors.

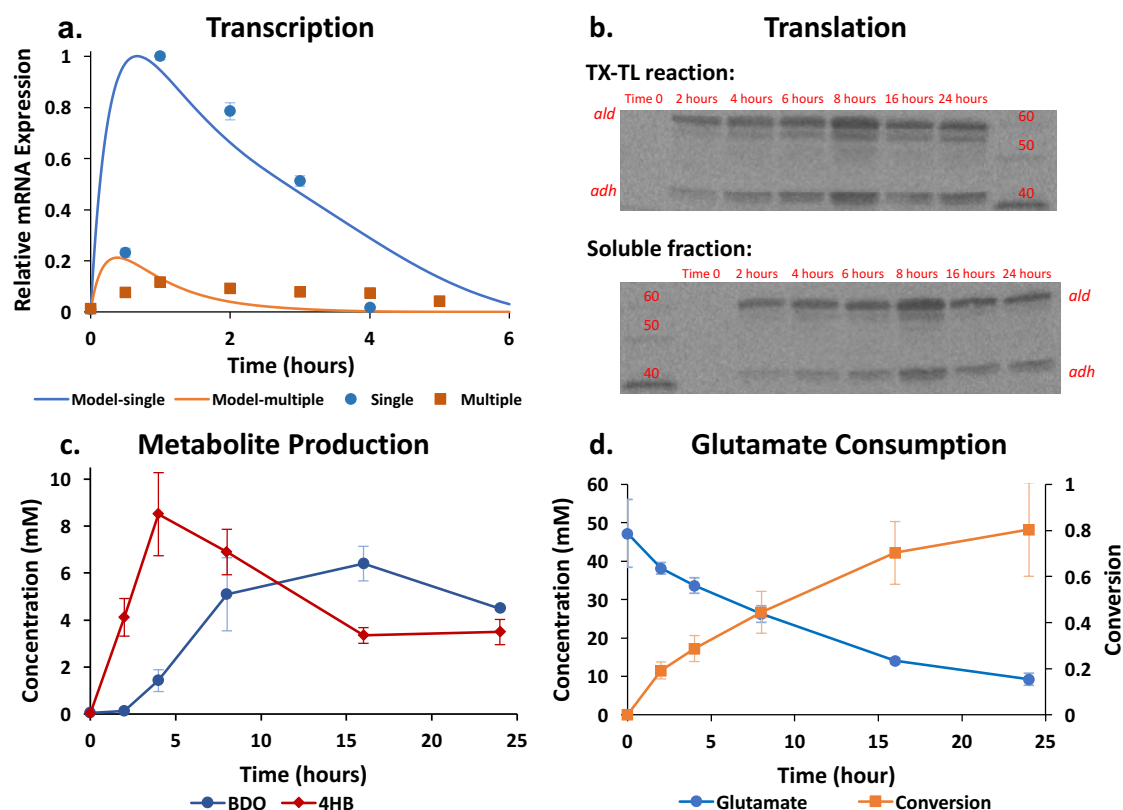


Figure 2: System-level studies of TX-TL. a. Transcription: Relative mRNA for *ald* (*C*) measured using RT-PCR in the first five hours of TX-TL reactions. Blue circles represent data points for mRNA expression of *ald* (*C*) from TX-TL reaction added with just linear DNA encoding *ald* (*C*), and orange squares represent data points for mRNA expression of *ald* (*C*) from TX-TL reaction added with linear DNA encoding all individual pathway enzymes. The blue and orange line represents the respective predictive dynamics. b. Translation: Protein expression analyzed using Western blotting over a 24-hour time range. The top half shows results from analysis of TX-TL reactions, and the bottom half shows results from analysis of supernatant of TX-TL reactions. Both were generated from the same blot. c. Metabolite Production: Measured concentrations of 1,4-BDO (blue circle) and 4HB (red diamond) over time in TX-TL. d. Glutamate Consumption: Glutamate consumption over time in TX-TL (blue circle) and conversion (orange square). Error bars show one standard deviation for $n \geq 3$ independent experiments.

Design space exploration

The expression levels of *ald* (*C*) and *adh* (*C*) are modulated to explore the design space of the BDO pathway. The expression levels of *sucD* (035), *4hbd* (036), and *cat2* (034) are held at fixed

levels. The expression levels of all pathway enzymes are adjusted by tuning the UTR. Constructing the BDO pathway using parts from the BCD library helps modulate pathway enzyme expression levels in TX-TL and *in vivo*. Details of construct design can be found in **Figure S5**. We generate constructs with varying enzyme expression levels, and the level of soluble *ald* (C) expression in TX-TL and *in vivo* are shown in **Figure 3a** and **Figure 3d**, respectively. The SDS-PAGE and Western blot images are shown in **Figure S6**. BCD 2, BCD 20, and BCD 22 are used to modify protein expression level to high, medium, and low. We have constructed plasmids with *ald* (C) and *adh* (C) using a convergent orientation to minimize genetic context effects.³⁹ The nine different constructs (shown in **Table S1**) show a range of protein expression levels in TX-TL and *in vivo*. The *ald* (C) expression levels are similar in TX-TL and *in vivo*. However, the expression of *ald* (C) with constructs containing BCD 22 show higher relative expression *in vivo* versus in TX-TL.

We examine the effect of the expression level of *ald* (C) on metabolite production in both systems. Although the metabolic output in TX-TL is not exactly the same as the one *in vivo*, we observe that the production of BDO is closely related to the expression level of the bottleneck enzyme *ald* (C) in TX-TL and *in vivo*, which matches with previous studies.²⁵ The BDO production level from the constructs in the two systems is different. **Figure 3b** and **3e** shows the production levels of BDO from the designed constructs. The BDO production from constructs with strong *adh* (C) expression are lower *in vivo* than in TX-TL. As shown in **Figure 3b** and **3c**, BDO production level is linearly correlated to *ald* (C) expression from constructs containing the same BCD in front of the *adh* (C) coding region. Overall, the concentration of BDO produced is linearly correlated to the relative expression levels of *ald* (C) in TX-TL regardless of the expression level of *adh* (C). As shown in **Figure 3e** and **3f**, the concentration of BDO produced is linearly correlated to the relative expression levels of *ald* (C) *in vivo* for some constructs. The expression of *adh* (C) is overexpressed for constructs containing BCD 2 in front of the coding region of *adh* (C). Protein overexpression seems to cause a metabolic burden so strong with high *adh* (C) expression that the production of BDO is limited to about 1 mM regardless of the expression level of *ald* (C). Data collected from both systems indicates that *ald* (C) is the bottleneck enzyme of the BDO pathway.

Resource limitation is a more prominent problem *in vivo* versus TX-TL. The *in vivo* system carries a larger metabolic burden with more complicating factors such as cell growth and the development of antibiotic resistance. While developing and maintaining the transcription and translation machinery, the *in vivo* system also fights against the antibiotics in the culture broth. Although we have used promoter pA⁴⁰ as part of the constructs, leaky enzyme expression is still a problem. Transformation of constructs containing BCD2 (for strong expression) is problematic. At 37 °C, the production of BDO *in vivo* is much lower due to metabolic burden (data not shown). At 30 °C, the production of BDO *in vivo* is comparable to the production of BDO in TX-TL. The production level difference can be the supplement of NADPH and acetyl-coA in the TX-TL system.

The mapping of design space of metabolic pathway from TX-TL to *in vivo* can be further tuned by repeating the design-build-test cycle shown in **Figure 1**. There are two ways to map the two systems more closely. The first way is to redesign plasmid constructs. Re-constructing the pathway with a different set of BCDs, perhaps using BCD 9 or BCD12 instead of BCD 2, can

alleviate the metabolic burden *in vivo*. Using weaker BCDs can potentially avoid overexpression of enzymes *in vivo*, which can more effectively show the linear correlation between *ald* (*C*) expression level and BDO production. The second way is to tune protein expression in TX-TL. Plasmid concentration within cells is hard to control, but the expression level of proteins in TX-TL can be tuned by adjusting the concentration of added DNA.⁶ In other words, the concentration of the added DNA can be adjusted to map protein expression levels in TX-TL to *in vivo*. In this work, we have chosen a relatively low DNA concentration to show a linear range of dynamics between *ald* (*C*) expression and BDO production. More design iterations will lead to better mapping results.

This work has only used batch mode reaction to compare the two systems strictly. However, both systems can benefit from pH control by feeding base. Cell growth is limited by pH (shown in **Figure S9**), and the pH can be tuned by adjusting buffer.

a. *ald* expression in TX-TL

<i>ald</i> \ <i>adh</i>	Strong	Medium	Weak
Strong	1.00	0.73	0.47
Medium	0.10	0.14	0.08
Weak	0.01	0.01	0.00

BDO production in TX-TL

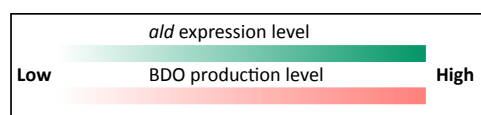
<i>ald</i> \ <i>adh</i>	Strong	Medium	Weak
Strong	7.37	8.67	4.00
Medium	3.06	3.90	2.65
Weak	1.77	1.37	0.92

b. *ald* expression *in vivo*

<i>ald</i> \ <i>adh</i>	Strong	Medium	Weak
Strong	0.93	1.00	0.83
Medium	0.71	0.64	0.62
Weak	0.22	0.23	0.25

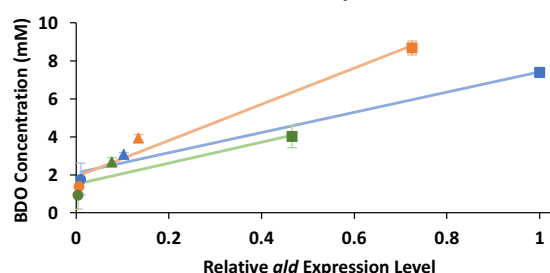
BDO production *in vivo*

<i>ald</i> \ <i>adh</i>	Strong	Medium	Weak
Strong	0.87	3.95	4.37
Medium	1.52	3.40	2.10
Weak	0.86	0.54	0.59



c.

BDO Production Vs. *ald* Expression in TX-TL



d.

BDO Production Vs. *ald* Expression *in vivo*

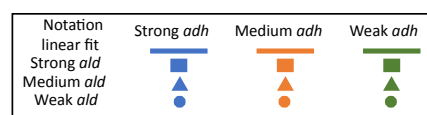
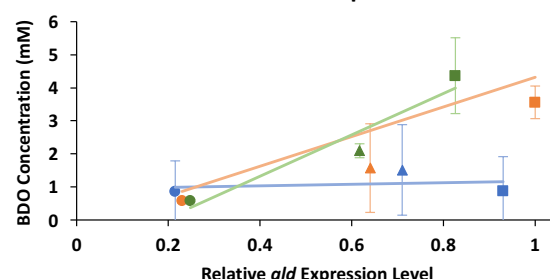


Figure 3: Design Space Exploration in TX-TL and *in vivo*. a. Heat map of *ald* (*C*) expression in TX-TL: Green—*ald* (*C*) expression as the RBS strength of the *ald* (*C*) varies. b. Heat map of the corresponding BDO production in TX-TL: Orange—BDO production (mM) as the RBS strength of the *ald* (*C*) varies. c. BDO production versus protein expression in TX-TL: blue—strong BCD for *adh* (*C*), orange—medium BCD for *adh* (*C*), green—weak BCD for *adh* (*C*), circle: weak BCD for *ald* (*C*), triangle: medium BCD for *ald* (*C*), square: strong BCD for *ald* (*C*). d. Heat map of *ald* (*C*) expression *in vivo*: Green—*ald* (*C*) expression as the RBS strength of the *ald* (*C*) varies. e. Heat map of the corresponding BDO production *in vivo*: Orange—BDO production (mM) as the RBS strength of the *ald* (*C*) varies. f. BDO production versus protein expression *in*

vivo: blue—strong BCD for *adh* (C), orange—medium BCD for *adh* (C), green—weak BCD for *adh* (C), circle: weak BCD for *ald* (C), triangle: medium BCD for *ald* (C), square: strong BCD for *ald* (C). Error bars show one standard deviation for $n \geq 3$ independent experiments.

Application of TX-TL

We apply the TX-TL system to compare three different enzyme combinations of the BDO pathway. The reaction conditions for testing the enzyme combinations are the same. The same *sucD* (035) and *4hbd* (036) are added into TX-TL. The three different enzyme combinations are A: *cat2* (034), *ald* (025B), and *adh* (012) previously published,³³ B: *cat2* (B), *ald* (B), and *adh* (B) previously published,³⁴ C: evolved version of combination B, *cat2* (C), *ald* (C), *adh* (C). As shown in **Figure 4**, we study the production of BDO, 4HB, and GBL in TX-TL. Combination C results with the greatest BDO production rate, and the production of byproduct GBL is the lowest respectively. The production of 4-HB is very similar for all three combinations, but this could be the conversion to 4-HB is reaching equilibrium by the end of the TX-TL reaction. Since the conversion from glutamate to 4-HB is redox neutral, the accumulated 4-HB is almost the same for all three combinations. Enzyme expression comparison between Combination B and Combination C is shown in **Figure 4b**. Notably, the expression of *ald* and *adh* in combination B is much weaker than the expression of *ald* and *adh* in combination C. The engineered enzymes from Combination C is evolved for stability and enzyme specificity for NADPH.

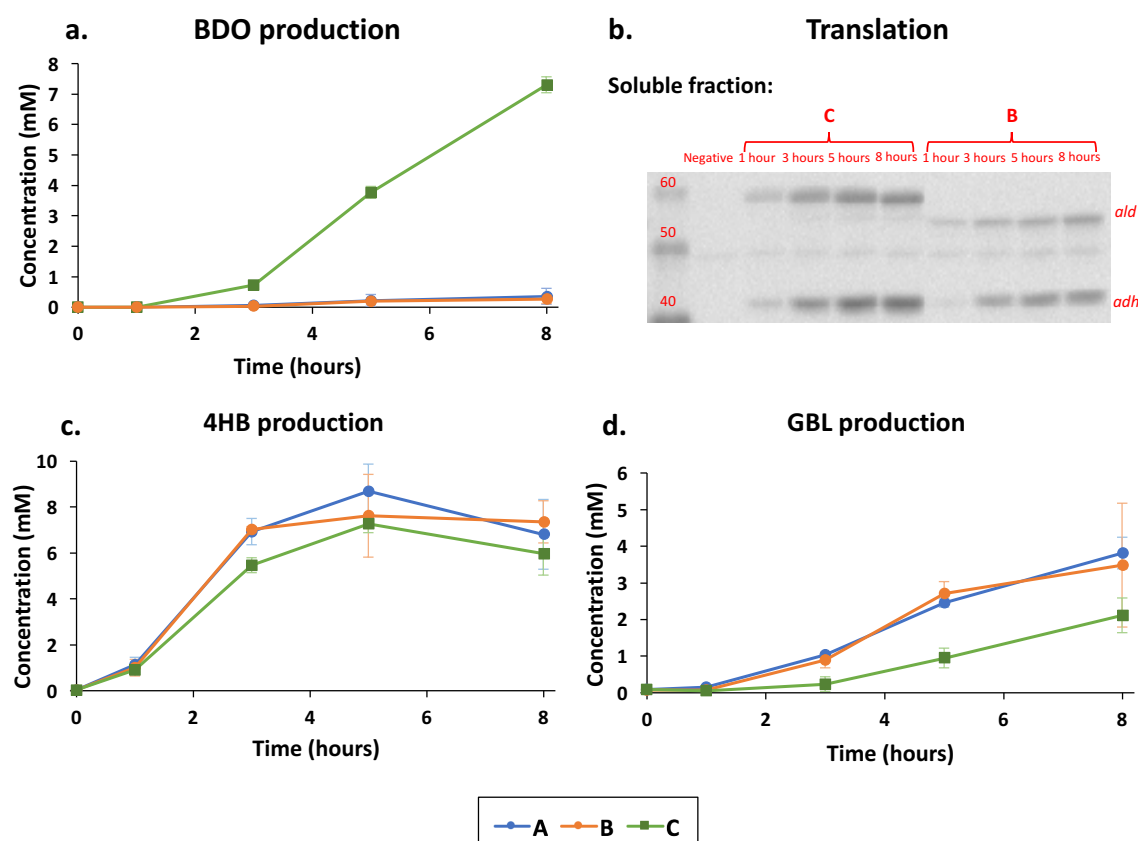


Figure 4: Application of TX-TL—pathway ranking. a. BDO production in the first 8 hours of the TX-TL reaction is shown. b. Soluble protein in the first 8 hours of the TX-TL reaction is shown. c. 4HB production in the first 8 hours of the TX-TL reaction is shown. d. GBL production in the first 8 hours of the TX-TL reaction is shown. Blue circle—enzyme combination A, orange

circle—enzyme combination B, green square—enzyme combination C. Error bars show one standard deviation for $n \geq 3$ independent experiments.

Conclusion

This work shows that TX-TL can be a reliable engineering platform for metabolic engineering. Through studying transcription, translation, and metabolism in TX-TL, we find that the limitation of the TX-TL system lies within rapid degradation of mRNA and cofactor imbalance. The main energy source of TX-TL is glutamate, which is different from the *in vivo* system but this does not impact its ability to produce target metabolites similar to small-scale *in vivo* synthesis. The ease of adding DNA of interests and tuning protein expression by DNA concentrations opens up unique opportunities for rapidly exploring pathway design space. Studying the BDO pathway in both TX-TL and *in vivo* helps confirm that TX-TL is a reliable tool for capturing pathway dynamics. Successfully demonstrating the viability of transcription and translation machinery in S138 extract and the production of BDO provides confidence for future prototyping and engineering in extracts made with advanced *E. coli* strain or potentially other organisms.

Materials and Methods

Cultivation

48-well plates are used for cell culture and metabolite production. M9 minimal salts medium ($6.78 \text{ g l}^{-1} \text{ Na}_2\text{HPO}_4$, $3.0 \text{ g l}^{-1} \text{ KH}_2\text{PO}_4$, $0.5 \text{ g l}^{-1} \text{ NaCl}$, $1.0 \text{ g l}^{-1} \text{ NH}_4\text{Cl}$, 1 mM MgSO_4 , 0.1 mM CaCl_2) is used, and it is supplemented with 10 mM NaHCO_3 , $20 \text{ g l}^{-1} \text{ D-glucose}$ and 100 mM MOPS to improve the buffering capacity, $10 \text{ } \mu\text{g ml}^{-1}$ thiamine and the appropriate antibiotics. 0.5 mM of IPTG is added to induce enzyme expression. Cell culture starts with overnight growth in LB followed by another overnight growth in minimal media and final culture. The final culture is placed on a shaker at 650 rpm at 30°C or 37°C for cultivation. The culture is centrifuged to collect pellet for protein analysis and supernatant for metabolite analysis.

Cell-Free Expression Preparation and Execution

Fermentations are performed with 1 L initial culture volume in 2-l Biostat B+ bioreactors (Sartorius Stedim Biotech). The temperature is held at 37°C , and the pH is held at 7.0 . Cells are grown aerobically to an optical density (O.D.) of approximately 10 , at which point the culture goes through a series of spin-down, re-suspension, and wash cycle according to protocol described by colleagues previously.⁴¹ Collected cell pellets are homogenized using M-110F Microfluidizer Processor and extracted according to the method described by Kwon *et al.*⁴² with the post-homogenization incubation period extended to 80 min instead of 60 min . Buffer preparation is done according to Sun *et al.*'s protocol⁴¹ with a supplement of 30 mM maltodextrin. The preparation results in extract with conditions: $8.9\text{--}9.9 \text{ mg/mL}$ protein, $4.5\text{--}10.5 \text{ mM}$ Mg-glutamate, $20\text{--}40 \text{ mM}$ K-glutamate, $0.33\text{--}3.33 \text{ mM}$ DTT, 1.5 mM each amino acid except leucine, 1.25 mM leucine, 50 mM HEPES, 1.5 mM ATP and GTP, 0.9 mM CTP and UTP, 0.2 mg/mL tRNA, 0.26 mM CoA, 0.33 mM NAD⁺, 0.75 mM cAMP, 0.068 mM folinic acid, 1 mM spermidine, 30 mM 3-PGA, 4 mM NADPH, 1 mM acetyl-coA, and 1 mM NADH. When needed, inducers such as IPTG, linear DNA or plasmid DNA are added to a mix of extract and buffer. TX-TL reactions are conducted in PCR tubes and kept at 29°C with incubation in PCR

machine. BioTeK Synergy H1 microplate reader is used to collect kinetic data for fluorescent protein and MG-Aptamer.

Plasmid DNA, PCR Product Preparation, and Cloning

PCR products are amplified using KOD (Novagen, EMD). Plasmids are miniprep using a QIAprep spin columns (Qiagen) and further concentrated with Millipore's Amicon Ultra-0.5 Centrifugal Filter Unit with Ultracel-30 membrane. All plasmids are processed at stationary phase. Before use in the cell-free reaction, PCR products undergo an additional PCR purification step using QIAquick PCR Purification Kit (Qiagen), which removes excess salt detrimental to TX-TL and are eluted and stored in water at -20°C for long-term storage. Gibson Assembly Ultra and Gibson Assembly HiFi Master Mix from SGI-DNA are used for plasmid assembly.

Analytical Methods

LCMS/MS is used to analyze TX-TL samples and supernatant of cell culture. TX-TL samples were first diluted with 1:3 volume ratio with methanol to remove proteins and other big molecules, and then diluted 1: 12.5 with diluent containing labelled internal standards before LCMS analysis. Cell culture samples were diluted 1:50 with same diluent before LCMS analysis. API3200 triple quadrupole system (AB Sciex, Life Technologies, Carlsbad, CA), interfaced with Agilent 1260 HPLC, utilizing electrospray ionization and MRM based acquisition methods is used. BDO and GBL are detected using positive ionization mode, while 4HB and related acidic compounds are detected using negative ionization mode. Zorbax Eclipse XDB C18 4.6x30mm (particle size 1.8 μm) was used. Column temperature is maintained at 40°C , flow rate of 0.7 mL/min. Injection volume is 5 μL . Eluents include water with 0.1% formic acid (A) and methanol with 0.1% formic acid (B). A fast 1.5 min 5-95% methanol gradient is used, resulting in 1.5 min long LCMS method.

Protein Gel and Western blotting

Culture pellets are resuspended with Novagen BugBuster Protein Extraction solution mixed with rLysozyme and benzonase. The amount of BugBuster added is normalized based on the OD value of 1mL culture. Samples are incubated at room temperature for 20 minutes. An equal volume of 2X Laemmli buffer containing 5% of β -mercaptoethanol is mixed with each sample. The mixture is subsequently boiled in a thermocycler at 99°C for 5 min. 10 μL of a mixture containing cell pellets is loaded into each lane, or 5 μL of a mixture containing TX-TL sample is loaded. Invitrogen MagicMarker XP Western Protein Standard and Bio-Rad Kaleidoscope is added as protein ladders. Bio-Rad 4-15% Criterion TGX (Tris-Glycine eXtended) Stain-Free precast gels are used. Gel imaging is done using Bio-Rad Gel Doc EZ Imager for measuring total protein intensity. Invitrogen iblot system is used for gel transfer. Western blot imaging is done using the Chemi Hi Resolution setting on a BioRad ChemiDoc MP imager. The intensity from Western blot is normalized by the total protein from each lane, and stain-free total protein is the loading control for Western Blots.³⁷ Protein band intensity is determined by using Image Lab 5.2.1.

Competing financial interests

Support for this work is provided by Genomatica Inc., a company pursuing commercialization of the 1,4-butanediol process discussed here. All authors except Y.Y.W. and R.M.M. are employees

of Genomatica Inc. at the time the work is performed. R.M.M is a co-founder and board member of Synvirobio, a for-profit company developing cell-free systems.

Acknowledgement

We thank Dr. Nathan Dalleska and the Environmental Analysis Center for the support and assistance using GC/MS for initial data collection. We thank Anna Lechner, Evan Ehrich, Emily Mitchell, Glenn Majer, and Sari Rizek for material analysis and preparation. We thank Jingyi Li, Jungik Choi, Jonathan Joaquin, Joseph R. Warner, Robin Osterhout, and Robert Haselbeck for helpful discussion. This material is based upon work supported in part by the Defense Advanced Research Projects Agency (DARPA/MTO) Living Foundries program; contract number HR0011-12-C-0065 (DARPA/CMO). Y.Y.W was supported by NIH/NRSA Training Grant 5 T32 GM07616 and the Gordon and Betty Moore Foundation Grant GBMF2809 to the Caltech Programmable Molecular Technology Initiative.

Reference

1. Jang, Y.-S.; Kim, B.; Shin, J. H.; Choi, Y. J.; Choi, S.; Song, C. W.; Lee, J.; Park, H. G.; Lee, S. Y., Bio-based production of C2–C6 platform chemicals. *Biotechnology and Bioengineering* **2012**, *109* (10), 2437-2459.
2. Peralta-Yahya, P. P.; Keasling, J. D., Advanced biofuel production in microbes. *Biotechnology Journal* **2010**, *5* (2), 147-162.
3. Alzari, P. M.; Berglund, H.; Berrow, N. S.; Blagova, E.; Busso, D.; Cambillau, C.; Campanacci, V.; Christodoulou, E.; Eiler, S.; Fogg, M. J.; Folkers, G.; Geerlof, A.; Hart, D.; Haouz, A.; Herman, M. D.; Macieira, S.; Nordlund, P.; Perrakis, A.; Quevillon-Cheruel, S.; Tarandeau, F.; van Tilbeurgh, H.; Unger, T.; Luna-Vargas, M. P. A.; Velarde, M.; Willmanns, M.; Owens, R. J., Implementation of semi-automated cloning and prokaryotic expression screening: the impact of SPINE. *Acta Crystallographica Section D* **2006**, *62* (10), 1103-1113.
4. Alper, H.; Miyaoku, K.; Stephanopoulos, G., Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nat Biotech* **2005**, *23* (5), 612-616.
5. Hong, K.-K.; Nielsen, J., Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cellular and Molecular Life Sciences* **2012**, *69* (16), 2671-2690.
6. Sun, Z. Z.; Yeung, E.; Hayes, C. A.; Noireaux, V.; Murray, R. M., Linear DNA for Rapid Prototyping of Synthetic Biological Circuits in an *Escherichia coli* Based TX-TL Cell-Free System. *ACS Synthetic Biology* **2014**, *3* (6), 387-397.
7. Noireaux, V.; Bar-Ziv, R.; Libchaber, A., Principles of cell-free genetic circuit assembly. *Proceedings of the National Academy of Sciences* **2003**, *100* (22), 12672-12677.
8. Niederholtmeyer, H.; Sun, Z. Z.; Hori, Y.; Yeung, E.; Verpoorte, A.; Murray, R. M.; Maerkl, S. J., Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife* **2015**, *4*, e09771.
9. Pardee, K.; Green, Alexander A.; Ferrante, T.; Cameron, D. E.; DaleyKeyser, A.; Yin, P.; Collins, James J., Paper-Based Synthetic Gene Networks. *Cell* **159** (4), 940-954.
10. 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.; Murray, R. M.; Lucks, J. B., Rapidly Characterizing the Fast Dynamics of RNA Genetic Circuitry with Cell-Free Transcription–Translation (TX-TL) Systems. *ACS Synthetic Biology* **2015**, *4* (5), 503-515.

11. Zhu, Z.; Kin Tam, T.; Sun, F.; You, C.; Percival Zhang, Y. H., A high-energy-density sugar biobattery based on a synthetic enzymatic pathway. *Nat Commun* **2014**, *5*.
12. Wang, Y.; Huang, W.; Sathitsuksanoh, N.; Zhu, Z.; Zhang, Y. H. P., Biohydrogenation from Biomass Sugar Mediated by In Vitro Synthetic Enzymatic Pathways. *Chemistry & Biology* **2011**, *18* (3), 372-380.
13. Zawada, J. F.; Yin, G.; Steiner, A. R.; Yang, J.; Naresh, A.; Roy, S. M.; Gold, D. S.; Heinsohn, H. G.; Murray, C. J., Microscale to manufacturing scale-up of cell-free cytokine production—a new approach for shortening protein production development timelines. *Biotechnology and Bioengineering* **2011**, *108* (7), 1570-1578.
14. Ng, P. P.; Jia, M.; Patel, K. G.; Brody, J. D.; Swartz, J. R.; Levy, S.; Levy, R., A vaccine directed to B cells and produced by cell-free protein synthesis generates potent antilymphoma immunity. *Proceedings of the National Academy of Sciences* **2012**, *109* (36), 14526-14531.
15. Pardee, K.; Slomovic, S.; Nguyen, Peter Q.; Lee, Jeong W.; Donghia, N.; Burrill, D.; Ferrante, T.; McSorley, Fern R.; Furuta, Y.; Vernet, A.; Lewandowski, M.; Boddy, Christopher N.; Joshi, Neel S.; Collins, James J., Portable, On-Demand Biomolecular Manufacturing. *Cell* **2016**, *167* (1), 248-259.e12.
16. Kay, J. E.; Jewett, M. C., Lysate of engineered *Escherichia coli* supports high-level conversion of glucose to 2,3-butanediol. *Metabolic Engineering* **2015**, *32*, 133-142.
17. Karim, A. S.; Jewett, M. C., A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. *Metabolic Engineering* **2016**, *36*, 116-126.
18. Jewett, M. C.; Swartz, J. R., Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnology and Bioengineering* **2004**, *86* (1), 19-26.
19. Shin, J.; Noireaux, V., An *E. coli* Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells. *ACS Synthetic Biology* **2011**, *1* (1), 29-41.
20. Caschera, F.; Noireaux, V., A cost-effective polyphosphate-based metabolism fuels an all *E. coli* cell-free expression system. *Metabolic Engineering* **2015**, *27*, 29-37.
21. Calhoun, K. A.; Swartz, J. R., Energizing cell-free protein synthesis with glucose metabolism. *Biotechnology and Bioengineering* **2005**, *90* (5), 606-613.
22. Jewett, M. C.; Calhoun, K. A.; Voloshin, A.; Wu, J. J.; Swartz, J. R., An integrated cell-free metabolic platform for protein production and synthetic biology. *Molecular Systems Biology* **2008**, *4*, 220-220.
23. Jewett, M. C.; Swartz, J. R., Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. *Biotechnology and Bioengineering* **2004**, *87* (4), 465-471.
24. Glick, B. R., Metabolic load and heterologous gene expression. *Biotechnology Advances* **1995**, *13* (2), 247-261.
25. Culler, S., An integrated bioengineering platform that harnesses the synthetic biology toolbox is the key to an economically viable commercial bioprocess. *Chemical Engineering Progress* **2016**, pp 42-51.
26. Du, J.; Yuan, Y.; Si, T.; Lian, J.; Zhao, H., Customized optimization of metabolic pathways by combinatorial transcriptional engineering. *Nucleic Acids Research* **2012**, *40* (18), e142.
27. Lee, M. E.; Aswani, A.; Han, A. S.; Tomlin, C. J.; Dueber, J. E., Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Research* **2013**.

28. Coussement, P.; Maertens, J.; Beauprez, J.; Van Bellegem, W.; De Mey, M., One step DNA assembly for combinatorial metabolic engineering. *Metabolic Engineering* **2014**, *23*, 70-77.
29. 1,4-Butanediol Market by Technology (Reppe process, Davy process, Butadiene process, Propylene oxide process and others), and by Application (THF, PBT, GBL, PU, and Others) - Global Trends & Forecasts to 2019. <http://www.marketsandmarkets.com/Market-Reports/1-4-butanediol-market-685.html> (accessed February 17, 2017).
30. Novamont opens world's first commercial plant for bio-production of a major intermediate chemical. <https://www.genomatica.com/> (accessed March 20, 2017).
31. Burgard, A.; Burk, M. J.; Osterhout, R.; Van Dien, S.; Yim, H., Development of a commercial scale process for production of 1,4-butanediol from sugar. *Current Opinion in Biotechnology* **2016**, *42*, 118-125.
32. Mutalik, V. K.; Guimaraes, J. C.; Cambray, G.; Lam, C.; Christoffersen, M. J.; Mai, Q.-A.; Tran, A. B.; Paull, M.; Keasling, J. D.; Arkin, A. P.; Endy, D., Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat Meth* **2013**, *10* (4), 354-360.
33. Yim, H.; Haselbeck, R.; Niu, W.; Pujol-Baxley, C.; Burgard, A.; Boldt, J.; Khandurina, J.; Trawick, J. D.; Osterhout, R. E.; Stephen, R.; Estadilla, J.; Teisan, S.; Schreyer, H. B.; Andrae, S.; Yang, T. H.; Lee, S. Y.; Burk, M. J.; Van Dien, S., Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat Chem Biol* **2011**, *7* (7), 445-452.
34. Barton, N. R.; Burgard, A. P.; Burk, M. J.; Crater, J. S.; Osterhout, R. E.; Pharkya, P.; Steer, B. A.; Sun, J.; Trawick, J. D.; Van Dien, S. J.; Yang, T. H.; Yim, H., An integrated biotechnology platform for developing sustainable chemical processes. *J Ind Microbiol Biotechnol* **2015**, *42* (3), 349-360.
35. Tuza, Z. A.; Singhal, V.; Kim, J.; Murray, R. M., An In Silico Modeling Toolbox for Rapid Prototyping of Circuits in a Biomolecular "Breadboard" System. In *52nd IEEE Conference on Decision and Control*, Florence, Italy, 2013.
36. Siegal-Gaskins, D.; Tuza, Z. A.; Kim, J.; Noireaux, V.; Murray, R. M., Gene Circuit Performance Characterization and Resource Usage in a Cell-Free "Breadboard". *ACS Synthetic Biology* **2014**, *3* (6), 416-425.
37. Gilda, J. E.; Gomes, A. V., Stain Free Total Protein Staining is a Superior Loading Control to β -Actin for Western Blots. *Analytical Biochemistry* **2013**, *440* (2), 10.1016/j.ab.2013.05.027.
38. Kao, P., Re: Ladder protein cocentration. Wu, Y. Y., Ed. 2017.
39. Yeung, E.; Dy, A. J.; Martin, K. B.; Ng, A. H.; Del Vecchio, D.; Beck, J. L.; Collins, J. J.; Murray, R. M., The Effect of Compositional Context on Synthetic Gene Networks. *bioRxiv* **2016**.
40. Lutz, R.; Bujard, H., Independent and Tight Regulation of Transcriptional Units in *Escherichia Coli* Via the LacR/O, the TetR/O and AraC/I1-I2 Regulatory Elements. *Nucleic Acids Research* **1997**, *25* (6), 1203-1210.
41. Sun, Z. Z.; Hayes, C. A.; Shin, J.; Caschera, F.; Murray, R. M.; Noireaux, V., Protocols for Implementing an *Escherichia coli* Based TX-TL Cell-Free Expression System for Synthetic Biology. **2013**, (79), e50762.
42. Kwon, Y.-C.; Jewett, M. C., High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Scientific Reports* **2015**, *5*, 8663.