Using Synthetic Biological Parts and Microbioreactors to Explore the Protein Expression Characteristics of *Escherichia coli*

Thomas E. Gorochowski,* Eric van den Berg, Richard Kerkman, Johannes A. Roubos, and Roel A. L. Bovenberg

DSM Biotechnology Center, P.O. Box 1, 2600 MA Delft, The Netherlands

S *Supporting Information*

ABSTRACT: Synthetic biology has developed numerous parts for the precise control of protein expression. However, relatively little is known about the burden these place on a host, or their reliability under varying environmental conditions. To address this, we made use of synthetic transcriptional and translational elements to create a combinatorial library of constructs that modulated expression strength of a green fluorescent protein. Combining this library with a microbioreactor platform, we were able to perform a detailed large-scale assessment of transient expression and growth characteristics of two *Escherichia coli* strains across several temperatures. This revealed significant differences in the robustness of both strains to differing types of protein expression, and a complex response of transcriptional and translational elements to differing temperatures. This study supports the development of reliable synthetic biological systems capable of working across different hosts and environmental contexts. Plasmids developed during this work have been made publicly available to act as a reference set for future research.

KEYWORDS: *protein expression, bioreactor, synthetic biology, characterization, biological parts*
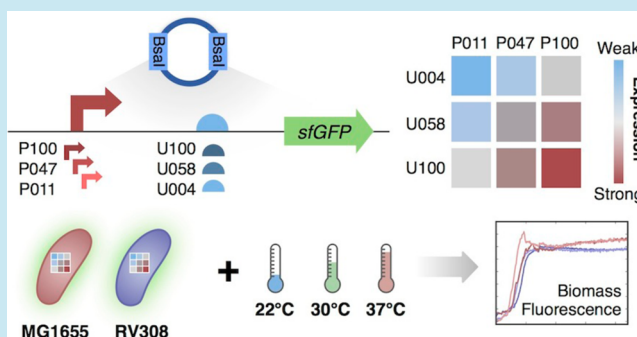
Protein synthesis is a highly dynamic and multistep process. It plays a central role in synthetic biology by providing the machinery needed to execute novel genetic programs. This importance has led to significant effort being made to develop genetic control elements able to precisely modulate various aspects of protein expression.[1−9] Such a capability is not only essential for the successful construction of more complex synthetic biological devices but will also provide the tools needed for the eventual 'tuning' of their function for improved performance and reliability.[10] Indeed, small libraries of such parts have already been successfully used to optimize the flow through a synthetic metabolic pathway and uncover the specific expression levels of individual enzymes that support high product yields.[11]

The multiple levels at which protein synthesis can be regulated has resulted in the development of many different types of parts capable of controlling transcriptional and translational aspects of this process. At the transcriptional level, libraries of promoters have been created spanning a wide range of expression levels.[2,3] In addition to the use of random mutagenesis and screening based techniques, efforts have also been make to understand potential rules governing promoter structure.[4−6] This opens up the opportunity to rationally engineer such elements to integrate multiple signals and allow regulation of expression in user defined ways.

Control of translation has seen similar libraries of ribosome binding sites (RBSs) generated[8] and rational approaches developed.[7] Biophysical models of the interactions between the ribosome and mRNA have successfully been used to predict relative ribosome initiation strengths and applied in a forward-engineering mode to suggest potential RBS sequences with a desired strength.[7] In addition to RBSs, the speed of translation has also been found to be strongly influenced by synonymous codon usage within the gene being expressed. Changes in codon usage have been shown to strongly effect overall expression levels,[12,13] influence the correct folding of active proteins,[14] and to enable dynamic responses to environmental stresses.[15]

A problem with many of the control elements developed so far is that their performance can often be influenced by the genetic context in which they are used.[16,17] Combining the same promoter with differing RBSs and genes can result in very different strengths of expression. Mutalik et al. quantified this effect showing that significant variability arose through interactions between the promoter and RBS, and especially the RBS and the gene of interest.[17] This supported previous findings that illustrated secondary structure formation near the RBS-gene junction strongly affects expression levels.[12] In an attempt to reduce this problem, a library of promoters and

RBSs were designed by the BIOFAB (http://www.biofab.org) to minimize these genetic contextual effects, and characterization of these new parts showed large improvements in reliability.[8]

While the creation of new parts to precisely control protein expression is vital for the development of synthetic biology, an aspect that has been neglected when characterizing part performance is the potential robustness to differing hosts and environmental conditions. Often performance is assessed under limited laboratory conditions and single time points, with little consideration as to the burden being placed upon the cell.[16] In this work we consider "burden" to relate to the additional demands placed on the cells natural protein synthesis capacity by additional expression constructs, that is, amino acids and energy being redrawn from the cell's own growth metabolism. This could then be manifested through changes in growth rate, yield on sugar, expression, or other cellular characteristics. These factors can result in variability in performance between laboratories carrying out identical experiments and lead to difficulties when scaling production in industrial biotechnology applications.[18] This is due to (i) limitations in experimental techniques that result in measurements that are not at sufficient detail to fully capture the variability present and (ii) a lack of attention by the synthetic biology community due to a greater emphasis being placed on the construction of complex designs in a less quantitative manner that work under specific (controlled) conditions. If fully predictive design of more complex of synthetic biological devices, porting of devices between organisms, and broader applications in real-world environments are to become a reality, more detailed and diverse characterization efforts will be essential.[1,18−21]

To tackle this issue, we built a combinatorial library of expression constructs in which transcriptional and translational aspects of expression of a superfolder green fluorescent protein (sfGFP)[22] were modulated in two different strains of *Escherichia coli* (MG1655 and RV308). Making use of this library in combination with a BioLector microbioreactor,[23−25] we were able to perform over 150 independent expression experiments under highly controlled conditions at multiple temperatures, while concurrently taking more than 35 000 independent biomass and fluorescence recordings. This enabled us to capture detailed information related to both part performance and cellular growth characteristics. Unlike many existing studies that use approaches (e.g., shake-flasks) where environmental factors are difficult to precisely control and other factors such as oxygen transfer may influence results, the BioLector platform enabled reliable and highly reproducible data that has also been shown to more accurately mimic larger scales.[24] Analysis of these data revealed differences between strains in terms of their growth dynamics and robustness to both expression strength and temperature. Furthermore, striking changes were found in the relative strengths of transcriptional and translational elements to differing temperatures. Such an experiment would have been infeasible using standard manual or semiautomated approaches for data collection. Therefore, this study also illustrates how recent advancements in microbioreactors offer an improved method for characterization of synthetic biological parts[20] and enables a better understanding of the potential burden they place on the cell.

We began by focusing on the modulation of both transcriptional and translational aspects of protein expression that would form the basis of our characterization. We constructed a set of expression vectors in which the sfGFP protein[22] was constitutively expressed using a range of three different strength constitutive promoters and three different ribosome binding sites (RBSs) (Figure 1; Methods). This
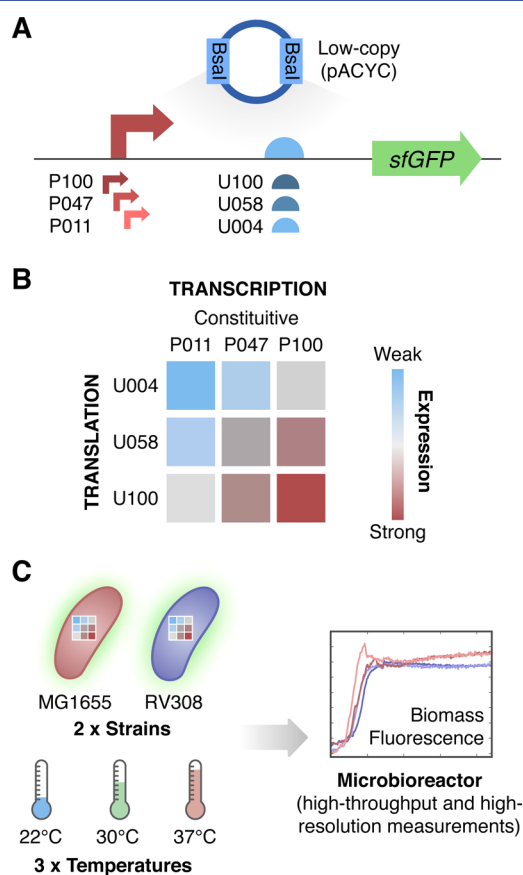

Figure 1. Using synthetic biological parts and microbioreactors to explore expression characteristics. (A) Overview of our expression construct design. (B) Combinatorial construction using Golden Gate cloning allowed us to generate a library of expression vectors covering different strengths of transcription (promoters P100, P047, and P011) and translation (RBSs U100, U058, and U004). (C) Our library was transformed into multiple *E. coli* strains (MG1655 and RV308) and tested under several different temperatures (22, 30, and 37 °C). To enable large numbers of expression experiments to be performed, a BioLector microbioreactor platform was used. This allowed for precise control over environmental parameters and high-resolution temporal measurements of both biomass and fluorescence.

protein was chosen due to its demonstrated robust folding that reduced the chance of strong expression leading to potentially nonactive protein products.[22] This was important as fluorescence was used as a proxy for protein levels, and so, the expression of inactive proteins would lead to an underestimate in overall levels.

To ensure that we spanned the space of potential expression strengths evenly, we made use of well-characterized and publicly available parts from the BIOFAB.[8] We selected elements covering weak, medium, and strong expression for both constitutive promoters and RBSs (Methods). A major advantage of using these parts was that existing characterization data covered a variety of genetic contexts; that is, promoters had been tested with a variety of RBSs and genes. This enabled us to assess not just the strength of expression under a single condition but also the expected variation due to differences in
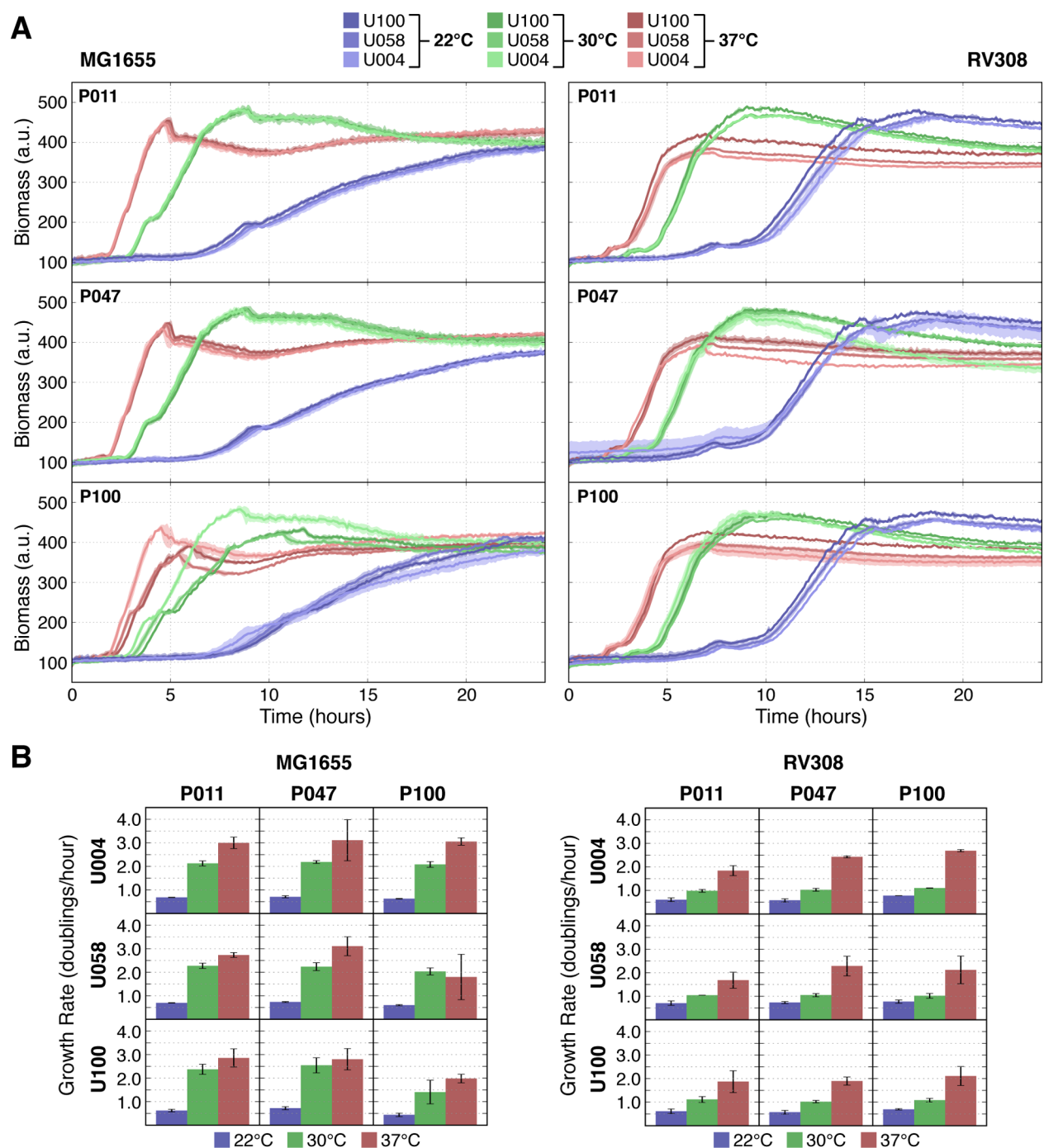
Figure 2. Response of growth dynamics to varying expression levels and temperatures. (A) Time-series of direct biomass measurements for each promoter-RBS combination at 22 °C (blue), 30 °C (green), and 37 °C (red) and for the MG1655 (left) and RV308 (right) strains. Lines represent average values and shaded regions denote the standard deviation. The intensity of the colored lines relates to RBS strength. (B) Maximum growth rates for each promoter-RBS combination at 22 °C (blue), 30 °C (green), and 37 °C (red) with error bars representing the standard deviation (Methods).

the promoter, RBS, and gene combinations. Furthermore, significant contextual effects often present between the RBS and gene of interest[17] were reduced due to the RBSs containing a bicistronic design that helps disrupt potential secondary structure in the mRNA near the ribosome initiation site, ensuring more reliable translational initiation rates.[8] An overview of the basic vector design for the expression library is shown in Figure 1A.

The expression library was transformed into two different *E. coli* K-12 strains: the commonly studied MG1655 and the more industrially relevant RV308 (Methods). These were chosen to quantify the differences that can occur between closely related strains and enabled us to assess how well synthetic parts, often tested in MG1655, might perform in the more industrially realistic background of RV308. The RV308 strain was specifically chosen due to a long history of industrial use starting in 1981 with the synthesis of insulin[26] and continuing
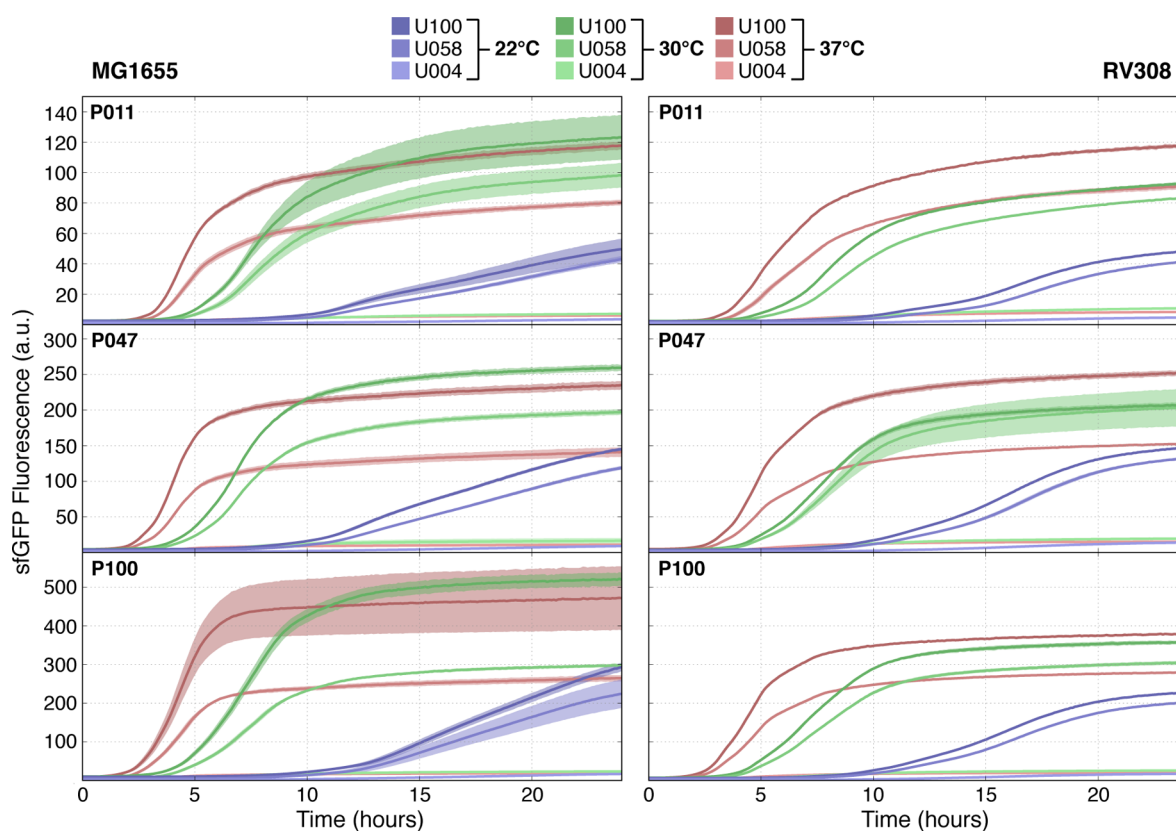
Figure 3. Changes in sfGFP fluorescence over time for various expression strengths and temperatures. Time-series of sfGFP fluorescence measurements for each promoter-RBS combination at 22 °C (blue), 30 °C (green), and 37 °C (red), and for the MG1655 (left) and RV308 (right) strains. Lines represent average values and shaded regions denote the standard deviation. The intensity of the colored lines relates to RBS strength.

with the production of many types of enzyme and therapeutic protein.[27–29] In addition to strain differences, we also evaluated how environmental factors impact upon protein expression characteristics, performing all experiments at multiple temperatures (22, 30, and 37 °C). A standard complex medium (LBC; Methods) was used throughout.

Cells were cultured in a BioLector microbioreactor under relevant conditions for a 24 h period (Methods). To ensure that cells initially contained the required expression plasmid, antibiotic selection was used in the overnight starter cultures. However, upon inoculation of the BioLector no antibiotic was present. This choice was made due to differences in the inherent resistance of each strain to the kanamycin antibiotic (RV308 requiring twice the concentration of MG1655; Methods). As the concentration of kanamycin will affect translational processes and general cellular physiology to differing extents, we did not want this factor influencing our comparisons between strains. Furthermore, the robustness of the strains to plasmid loss was another aspects we wanted to explore, especially due to the fact that antibiotic use is often not viable for larger-scale industrial applications. However, to ensure that plasmid loss did not dominate this process, preliminary experiments were performed with and without kanamycin antibiotic selection for three plasmids of varying expression strength (Methods). In virtually all cases, no significant differences were observed (Supporting Information Figure S1). Some deviation was seen at the highest expression strength for the MG1655 strain (potentially due to a greater selective pressure for plasmid loss due to increased expression

stress), but strong sfGFP fluorescence was still observed (Supporting Information Figure S1).

For the main expression experiments, biomass and sfGFP fluorescence measurements were taken every 6 min to capture a detailed picture of the cellular dynamics. Time-series of these data are shown in Figure 2 for biomass and Figure 3 for sfGFP fluorescence. It should be noted that the biomass measurements we report are based on light-scattering, which has been shown to adhere to a linear relationship with dry cell weight over a wide range of cellular densities.[23] For conversion to more commonly reported OD_{600} values, a calibration experiment was performed with details available in the Methods section.

The growth time-series data revealed clear differences between strains, both in terms of the specific shape of the growth curves and their response to temperature (Figure 2A). Furthermore, the complex medium (Methods) led to a biphasic growth curve with two exponential growth phases separated by a plateau which is likely due to a shift in carbon source. Interestingly, this shift occurs at approximately the same cellular density for differing temperatures and expression strengths, but differs between the strains (OD_{600} ~1.3 for MG1655 and ~0.9 for RV308; Figure 2A).

The MG1655 strains were highly sensitive to temperature, displaying a significant slowing of growth dynamics at 30 °C and even more so at 22 °C. At each temperature, growth dynamics were similar for most levels of protein expression. However, a breakdown did occur for the strongest promoter (P100) and two strongest RBSs (U100 and U058). Expression at these levels at all temperatures leads to a large slowing of
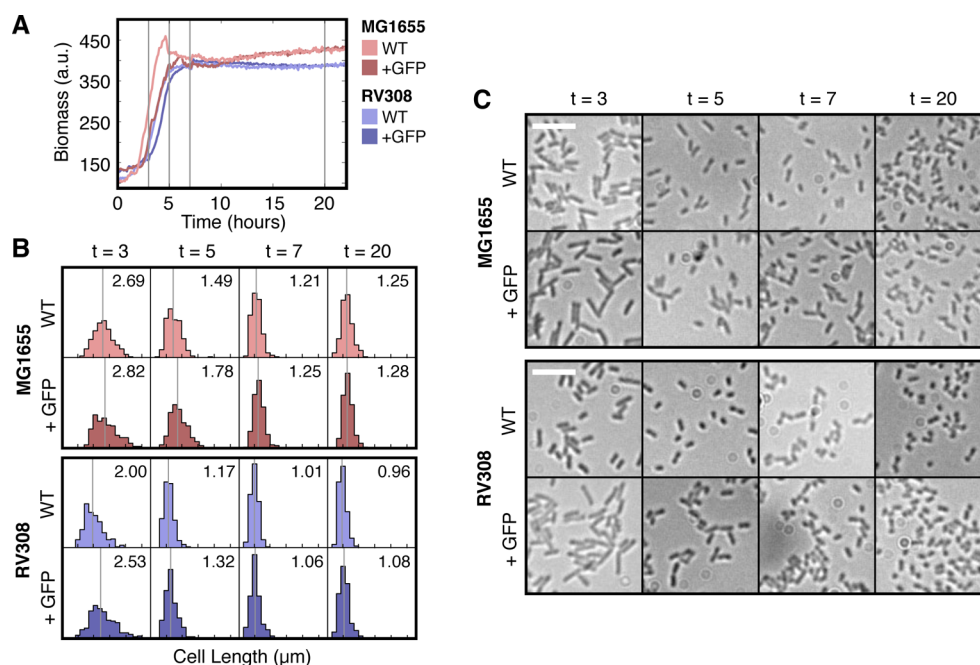
Figure 4. Changes in cell morphology during growth and expression. (A) Biomass time-series with gray vertical lines marking the sample points at 3, 5, 7, and 20 h. (B) Distributions of cell length at each sample point. Gray vertical lines and numbers in the top right corner of each plot show the median values of the distribution. (C) Microscope images of cells from each of the samples (see Supporting Information Figure S3 for original versions). White markers in the top left corner of the images measures 5 μm.

growth ($P = 0.0414$ at 22 °C, 0.0017 at 37 °C, Student's t test; Figure 2B). This is most prominent at 30 and 37 °C, whereas at 22 °C the change is less visible from the time-series plots. For all MG1655 strains, upon reaching stationary phase a sudden decrease in the biomass measurements was observed, before an eventual recovery to a similar steady state level (Figure 2A). The speed of this change was fastest at higher temperatures taking ~15 min at 37 °C for the initial drop. It is unlikely that cell lysis would lead to such a rapid drop in biomass, and so, later on in the text, we investigate how potential shifts in cell morphology may act as a mechanism for this behavior and thereby change the effect of light-scatter based measurement.

In contrast, the RV308 strains displayed far more robust and smooth growth dynamics, maintaining similar growth curves at all temperatures (Figure 2A). A breakdown in these dynamics was also not observed at the strongest expression levels. The only small difference was found between 37 °C and the other lower temperatures with a noticeable reduction in average growth rates (Figure 2B). Interestingly, the steady state biomass measurements for the RV308 strains displayed an approximate ordering based on temperature and the strength of the RBS. Lower temperatures and the use of a stronger RBS for a particular promoter resulted in less of a delay before exponential increase in biomass, and a small increase in the maximum biomass reached. This seems counterintuitive given that expression of foreign proteins will exert additional burden on the cell. However, these additional demands may not actually impact normal growth (as we observe similar growth profiles for cells producing lower levels of protein) and instead may be linked to minor changes in cell shape, as described later in the text.

In terms of sfGFP expression, we find that all strains displayed smooth production profiles over time (Figure 3) with production rates matching the strengths of expression (Supporting Information Figure S2). The only minor exception

was for RV308 at 37 °C, which exhibits initial exponential production followed by a period (~3 h) of linear production before entering stationary phase. With the start of this feature coinciding with the transition to stationary phase growth (Figure 2A), it is likely the result of internal shifts in cellular state as the growth rate is reduced or changes in the complex media during growth.[30]

Variation in sfGFP expression for the majority of experiments and both strains was very small. The few cases where larger differences were observed mostly occurred for the MG1655 strains and stronger promoter strengths, but were found at a range of other temperatures and RBS strengths. Closer inspection of the actual expression values showed that the increased variability was due to a single lower-level reading. Because all experiments were performed with no antibiotic selection, such behavior is consistent with the hypothesis that plasmid loss leads to a subpopulation containing no expression vector. Another potential cause could be the accumulation of point mutations that reduce protein expression.[31] However, as such events are generally rare, plasmid loss is likely to be the cause with the industrial RV308 strain providing greater plasmid stability.

The striking temporal shifts we observed in our biomass measurements for the MG1655 strains are unlikely to be the result of changes in the number of cells through replication or lysis due to the rapid speed with which they occur, some in the order of minutes (see Figure 2A and the time-series for MG1655 at 37 °C where a sharp drop in biomass occurs at ~4.5 h). Instead, we hypothesized that the cells may instead undergo shifts in their morphology induced to differing degrees by the environmental conditions (e.g., availability of nutrients and cell density), genetic differences between MG1655 and RV308, and the internal expression demands we were placing upon the cells. Indeed, recent single cell studies of *E. coli* growth have shown that the dynamics of replication are more

Table 1. Analysis of the Influence of Transcriptional, Translational, and Environmental (Temperature) Factors on Protein Production and Growth Characteristics[a]

	MG1655						RV308					
	$\overline{x}_{22}$	$\overline{x}_{30}$	$\overline{x}_{37}$	CV_{22}	CV_{30}	CV_{37}	$\overline{x}_{22}$	$\overline{x}_{30}$	$\overline{x}_{37}$	CV_{22}	CV_{30}	CV_{37}
sfGFP Production (after 24 h)												
Transcriptional												
P100	100	100	100	7.7	2.1	11.1	100	100	100	1.3	2.5	2.2
P047	54	62	54	8.0	22.0	6.8	70	67	65	11.5	14.5	13.7
P011	20	29	29	13.5	16.5	8.0	23	31	34	13.8	24.5	12.3
Translational												
U100	100	100	100	2.4	7.4	10.7	100	100	100	0.3	1.3	1.0
U058	83	71	62	13.6	15.8	7.5	88	91	70	2.4	10.7	11.1
U004	7	5	5	14.7	21.9	9.3	9	9	6	11.6	20.8	11.5
Temperature												
	66	121	100	22.9	15.2	6.9	65	108	100	26.8	20.3	1.7
Maximum sfGFP Production Rate												
Transcriptional												
P100	100	100	100	8.4	4.1	6.3	100	100	100	2.1	3.5	3.1
P047	47	62	51	10.7	14.5	24.2	70	65	63	8.3	12.7	18.9
P011	15	24	26	16.3	16.0	21.5	21	24	22	3.4	18.0	16.4
Translational												
U100	100	100	100	1.5	7.5	3.4	100	100	100	0.4	1.5	0.2
U058	87	65	57	14.9	14.4	14.2	93	78	58	4.1	14.0	10.6
U004	6	4	3	18.0	4.6	24.4	7	6	4	10.1	12.5	11.0
Temperature												
	20	79	100	33.7	20.0	5.0	34	74	100	24.4	21.0	2.9
Maximum Growth Rate												
Transcriptional (compared to nonexpressing strain)												
P100	79	84	79	18.4	23.0	36.0	106	104	106	7.3	6.3	18.9
P047	105	106	103	5.9	11.3	18.4	89	101	101	15.3	4.7	14.4
P011	97	102	98	7.4	7.5	9.2	91	103	83	13.6	8.2	16.0
Translational (compared to nonexpressing strain)												
U100	84	96	89	24.3	29.3	21.0	89	105	90	12.8	7.5	15.7
U058	100	99	87	9.2	7.7	30.9	105	102	94	8.7	5.5	22.3
U004	98	96	105	6.4	4.5	15.1	93	101	107	16.4	6.3	17.3
Temperature												
	24	81	100	14.7	21.0	18.9	33	51	100	17.8	14.8	12.6

[a]Within each subsection of the table, values are normalized for either the strongest promoter, P100 (transcriptional), strongest RBS, U100 (translational), or highest temperature, 37 °C. The only exceptions are for the transcriptional and translational sub-sections of the maximum growth rate section, which are normalized to a strain that contains no expression plasmid. This enables us to see the impact that expression has on normal growth. Normalized values at the different temperatures ($\overline{x}_{22}$, $\overline{x}_{30}$, and $\overline{x}_{37}$) are given as relative percentages, with the strongest expression element (P100 or U100), highest temperature, or wild-type strain representing 100%. CV_x values correspond to the coefficient of variation at temperature x.

complex than previously thought, with multiple forms present across genetically identical populations.[32]

To investigate this possibility further, we performed a sampling of a subset of strains under normal conditions (37 °C) and at key points during the experiment (Figure 4A). These covered the initial growth after inoculation, the end of exponential phase, early stationary phase, and late stationary phase growth. We also chose to analyze both the MG1655 and RV308 strains, with and without strong protein production (sfGFP expressed using P100 and U100). By viewing these samples under the microscope and manually measuring cell lengths, we could determine the morphological distributions at each stage (Methods).

Figure 4B shows the cell length distributions at various points during the experiment. A clear trend is observed in all cases with significant shifts from longer filamentous-like structures to shorter spherical forms over time (Figure 4C, Supporting Information Figure S3 for original microscope images; $P < 2.2 \times 10^{-16}$, Supporting Information Table S1). We

find that the MG1655 strains exhibit longer cell lengths at all stages and a slower narrowing of the distribution to smaller cell lengths during growth. Strains strongly expressing sfGPF were found to have significantly longer lengths during exponential phase growth (4.8% and 19.5% for MG1655 and 26.5% and 12.8% for RV308 after 3 and 5 h, respectively; $P < 0.03$, Supporting Information Table S1), while converging to similar steady state lengths for each type of strain at the stationary phase (RV308 strains being ~13% shorter than MG1655 strains). Interestingly, although the absolute shifts in length varied between MG1655 and RV308, the relative ratios of the final to initial lengths were approximately the same for both (0.45 and 0.46 for MG1655, and 0.43 and 0.48 for RV308 with and without sfGFP expression, respectively). This suggests that the underlying mechanism for the differences may be strain independent but potentially linked to other factors such as media composition or environmental conditions such as pH or temperature.

Expression of recombinant proteins often involves the optimization of conditions to ensure good yields of active forms and temperature is a common environmental parameter to adjust.[33] Even so, characterization efforts of synthetic parts to control protein expression[1,8] have neglected to consider the influence this variable might have on their performance. Therefore, we attempted to look in more detail at the impact of temperature on the differing strengths of synthetic transcriptional and translational elements used in our expression library.

To assess the effect of temperature on the performance of the transcriptional and translational parts in terms of changes to expression and growth characteristics, we performed an analysis that broke down the relative performance of the strains in relation to transcriptional, translational, and temperature components. For example, given a factor of interest such as the maximum sfGFP production rate, we assessed the relative performance of the transcriptional elements separately (promoters P100, P047, and P011) by normalizing the associated rate for each promoter-RBS combination by the average rate for the strongest promoter (P100) and same RBS (Supporting Information Figure S4). This allowed for relative rates from differing strength RBSs to be combined to give an average relative performance for each promoter in isolation. These could then be calculated at each temperature to capture changes in the relative strengths of the promoters and their sensitivity to temperature. The same methodology was also used to assess translational differences by normalizing to the strongest RBS (U100) and temperature differences by normalizing each promoter-RBS combination value to the average for the same combination at 37 °C (Supporting Information Figure S4). For the transcriptional and translational element analysis of the growth rate data, we instead normalized to growth rate data from the same strain containing no expression plasmid. This allowed us to more easily see the impact that different expression elements had on normal growth. It should be noted that this approach assumes minimal contextual effects between promoter and RBS combinations, as these will influence the overall averages of the pooled values. This is supported by recent experimental results that show such effects to be small.[8]

Table 1 summarizes the results from this analysis (see Supporting Information Figure S4 for a description of how various entries are calculated and Figure S5 for a visualization of the underlying data). As we would expect, a close relationship was found between the total sfGFP production after 24 h and the maximum production rate achieved, with transcriptional and translational elements giving very similar results between these factors. We find that transcriptional elements for both MG1655 and RV308 strains see similar performance across different temperatures, which is in contrast to the translational elements where the relative production total and rate increase at lower temperatures. This is especially prominent for the maximum sfGFP production rate with RBS U058 at 22 °C, where an increase in the relative rate (compared to U100) is seen of 53% and 60% in comparison to 37 °C for the MG1655 and RV308 strains, respectively. For both sfGFP production rate and total, the RV308 strains saw all weaker transcriptional and translational elements perform at stronger strengths than for MG1655.

Unlike expression characteristics, interesting differences were found between the strains when comparing their growth rate analyses. The MG1655 strains displayed a decrease in growth

rates for the strongest promoter (P100) with the differences greatest at high and low temperatures (37 and 22 °C). As both P047 and P011 promoters displayed little impact on growth across different temperatures, this suggests that the P100 promoter may reach a level of mRNA expression that begins to noticeably impact upon cells ability to grow normally (Figure 2B). In contrast, translational elements exhibited less of a trend but also saw a smaller decrease in growth rates for the strongest element (U100).

The RV308 strains displayed very different characteristics with less differences in the growth rate for all transcriptional and translational elements in addition to less variation. This matched the robust behavior observed earlier in the growth time-series data (Figure 2). While we do not have sequence information for the RV308 strain to potentially attribute this robustness to particular genetic differences, the evolution of this strain for improved protein production traits is likely to have resulted in mutations possessing this phenotype; that is, temperature is often varied in production processes,[33] and the need for highly reproducible dynamics is of great importance for industrial-scale processes.

In this work, we combined the use of recently developed synthetic parts for the precise control of protein synthesis, with a microbioreactor platform to explore the protein expression characteristics of two E. coli strains (MG1655 and RV308) across differing temperatures. We have shown that growth and expression of the industrially relevant strain (RV308) are more robust to temperature changes and levels of sfGFP synthesis. In contrast, a significant breakdown in growth dynamics was found for MG1655 when strongly expressing the sfGFP protein.

The detailed time-series data recorded by the microbioreactor revealed transient drops in biomass measurements. These were independent of the breakdown in growth described above during strong expression and were present for all strains with most rapid changes seen for the MG1655 strain. This could be potentially explained after analysis of cell morphology during growth. Significant changes in cell shape over time were found, with a shift to shorter and less variable cell lengths. Furthermore, significant differences were seen in average cell lengths between the different strains (RV308 being shorter than MG1655) and strains strongly expressing sfGFP (wild-type strains were significantly shorter than the sfGFP expressing strains, although to a lesser degree than the between strain differences). Because biomass measurements in the BioLector are performed using a light-scattering technique, it is sensitive to changes in cell size. This means that the reductions in size we observe would lead to lower recorded biomass measurements even though the same number of cells were present.

Finally, by separating the relative influence of transcriptional and translational factors, we found that the transcriptional elements were able to better maintain their relative strength across different temperatures, while translational elements saw a relative increase in the weaker RBSs at lower temperatures. We also showed a strain specific response to these elements. In particular, both types of element were found to have higher relative strengths when placed in the RV308 strain. This strain also displayed more robust growth characteristics with little change between the use of different transcriptional and translational elements. In contrast, the MG1655 strain saw a large increase in growth rate when using weaker transcriptional elements, related to the breakdown observed in the overall growth dynamics.

Table 2. Expression Elements (Promoters and Ribosome Binding Sites)[a]

name	BIOFAB ID	strength (au)	sequence
Promoters			
P100	apFAB95	1594	AAAAAATTTATTTGCTTTCGCATCTTTTTGTACCTATAATGTGTGGA
P047	apFAB45	759	AAAAAGAGTATTGACTTCGCATCTTTTTGTACCTATAATGTGTGGA
P011	apFAB65	174	TTGACATCAGGAAAATTTTTCTGTATAATGTGTGGA
Ribosome Binding Sites			
U100	apFAB682	1594	GGGCCCAAGTTCACTTAAAAAGGAGATCAACAATGAAAGCAATTTTCGTACTGAAACATCTTAATCATGCTAAGGAGGTTTTCTA
U058	apFAB690	948	GGGCCCAAGTTCACTTAAAAAGGAGATCAACAATGAAAGCAATTTTCGTACTGAAACATCTTAATCATGCTGCGGAGGGTTTCTA
U004	apFAB702	120	GGGCCCAAGTTCACTTAAAAAGGAGATCAACAATGAAAGCAATTTTCGTACTGAAACATCTTAATCATGCGATGGACGGTTTCTA

[a]Strength given in arbitrary fluorescence units.[8] The numeric suffix in the name of the element approximately corresponds to the strength as a percentage of the strongest expression element of that type.

Comparing our results to those performed by the BIOFAB,[8] we found that all parts maintain the same rank performance across different temperatures and hosts and saw similar reliability. However, there were some specific differences in the relative performance of the individual parts that our broader characterization highlighted. Specifically, we found that while the part strengths measured by the BIOFAB matched ours for the MG1655 strain at 37 °C, it was clear to see that, in general, the relative strength of the U004 and U058 RBSs grew significantly with lower temperatures in comparison to U100 (see Figure 3 MG1655 strain and Table 1). Furthermore, these changes were also present for the RV308 strain and all temperatures (Figure 3; Table 1).

A major challenge currently facing synthetic biology is the need to improve the reliability of parts to enable the more predictable construction of larger systems. To meet this goal, significant effort has so far has revolved around the development of standards to ensure key characteristics of performance are captured.[19−21] Unfortunately, most standards to date have focused predominantly on device performance under limited conditions,[20] with little consideration of the impact on the host, or changes in performance due to environmental factors. Some studies have looked at the potential role of media and the host,[1,18] but as we show here, relative performance of even basic expression parts is also strongly affected by temperature, something that has not been considered by previous characterization efforts. Furthermore, many of the performance indicators are given as single values, ignoring the dynamic nature of many parts due to intrinsic links with growth characteristics of the host.

A hurdle that has restricted broader characterization efforts covering environmental and temporal aspects of expression and growth has been the practicalities of performing the large numbers of experiments and measurements necessary. While DNA construction and manipulation techniques have advanced to the stage where libraries of thousands of strains can be built, assaying the performance of these across multiple conditions and at a high temporal resolution remains difficult. The BioLector microreactor platform we use here offers a partial solution, enabling larger numbers of strains to be assessed in detail and under controlled environmental conditions. However, miniaturization, potentially using microfluidic approaches, will still be required in the future to meet the potential combinatorial explosion in the number of experiments needed as the range of potential factors required during characterization increases.

There are several potential future directions for this work. First, it is known that media plays a key role in expression characteristics.[1,18] Broadening this study to encompass other widely used complex and defined medias would help clarify if the temperature changes we see are general or media-specific. Second, the size of our library at present is constrained, and we currently to not have the capacity to make predictions about the performance of other parts. This will be essential if engineering of biological systems is to become a reality. By integrating existing characterization data,[9] with models and experiments that capture the demands of synthetic circuits[34] and host interactions,[16] it will be possible to build more complete models that not only cover part performance under known experimental conditions but also enable accurate predictions of part and host response under those not yet tested. Such models will enable a substantial speed-up in the design process of synthetic biological systems and will be essential in making it a true engineering discipline. Finally, the BioLector we have used for data collection has been shown to closely mimic those found in larger reactors, easing potential scale-up.[24] While large-scale production is one potential use for these expression parts, other applications may not be able to provide such controlled environmental conditions. For this reason, it is important to assess the library we have developed using differing apparatus with realistic real-world conditions to gauge the variability that will need to be accommodated by any system considering their use. Such efforts will be crucial if synthetic biological systems are to function reliably and eventually find broad applications outside the laboratory.

To aid in these efforts, the plasmids used in this study have been deposited for public use (Addgene plasmid IDs: 48264−48272) with the aim of them forming a starting point for the future development of a standard reference set to assess the variability of transcriptional and translational processes across different experimental set ups (e.g., microbioreactors, shake flasks, full-scale bioreactors, etc.) and between laboratories (academic and industrial). In addition, they also can be of great use to those testing how other determinants of translational speed, such as codon usage,[12,13,35] are affected by other aspects of protein synthesis. This effort builds on the idea that the foundations of synthetic biology require a coordinated effort to make predictable construction of synthetic biological systems a reality and to achieve the ambitious goals of the field.

■ METHODS

Microorganisms and Media. Cloning was performed using *E. coli* NEB 10-β strains (New England Biolabs, U.S.A.;

part No. C3019H). Expression experiments were carried out using *E. coli* K-12 MG1655 (ATCC# 700926) and RV308 (ATCC# 31608; Su-, lac X 74, gal ISII: OP308, strA) strains. Cells were cultured using Luria−Bertani medium supplemented with casamino acids (LBC) consisting of: 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, and 1 g/L casamino acids, pH 7.0.

Expression Elements. To ensure reliable expression characteristics across the differing constructs, we made use of promoters and RBSs developed and characterized by the BIOFAB.[8] These contain features, such as a bicistronic design for the RBSs, that help reduce the contextual effects often present between expression elements and the gene of interest.[17] Table 2 details the specific designs that we used. All expression elements and the sfGFP gene[22] were synthesized by DNA2.0, U.S.A.

Expression Vectors. The pJ251 plasmid (DNA2.0, U.S.A.) formed the basis for our expression vectors. This contains a low copy origin of replication (pACYC), kanamycin resistance and includes BsaI sites flanking a visual marker for testing successful insertion of an expression cassette. Each of our promoters, RBSs, and the sfGFP gene were synthesized and cloned into a similar vector with ampicillin resistance (pJ254; DNA2.0, U.S.A.). To enable scar-less assembly of the expression vectors, all elements were flanked by appropriate BsaI sites producing 4 bp complementary overhangs after digestion that ensured correct ordering of elements during ligation. The following overhangs were chosen: vector−promoter AGTG; promoter−RBS TGGA; RBS−sfGFP TCTA; sfGFP−vector CCCC. Assembly of all vectors was performed using the standard Golden Gate cloning protocol as described by Engler et al.[36] All expression vectors are publicly available through Addgene (http://www.addgene.org) with plasmid IDs: 48264−48272.

Expression Experiments. Expression experiments were carried out using the BioLector microbioreactor platform (m2p-laboratories GmbH, Germany). To reduce variation in our measurements, the same physical BioLector machine was used for all experiments. Biomass concentrations were measured via scattered light at 620 nm excitation and GFP fluorescence through an excitation filter of 485 nm and an emission filter of 520 nm. Common gains of 20 and 40 were used for the biomass and GFP measurements, respectively.

Starter cultures were grown from single colonies in LBC media that was supplemented with kanamycin (50 μg/mL for MG1655 strains and 100 μg/mL for RV308 strains) at 37 °C overnight. These were then diluted 100-fold in LBC media containing no antibiotic and expression performed in 48-well FlowerPlate microtiter plates (m2p-laboratories GmbH, Germany; part number: MTP-48-B) with 1 mL culture volumes shaken at 900 rpm. Humidity control was enabled on the BioLector and biomass and GFP readings were taken every 6 min. Experiments with the MG1655 strain were performed in triplicate and RV308 strain in duplicate.

Plasmid Stability Experiments. To ensure the stability of the expression plasmids with no kanamycin antibiotic selection, we performed similar expression experiments for both the MG1655 and RV308 strains, three promoters (P100, P047, P011) and the strongest RBS (U100) for media with (50 μg/mL for MG1655 strains and 100 μg/mL for RV308 strains) and without kanamycin antibiotics. This was carried in a single 48-well FlowerPlate microtiter plate with three biological replicates for each strain and performed as described in the previous section.

BioLector Biomass Conversion to OD$_{600}$ Values. To enable comparison to more widely used OD$_{600}$ values of cell density, we performed a calibration experiment in which MG1655 and RV308 strains were grown to saturation overnight, their OD$_{600}$ measured, 5 dilutions made, and then biomass measurements taken in the BioLector at a gain of 20 (used throughout these experiments). Given that biomass measurements from the BioLector follow a linear relationship with cell density,[23] we used the standard linear form $y = ax + b$ and performed a least-squares fit of our biomass measurements to the calculated OD$_{600}$ values at each dilution. For LBC media, this gave $a = 0.01336$ and $b = -1.35251$ for the MG1655 strain and $a = 0.01805$ and $b = -1.61262$ for the RV308 strain.

Calculating Growth and Maximum sfGFP Production Rates. Growth rates were calculated by first removing the background light-scatter of the media (80 au) from the direct biomass measurements. These adjusted values were then $\log_2$ transformed and linear fits calculated using Spotfire 4.5.0. This was performed for each expression experiment separately, and then averages were taken of the biological replicates. Maximum production rates for sfGFP were calculated in a similar way by generating linear fits using Spotfire 4.5.0 over the steepest regions of the expression profiles during the exponential growth phase. Again, these individual rates were then averaged over the biological replicates.

Analysis of Single-Cell Morphology. Cells were prepared for imaging by mixing 50 μL of cell culture with 50 μL of 4% formaldehyde solution. A small droplet ∼1.5 μL was added to a microscope slide and heat fixed. To stain the cells, they were first covered in 10 μL of 5% Crystal Violet solution and then after 1 min washed and dried before microscopy. Samples were viewed using a standard light microscope (Leica DMLA) with a 100× oil immersion objective. Bright-field images were taken by camera (Photometrics CoolSNAP HQ[2]) and distributions of cell length generated using the ImageJ software.[37]

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Figures S1−S5 and Table S1 as described in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: +31 15 279 3135. E-mail: thomas.gorochowski@dsm.com.

Author Contributions

T.E.G., J.A.R., and R.B. conceived the study. T.E.G. performed all experiments and analyses. E.B. and R.K. provided input during construct design and supervised T.E.G.'s experiments. T.E.G. and J.A.R. wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Lee, T.; Krupa, R.; Zhang, F.; Hajimorad, M.; Holtz, W.; Prasad, N.; Lee, S.; Keasling, J. BglBrick vectors and datasheets: A synthetic biology platform for gene expression. *Journal of Biological Engineering* 2013, 5, Epub Sep 20, 2011. DOI: 10.1186/1754−1611−5−12

(2) Mey, M. D.; Maertens, J.; Lequeux, G.; Soetaert, W.; Vandamme, E. Construction and model-based analysis of a promoter library for *E. coli*: an indispensable tool for metabolic engineering. *BMC Biotechnology* 2007, 7, Epub Jun 18, 2007. DOI: 10.1186/1472−6750−7−34

(3) Hartner, F.; Ruth, C.; Langenegger, D.; Johnson, S.; Hyka, P.; Lin-Cereghino, G.; Lin-Cereghino, J.; Kovar, K.; Cregg, J.; Glieder, A. Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. *Nucleic Acids Res.* 2008, 36, Epub Jun 6, 2008. DOI: 10.1093/nar/gkn369

(4) Blount, B.; Weenink, T.; Vasylechko, S.; Ellis, T. Rational Diversification of a Promoter Providing Fine-Tuned Expression and Orthogonal Regulation for Synthetic Biology. *PLoS ONE* 2012, 7, Epub Mar 19, 2012. DOI: 10.1371/journal.pone.0033279

(5) Blazeck, J., and Alper, H. (2013) Promoter engineering: Recent advances in controlling transcription at the most fundamental level. *Biotechnology Journal* 8, 46−58.

(6) Lubliner, S., Keren, L., and Segal, E. (2013) Sequence features of yeast and human core promoters that are predictive of maximal promoter activity. *Nucleic Acids Res.* 41, 5569−5581.

(7) Salis, H., Mirsky, E., and Voigt, C. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* 27, 946−950.

(8) Mutalik, V., Guimaraes, J., Cambray, G., Lam, C., Christoffersen, M., Mai, Q.-A., Tran, A., Paull, M., Keasling, J., Arkin, A., and Endy, D. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature Methods* 10, 354−360.

(9) Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005) Tuning genetic control through promoter engineering. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12678−12683.

(10) Arpino, J., Hancock, E., Anderson, J., Barahona, M., Stan, G.-B., Papachristodoulou, A., and Polizzi, K. (2013) Tuning the Dials of Synthetic Biology. *Microbiology* 159, 1236−1253.

(11) Zelcbuch, L.; Antonovsky, N.; Bar-Even, A.; Levin-Karp, A.; Barenholz, U.; Dayagi, M.; Liebermeister, W.; Flamholz, A.; Noor, E.; S..Amram,; Brandis, A.; Bareia, T.; Yofe, I.; Jubran, H.; Milo, R. Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res.* 2013, 8, Epub Mar 6, 2013. DOI: 10.1093/nar/gkt151

(12) Kudla, G., Murray, A., Tollervey, D., and Plotkin, J. (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 324, 255−258.

(13) Welch, M.; Govindarajan, S.; Ness, J.; Villalobos, A.; Gurney, A.; Minshull, J.; Gustafsson, C. Design parameters to control synthetic gene expression in *Escherichia coli*. *PLoS ONE* 2009, 4, Epub Sep 14, 2009. DOI: 10.1371/journal.pone.0007002

(14) Zhang, G., Hubalewska, M., and Ignatova, Z. (2009) Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nature Structural and Molecular Biology* 16, 274−280.

(15) Wohlgemuth, S., Gorochowski, T., and Roubos, J. (2013) Translational sensitivity of the *Escherichia coli* genome to fluctuating tRNA availability. *Nucleic Acids Res.* 41, 8021−8033.

(16) Cardinale, S., Joachimiak, M., and Arkin, A. (2013) Effects of Genetic Variation on the *E. coli* Host-Circuit Interface. *Cell Reports* 4, 231−237.

(17) Mutalik, V., Guimaraes, J., Cambray, G., Mai, Q.-A., Christoffersen, M., Martin, L., Yu, A., Lam, C., Rodriguez, C., Bennett, G., Keasling, J., Endy, D., and Arkin, A. (2013) Quantitative estimation of activity and quality for collections of functional genetic elements. *Nature Methods* 10, 347−353.

(18) Moser, F., Broers, N., Hartmans, S., Tamsir, A., Kerkman, R., Roubos, J., Bovenberg, R., and Voigt, C. (2012) Genetic Circuit Performance under Conditions Relevant for Industrial Bioreactors. *ACS Synthetic Biology* 1, 555−564.

(19) Arkin, A. (2008) Setting the standard in synthetic biology. *Nat. Biotechnol.* 26, 771−774.

(20) Canton, B., Labno, A., and Endy, D. (2008) Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol.* 26, 787−793.

(21) Kahl, L.; Endy, D. A survey of enabling technologies in synthetic biology. *Journal of Biological Engineering* 2013, 7, Epub May 10, 2013. DOI: 10.1186/1754−1611−7−13

(22) Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T., and Waldo, G. (2005) Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* 24, 79−88.

(23) Kensy, F.; Zang, E.; Faulhammer, C.; Tan, R.-K.; Büchs, J. Validation of a high-throughput fermentation system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates. *Microbial Cell Factories* 2009, 8, Epub Jun 4, 2009. DOI: 10.1186/1475−2859−8−31

(24) Kensy, F.; Engelbrecht, C.; Büchs, J. Scale-up from microtiter plate to laboratory fermenter: evaluation by online monitoring techniques of growth and protein expression in *Escherichia coli* and *Hansenula polymorpha* fermentations. *Microbial Cell Factories* 2009, 8, Epub Dec 22, 2009. DOI: 10.1186/1475−2859−8−68

(25) Huber, R.; Ritter, D.; Hering, T.; Hillmer, A.-K.; Kensy, F.; Müller, C.; Wang, L.; Büchs, J. Robo-Lector − a novel platform for automated high-throughput cultivations in microtiter plates with high information content. *Microbial Cell Factories* 2009, 8, Epub Aug 1, 2009. DOI: 10.1186/1475−2859−8−42

(26) Wetzel, R., Kleid, D., Crea, R., Heyneker, H., Yansura, D., Hirose, T., Kraszewaki, A., Riggs, A., Itakura, K., and Goeddel, D. (1981) Expression in *Escherichia coli* of a chemically synthesized gene for a "mini-C" analog of human proinsulin. *Gene* 16, 63−71.

(27) Birch, G., Black, T., Malcolm, S., Lai, M., Zimmerman, R., and Jaskunas, S. (1995) Purification of recombinant human rhinovirus 14 3C protease expressed in *Escherichia coli*. *Protein Expression and Purification* 6, 609−618.

(28) Horn, U., Strittmatter, W., Krebber, A., Knupfer, U., Kujau, M., Wenderoth, R., Muller, K., Matzku, S., Pluckthum, A., and Riesenberg, D. (1996) High volumetric yields of functional dimeric miniantibodies in *Escherichia coli*, using an optimized expression vector and high-cell-density fermentation under non-limited growth conditions. *Appl. Microbiol. Biotechnol.* 46, 524−532.

(29) Belagaje, R., Reams, S., Ly, S., and Prouty, W. (1997) Increased production of low molecular weight recombinant proteins in *Escherichia coli*. *Protein Sci.* 6, 1953−1962.

(30) Baev, M., Baev, D., Radek, A., and Campbell, J. (2006) Growth of Escherichia coli MG1655 on LB medium: monitoring utilization of sugars, alcohols, and organic acids with transcriptional microarrays. *Appl. Microbiol. Biotechnol.* 71, 310−316.

(31) Kawe, M.; Horn, U.; Plückthun, A. Facile promoter deletion in *Escherichia coli* in response to leaky expression of very robust and benign proteins from common expression vectors. *Microbial Cell Factories* 2009, 8, Epub Jan 26, 2009. DOI: 10.1186/1475−2859−8−8

(32) Reshes, G., Vanounou, S., Fishov, I., and Feingold, M. (2008) Cell Shape Dynamics in *Escherichia coli*. *Biotechnology Journal* 94, 251−264.

(33) Weickert, M., Doherty, D., Best, E., and Olins, P. (1996) Optimization of heterologous protein production in *Escherichia coli*. *Current Opinion in Biotechnology* 7, 494−499.

(34) Leveau, J., and Lindow, S. (2001) Predictive and Interpretive Simulation of Green Fluorescent Protein Expression in Reporter Bacteria. *J. Bacteriol.* 183, 6752−6762.

(35) Bentele, K.; Saffert, P.; Rauscher, R.; Ignatova, Z.; Blüthgen, N. Efficient translation initiation dictates codon usage at gene start. *Molecular Systems Biology* 2013, 9, Epub Jun 18, 2013. DOI: 10.1038/msb.2013.32

138

dx.doi.org/10.1021/sb4001245 | *ACS Synth. Biol.* 2014, 3, 129−139

(36) Engler, C.; Gruetzner, R.; Kandzia, R.; Marillonnet, S. Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS ONE* 2009, *4*, Epub May 14, 2009. DOI: 10.1371/journal.pone.0005553

(37) Schneider, C., Rasband, W., and Eliceiri, K. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods 9*, 671−675.

139

dx.doi.org/10.1021/sb4001245 | *ACS Synth. Biol.* 2014, 3, 129−139