Synthetic Biology, 2022, 7(1), 1–9
DOI: https://doi.org/10.1093/synbio/ysac009
Advance access publication date: 5 July 2022
Research Article

Signal amplification of *araC pBAD* using a standardized translation initiation region

Patrick J. Shilling[1,$,*], Diana Khananisho[1], Alister J. Cumming[1], Bill Söderström[2], and Daniel O. Daley[1,*]

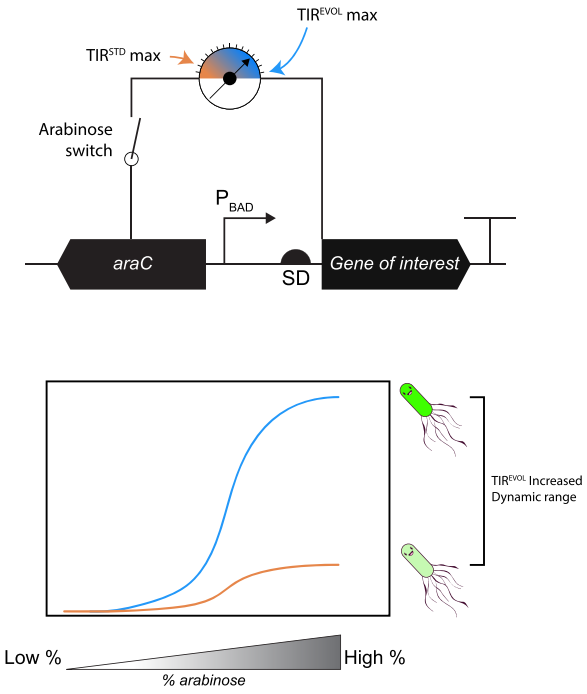
[1]Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden
[2]Australian Institute for Microbiology and Infection, University of Technology Sydney, Sydney, NSW, Australia
[$]Current/Present address: Commonwealth Scientific and Industrial Research Organisation, Clayton, Victoria, Australia.
*Corresponding authors: E-mails: patrick@shilling-lab.com and ddaley@dbb.su.se

Abstract

araC pBAD is a genetic fragment that regulates the expression of the *araBAD* operon in bacteria, which is required for the metabolism of L-arabinose. It is widely used in bioengineering applications because it can drive regulatable and titratable expression of genes and genetic pathways in microbial cell factories. A notable limitation of *araC pBAD* is that it generates a low signal when induced with high concentrations of L-arabinose (the maximum ON state). Herein we have amplified the maximum ON state of *araC pBAD* by coupling it to a synthetically evolved translation initiation region (TIR^{EVOL}). The coupling maintains regulatable and titratable expression from *araC pBAD* and yet increases the maximal ON state by >5-fold. The general principle demonstrated in the study can be applied to amplify the signal from similar genetic modules.

Key words: *araC pBAD*; genetic sensor module; synthetic evolution; translation initiation region; pBAD/HisB

Graphical Abstract



1. Introduction

Genetic sensor modules are cellular components that confer the ability to detect small molecules and respond by regulating gene expression (1). They are widely used in bioengineering for expressing recombinant proteins, manipulating metabolic pathways and assembling synthetic genetic circuits (2–5). Genetic

Submitted: 28 January 2022; **Received (in revised form):** 11 May 2022; **Accepted:** 4 July 2022

sensors typically contain an input device, such as an allosteric transcriptional regulator that binds small molecules, metabolites or high-value chemicals. They also contain an output device, such as a promoter that provides a readout such as transcription of a gene encoding a fluorescent protein, enzyme or antibiotic selection marker. The transcribed messenger RNA (mRNA) is subsequently translated by the cell, allowing easy detection (4). The performance of genetic sensor modules is assessed by titrating a small molecule and monitoring the response. The resulting dose–response curve can be used to calculate important parameters that define performance, such as ON/OFF states, dynamic range, operational range and sensitivity (5). Dynamic range is the difference between maximal (ON) and minimal (OFF) values of the reporter gene output. Operational range is the range of input values (i.e. concentration of small molecule) for which the sensor shows a change in output. And sensitivity is the concentration of small molecule at which the sensor is 50% activated.

Genetic sensor modules derived from nature do not always possess the performance parameters sought for bioengineering applications, such as a broad operational or dynamic range. Strategies to improve or modify biosensors or their downstream response are therefore highly sought after. Proven strategies include modifications to the input device, such as site-specific mutations in the allosteric transcriptional regulator to increase sensitivity, decrease cross-talk and even change the input sensed (6–9). Furthermore, modifications to the output device, such as nucleotide mutations in the promoter, operator or translational initiation region (*TIR*) can amplify the signal from the output module (10–14). Output signal amplification is also possible by increasing the copy number of the sensor (15) or upregulating the expression level of the allosteric transcriptional regulator (16).

araC pBAD from *Escherichia coli* is one of the more widely used genetic sensor modules in bioengineering. It contains the *pBAD* promoter, which regulates the transcription of the genes required for the uptake and conversion of L-arabinose to D-xylulose-5-phosphate (the *araBAD* operon) (17, 18). It also encodes a transcriptional regulator protein AraC, which negatively regulates the *pBAD* promoter in the absence of L-arabinose by binding to distal I_1 and O_2 sites and looping the DNA so that transcription by RNA polymerase is sterically blocked. In the presence of L-arabinose, AraC undergoes a conformational change and moves to the I_1 and I_2 sites on the *pBAD* promoter, allowing the RNA polymerase to initiate transcription from *pBAD*. This event also negatively regulates the *pC* promoter for *araC*. The *araC pBAD* module has been extensively used in bacterial expression plasmids as it allows regulatable and titratable induction of cloned genes in the presence of L-arabinose (16, 19, 20). It has also been built into sensor arrays (16) and biomolecular feedback controllers (21, 22). Moreover, its functionality has been expanded by evolving the AraC sensor, making it more sensitive to L-arabinose and insensitive to IPTG (6) and responsive to D-arabinose (8), mevalonate (9), triacetic acid lactone (23), ectoine (24) and blue light (25). Even though the *araC pBAD* genetic sensor module has been used in a multitude of applications, it has a major limitation: the maximal ON state is relatively low compared to other genetic sensors (16, 20, 26).

In this study, we have increased the maximum ON state of *araC pBAD* by coupling to a synthetically evolved *TIR*. The *TIR* is the nucleotide sequence recognized by the 30S ribosomal subunit during translation initiation (the first ribosomal 'footprint'). It stretches from the Shine–Dalgarno sequence to the fifth codon of the coding sequence (27) and controls ribosomal recruitment during translation initiation. This is the rate-limiting step in protein synthesis (28–31). The synthetically evolved *TIR* identified in this study (TIR^{EVOL}) is partially embedded in an N-terminal purification/detection tag that is part of the downstream output device and can be used as a standardized genetic element for amplifying the maximal ON state of *araC pBAD* in different contexts.

2. Materials and methods

2.1 Molecular cloning

The pBAD/HisB expression plasmid was obtained from Addgene (plasmid #31931). This plasmid contained the coding sequence for PAmCherry1 downstream of the coding sequence for a poly-Histidine sequence, the Xpress™ epitope and the enterokinase protease cleavage site as originally described (32). The coding sequences for super folder green fluorescent protein (sfGFP) (33), mNeonGreen (34) and mEos3.2 (35) were obtained from Genscript Biotech (Netherlands) (Online Supplement, Supplementary Table S1). These coding sequences were cloned into pBAD/HisB using the *in vivo* DNA assembly method (36) and the *E. coli* strain MC1061. The coding sequences were inserted downstream of the coding sequence for the poly-Histidine sequence, the T7 tag, the Xpress™ epitope and the enterokinase protease cleavage site, so that the TIR (either TIR^{STD} or TIR^{EVOL}) was not disturbed.

2.2 Synthetic evolution of the TIR

TIR libraries ($TIR^{LIBRARIES}$) were generated by polymerase chain reaction (PCR) amplification of the pBAD/HisB-sfGFP expression plasmid using overlapping primers as previously described (37–39). For each library, the forward primer incorporated six degenerate nucleotides on either side of the initiating start codon. The reverse primer overlapped with the forward primer by 13 nucleotides, allowing the circularization of the PCR product using the *in vivo* DNA assembly method (36) in the *E. coli* strain MC1061. Primers can be found in the Online Supplement, Supplementary Table S2. The PCR was carried out using the Q5® High-Fidelity DNA Polymerase (New England Biolabs) and a program that consisted of a denaturation step at 95°C for 2 min, followed by 30 cycles of 95°C for 45 s, 44–60°C for 45 s (using a gradient thermocycler) and 72°C for 5 min and a final elongation step of 68°C for 5 min. PCR products that were successfully amplified at the lowest annealing temperature with no nonspecific PCR fragments were used for subsequent steps. A 25 μl aliquot of the PCR was treated with DpnI and then transformed into 200 μl of chemically competent *E. coli* MC1061 using standard protocols that included a 30-min incubation on ice, a 1-min heat shock at 42°C and a 60-min recovery at 37°C. The transformation mixture was plated onto Lysogeny Broth (LB)-agar plates supplemented with 100 μg/ml of ampicillin and 0.1% (w/v) L-arabinose and incubated face down overnight at 37°C.

$TIR^{LIBRARIES}$ were screened by selecting 50 colonies that exhibited high levels of fluorescence on a blue-light box (Syngene). These colonies were further analyzed by growing the bacteria in liquid culture and quantifying the fluorescence from cell pellets. Here, overnight cultures were back-diluted (1:100) into 5 ml of LB containing 100 μg/ml of ampicillin in a 5 ml 24-well growth plate and incubated at 37°C with shaking at 180 RPM until an optical density at 600 nm (OD_{600}) of 0.5 was reached. Cultures were induced by addition of 0.1% (w/v) L-arabinose and then incubated for 2 h at 37°C with shaking at 180 RPM. Following incubation, the OD_{600} was determined, and 1 ml of culture was collected by centrifugation at 3220 × *g* for 15 min. The medium was removed, and the pelleted cells were resuspended in 200 μl buffer (50 mM Tris–HCl pH 8.0, 200 mM NaCl, 15 mM ethylenediaminetetraacetic acid (EDTA)). The cell suspension was transferred to a 96-well optical bottom black-wall plate (Thermo Scientific) and incubated at

room temperature for 2 h, enabling sfGFP maturation. Fluorescence was determined in a SpectraMax M2e (Molecular Devices) at an excitation and emission wavelength of 485 nm and 513 nm, respectively.

Expression plasmids were extracted from selected colonies, and the *TIR* was sequenced. The mutations identified were engineered back into the original expression plasmid using the method of Liu and Naismith (40). Briefly, primers contained complementarity to the template at their 3′ ends and primer–primer complementarity at their 5′ ends. A list of primers can be found in the Online Supplement, Supplementary Table S2. To assist PCR specificity, primer–template complementarity regions possessed a melting temperature 10°C higher than the primer–primer complemented regions.

2.3 Expression conditions

Coding sequences cloned into pBAD/HisB expression plasmids and harbored in the MC1061 strain were expressed using two different protocols. The first protocol was originally described by Guzman and colleagues (20). In short, overnight cultures were back-diluted (1:100) into 5 ml of LB containing 100 μg/ml of ampicillin and L-arabinose concentrations ranging from 0% (w/v) to 0.2% (w/v), in a 5 ml 24-well growth plate. The cultures were incubated at 37°C with shaking at 180 RPM for 3 h until an OD_{600} between 0.3 and 0.5 was reached. The second protocol is described in the instruction manual distributed by Thermo Scientific (41). Here, overnight cultures were back-diluted (1:100) into 5 ml of LB containing 100 μg/ml of ampicillin in a 5 ml 24-well growth plate and incubated at 37°C with shaking at 180 RPM until an OD_{600} between 0.3 and 0.5 was reached. Cultures were then induced by addition of L-arabinose concentrations ranging from 0% (w/v) to 0.2% (w/v). Cultures were incubated for an additional 3 h at 37°C with shaking at 180 RPM.

Activation and repression data, presented in Figure 1, were collected using a modified version of the protocol described by Guzman and colleagues (20). Three colonies of MC1061 harboring pBAD/HisB were grown overnight (i.e. in triplicate), then back-diluted (1:100) into 5 ml of LB and induced with 0.2% (w/v) L-arabinose for 2 h (i.e. activation) and separately, with 0.2% (w/v) D-glucose for 2 h (i.e. repression). Thus, activation and repression data were collected simultaneously from the same colonies.

Coding sequences cloned into the pET expression plasmids and harbored in the BL21(*DE3*) strain were expressed as described previously (37, 39, 42). Briefly, overnight cultures were back-diluted (1:100) into 5 ml of LB containing either 50 μg/ml of kanamycin or 100 μg/ml of ampicillin in a 5 ml 24-well growth plate and incubated at 37°C with shaking at 180 RPM until an OD_{600} between 0.3 and 0.5. IPTG was added at a final concentration of 1 mM and cultures were incubated for either 3 or 20 h at 37°C.

2.4 Quantification of sfGFP fluorescence

One milliliter of bacterial culture was collected by centrifugation at 3220 × *g* for 15 min. The culture medium was removed, and the pelleted cells were resuspended in 200 μl of buffer (50 mM Tris–HCl pH 8.0, 200 mM NaCl and 15 mM EDTA). The cell suspension was transferred to a 96-well optical bottom black-wall plate (Thermo Scientific) incubated at room temperature for 3 h, enabling maturation of the sfGFP. Fluorescence was determined in a SpectraMax M2e (Molecular Devices) at an excitation and emission wavelength of 485 nm and 513 nm, respectively.

2.5 Microscopy

Images were acquired using a NIKON TiE2 N-STORM v5 microscope equipped with an SR HP Apo TIRF 100 × 1.49 NA oil objective and integrated PFS4 (perfect focus system) for drift correction. EPI fluorescence images were acquired using an intensilight C-HGFIE light source with acquisition times of 2–20 ms. The emitted light was guided through a quad band filter (405:488:561:647) before being captured using an OCRA-Flash4 v3 sCMOS camera (Hamamatsu). Note that mEos3.2 was imaged in its green state. To visualize PAmCherry, it was photo-activated using a 405-nm laser line operated at ∼500 mW cm^{-1}, and images were acquired using a 561 nm read-out laser operated at 2 kW cm^{-1}.

Microscopy data were visualized and quantified in Fiji/ImageJ. Plots and statistics were extracted using the webapp PlotsOfData (43).

3. Results

The pBAD/HisB expression plasmid contains the *araC pBAD* genetic module. This module regulates the transcription of an open reading frame encoding an N-terminal extension consisting of a poly-Histidine purification tag, an Xpress™ epitope for detection of recombinant proteins and an enterokinase cleavage site fused in frame to a gene of interest (Figure 1A). Coding sequences expressed from pBAD/HisB, therefore, use a standardized genetic module for transcription (*araC pBAD*) and translation initiation (herein denoted TIR^{STD}) (Figure 1B). To assess the performance of these two genetic modules together, the coding sequence of the sfGFP was cloned into the pBAD/HisB and *E. coli* MC1061 cells containing the plasmid were grown in the presence of either 0.2% (w/v) L-arabinose ('activation') or 0.2% (w/v) D-glucose ('repression'). As anticipated, sfGFP fluorescence was detected from whole-cell pellets when L-arabinose was present in the culture medium (Figure 1C) but not when 0.2% (w/v) D-glucose was present (Figure 1D). The activation:repression ratio was calculated to be 43 (Figure 1E).

We reasoned that TIR^{STD} might not be optimal (since it has presumably been assembled *ad hoc* during the construction of the pBAD/HisB expression plasmid). We, therefore, utilized a synthetic evolution experiment to select a new *TIR* (37, 38). In the experiment, a plasmid library containing >16 × 10^6 theoretical permutations of the *TIR* was constructed and then transformed into MC1061 and plated on LB agar containing 0.1% (w/v) L-arabinose and 100 μg/ml of ampicillin. Fifty of the most fluorescent colonies were selected by visualization on a blue-light box and then a further eight of these were selected by quantification of whole-cell fluorescence in a spectrophotometer. DNA sequencing identified seven novel *TIR* sequences, which differed from TIR^{STD} and from each other (Figure 1B). Moreover, the *TIRs* encoded different amino acids in the second and third positions of the N-terminal extension due to allowance of non-synonymous codon changes in these positions during the synthetic evolution process (Figure 1B). To assess the performance of the synthetically evolved *TIRs*, cells harboring pBAD/HisB-sfGFP were again grown in the presence of either 0.2% (w/v) L-arabinose ('activation') or 0.2% (w/v) D-glucose ('repression'). sfGFP fluorescence from whole-cell pellets confirmed that the average activation was higher when all seven *TIR* sequences were used, regardless of whether it was calculated volumetrically or normalized by the cell density (Figure 1C). Alternatively, repression was unchanged (Figure 1D). The activation:repression ratios indicated that all seven evolved *TIRs* outperformed TIR^{STD} (Figure 1E). TIR^{22} (hereafter referred to

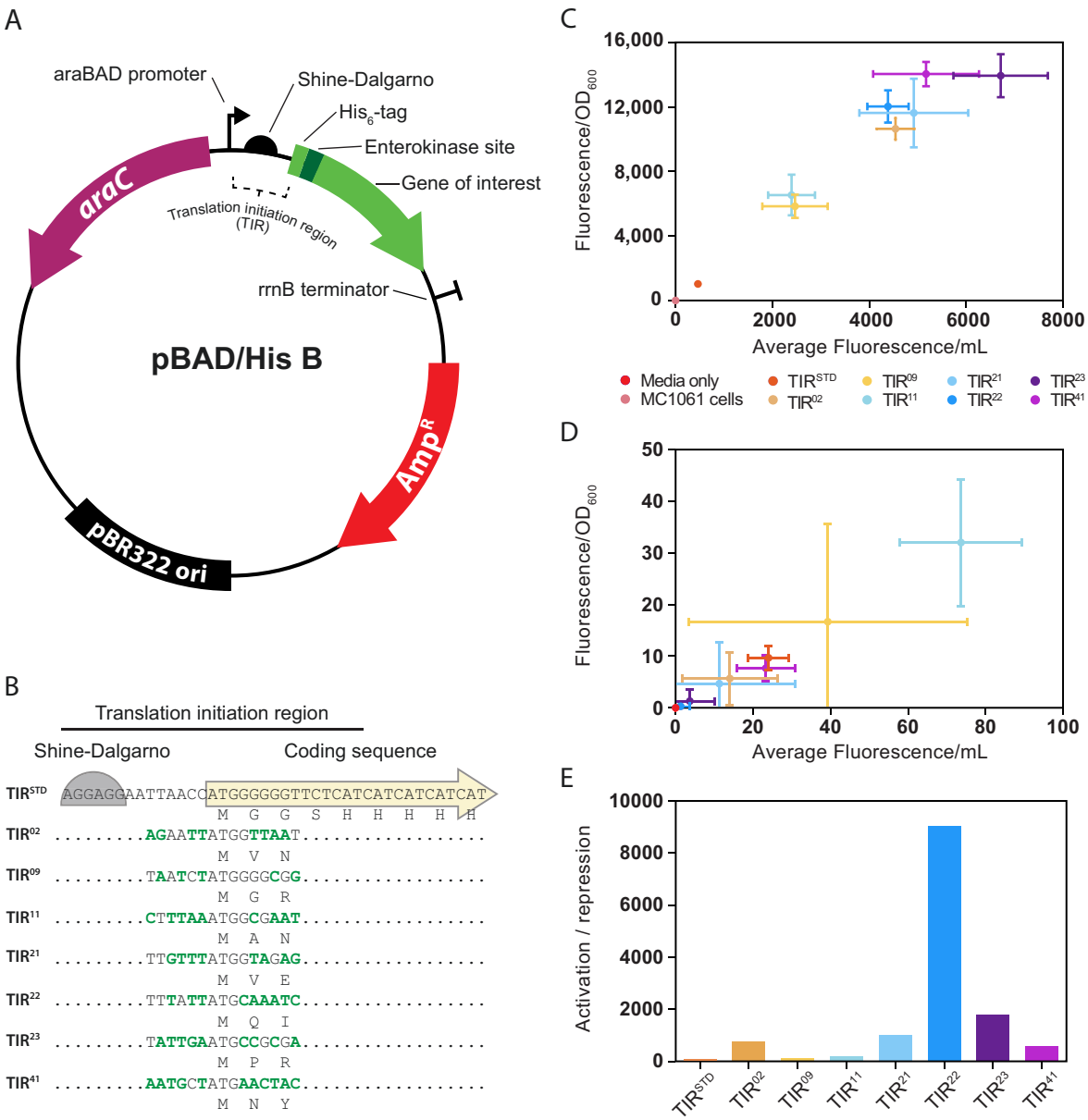


Figure 1. Performance of the *araC pBAD* genetic sensor in the pBAD/HisB expression plasmid. (A) Illustration of the pBAD/HisB expression plasmid that was used in the study. Genes of interest are cloned downstream of a region encoding for a poly-Histidine purification tag, an Xpress™ epitope for detection of recombinant proteins and an enterokinase cleavage site. Resistance to ß-lactam antibiotics is conferred by the *bla* gene encoding the TEM-1 ß-lactamase (denoted Amp^R). Amp^R is harbored on a *Tn3.12* fragment (44). (B) The translation initiation region embedded in pBAD/HisB, denoted TIR^{STD}, stretches from the Shine–Dalgarno sequence to the poly-Histidine purification tag. *TIRs* identified in this study, isolated from individual colonies, are denoted with a number. The resulting nucleotide changes are marked in bold text and corresponding amino acid changes are denoted. Note that TIR^{22} was renamed TIR^{EVOL}. (C) The efficiency of *araC pBAD* with various *TIRs* was calculated by transforming pBAD/HisB into MC1061 and inducing cells with 0.2% (w/v) L-arabinose for 2 h (i.e. activation). Data are presented as mean ± S.D. ($n = 3$). (D) As for (C) except that 0.2% (w/v) D-glucose was used (i.e. 'repression'). (E) The activation:repression ratio (i.e. average activation/average repression) of *araC pBAD* together with various *TIRs* was calculated using values obtained in (C) and (D).

as TIR^{EVOL}) and TIR^{23} were chosen for further characterization as they gave the highest activation:repression ratios (9025 and 1785, respectively).

We compared the titratability of *araC pBAD* with TIR^{STD}, TIR^{EVOL} and TIR^{23} by varying the concentration of L-arabinose and measuring the sfGFP fluorescence. Two induction protocols were used in these experiments. In the first protocol, based on the work of Guzman et al. (4), liquid cultures were started in the presence of varying concentrations of L-arabinose and grown to mid-exponential phase. sfGFP fluorescence was then quantified from whole-cell pellets (Figure 2A). At <0.002% (w/v) L-arabinose, fluorescence was equivalent to the MC1061 strain. At >0.002% (w/v) L-arabinose, titratable expression of sfGFP was observed with all three *TIRs* (Figure 2A). sfGFP fluorescence was 13.5 times higher when TIR^{EVOL} was used and six times when TIR^{23} was used (compared to TIR^{STD}). But since the readings did not plateau at these concentrations of L-arabinose, performance parameters such as the maximal ON state sensitivity, dynamic and operational range could not be calculated. Visualization of cells by fluorescence microscopy confirmed that the increase in fluorescence

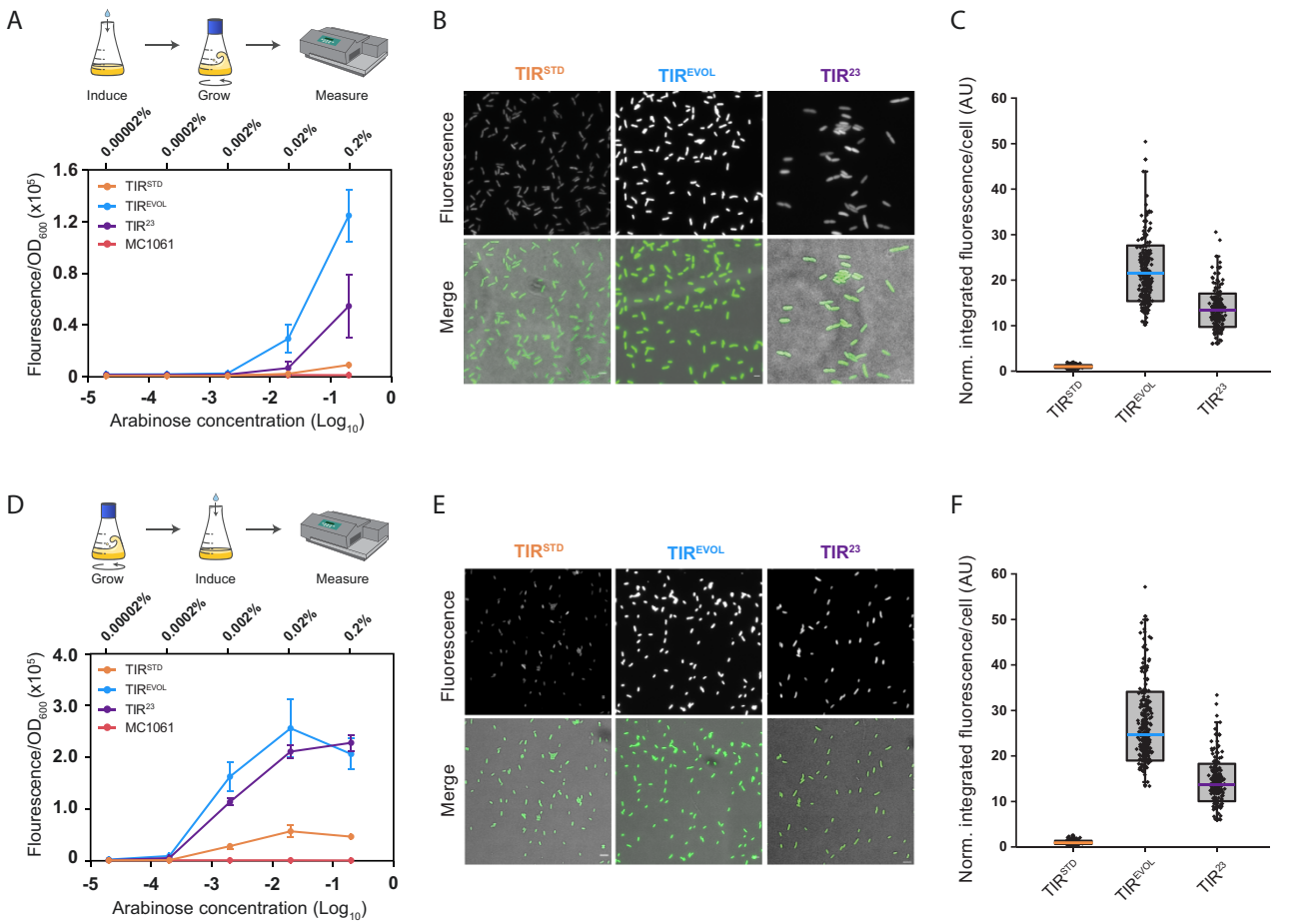


Figure 2. *araC pBAD* is titratable with different TIRs. (A) MC1061 cells harboring pBAD/HisB-sfGFP with either TIR^{STD}, TIR^{EVOL} or TIR^{23} were incubated in the presence of varying concentrations of L-arabinose and grown to mid-exponential phase as described by Guzman et al. (4). sfGFP fluorescence was then quantified from whole-cell pellets. Data are presented as mean ± S.D. ($n = 3$). (B) Visualization of cells by fluorescence microscopy. These cells were induced with 0.2% (w/v) L-arabinose, as described in Panel (A). Scale bars = 4 µm. (C) Box and whisker plot of fluorescence values obtained from individual cells is shown in Panel (B). Values were normalized to TIR^{STD}. $n = 244$ for TIR^{STD}, $n = 257$ for TIR^{EVOL} and $n = 226$ for TIR^{23}. Fluorescence intensity per cell was on average 21.5 times higher when TIR^{EVOL} was used and 13.5 times higher when TIR^{23} was used (i.e. compared to TIR^{STD}). (D) As for Panel (A) except that cells were grown to mid-exponential phase before varying concentrations of L-arabinose were added, as per the manufacturers' instructions (41). Data are presented as mean ± S.D. ($n = 3$). (E) Visualization of cells by fluorescence microscopy. These cells were induced with 0.2% (w/v) L-arabinose, as described in Panel (D). Scale bars = 4 µm. (F) Box and whisker plot of fluorescence values obtained from individual cells shown in Panel (E) ($n = 242$ for TIR^{STD}, $n = 262$ for TIR^{EVOL} and $n = 211$ for TIR^{23}). The fluorescence intensity per cell was on average 26.5 times higher when TIR^{EVOL} was used and 14 times higher when TIR^{23} was used (compared to TIR^{STD}). Note in (C) and (F) that raw fluorescence values of TIR^{EVOL} and TIR^{23} were normalized by the average value of the respective TIR^{STD}. The edges of the box show S.D., colored mid line is mean and whiskers indicate the 1–99% interval. Note also that variability in the data was partly caused by the fact that measurements were made on cells that were at different stages of their cell cycle (which affects cell size and, therefore, total fluorescence).

was in all cells in the population, as expected from previous work (45, 46) (Figure 2B). Quantification of the fluorescence from individual cells confirmed that fluorescence was higher when TIR^{EVOL} and TIR^{23} were used.

In the second protocol, based on the manufacturers' instructions (41), liquid cultures were grown to mid-exponential phase before varying concentrations of L-arabinose were added (Figure 2D). Cells were cultivated for a further 3 h and then sfGFP fluorescence was quantified from whole-cell pellets. Titratable expression of sfGFP was again observed with all three *TIRs* (Figure 2D). The maximal ON state was five times higher for both TIR^{EVOL} and TIR^{23} compared to TIR^{STD}. As a consequence, the dynamic range of *araC pBAD* was increased but the sensitivity and the operational range were unaffected (Figure 2D). Visualization of cells again confirmed that the increase in expression was in all cells in the population (Figure 2E and F). And quantification of the fluorescence from individual cells confirmed that the increase in the maximal ON state was higher when TIR^{EVOL} and TIR^{23} were used.

To determine whether *araC pBAD* and TIR^{EVOL} would work efficiently in the production of other recombinant proteins, the coding sequences for PAmCherry1 (32), mNeonGreen (34) and mEos3.2 (35) were cloned into pBAD/HisB (TIR^{EVOL}) and pBAD/HisB (TIR^{STD}). These fluorescent proteins originate from different organisms and their coding sequences have no homology to each other or to the coding sequence of sfGFP (33). In this set of experiments, we induced cells with 0.2% (w/v) L-arabinose and then visualized them by fluorescence microscopy. The fluorescence intensity from individual cells was then quantified. Cells harboring pBAD/HisB (TIR^{EVOL}) were always more fluorescent than those harboring pBAD/HisB (TIR^{STD}) (Figure 3A-C). This observation was validated by inducing cells with different concentrations of L-arabinose, separating whole-cell lysates by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) and

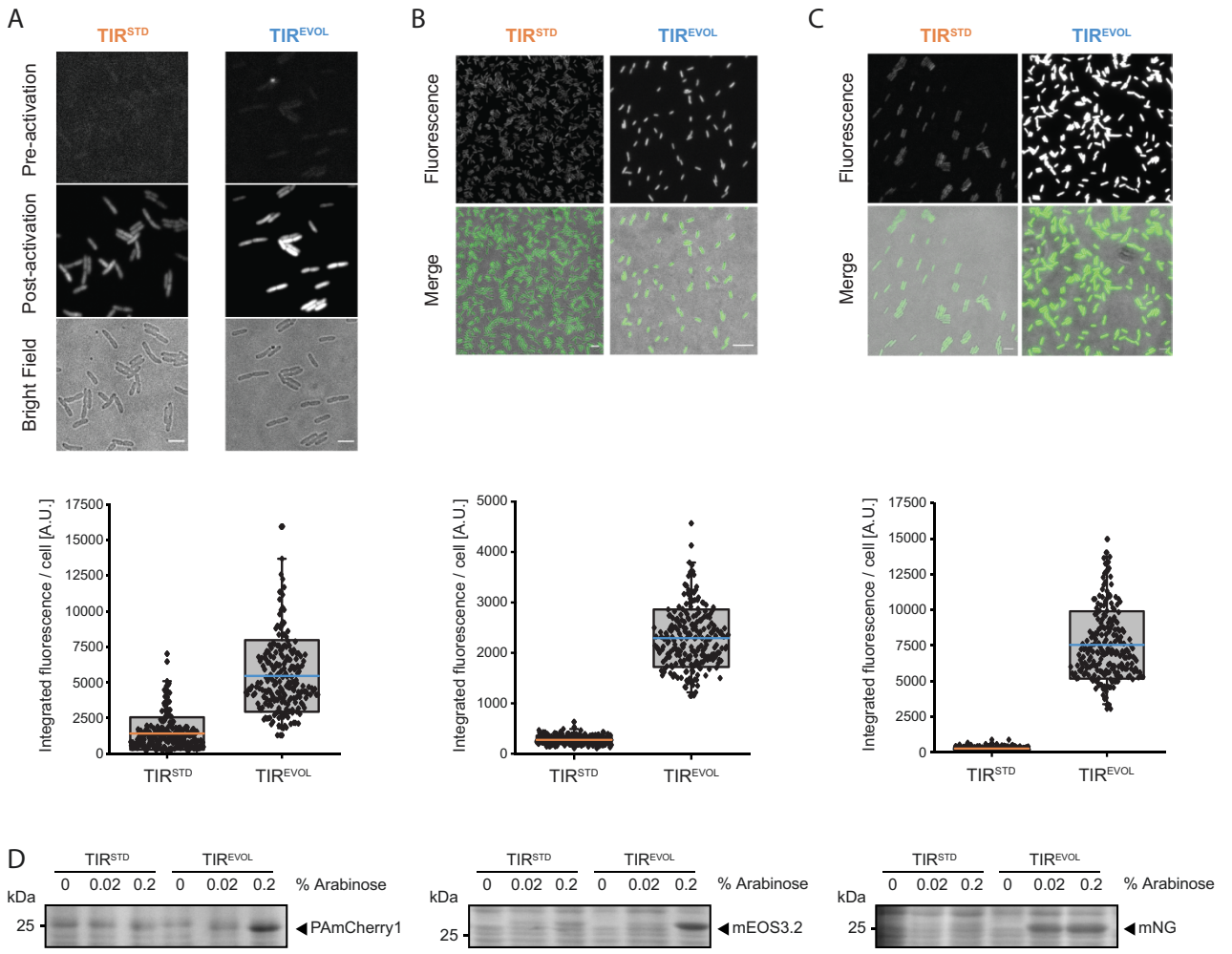


Figure 3. Output from *araC pBAD* and TIR^{EVOL} is increased across multiple coding sequences. Expression of (A) PAmCherry1, (B) mEos3.2 and (C) mNeonGreen from pBAD/HisB-sfGFP (TIR^{STD} and TIR^{EVOL}). All plasmids were harbored in the MC1061 strain and induced with 0.2% (w/v) L-arabinose. Cells were visualized by fluorescence microscopy (top panels) (scale bars = 4 μm). Note that PAmCherry1 has to be activated for visualization and that mEos3.2 was imaged in its green state. Fluorescence intensity values from individual cells were quantified (bottom panels): PAmCherry1 (TIR^{STD}) = 1414 ± 1137 (mean ± S.D.) (n = 244), PAmCherry1 (TIR^{EVOL}) = 5457 ± 2516 (n = 219). mEos3.2 (TIR^{STD}) = 257 ± 81 (n = 217), mEos3.2 (TIR^{EVOL}) = 2291 ± 571 (n = 243). mNeonGreen (TIR^{STD}) = 227 ± 67 (n = 251), mNeonGreen (TIR^{EVOL}) = 7037 ± 1456 (n = 229). The edges of the box show S.D., colored mid line is mean and whiskers indicate the 1–99% interval. Note also that variability in the data was partly caused by the fact that measurements were made on cells that were at different stages of their cell cycle (which affects cell size and, therefore, total fluorescence) (D) Cells were induced with different concentrations of L-arabinose and then separated by SDS-PAGE and stained with Coomassie. Unprocessed gel images are available in the Online Supplement, Supplementary Figures S2 and S3.

visualizing proteins by Coomassie staining (Figure 3D). Similar observations were made when *araC pBAD* and TIR^{23} were combined to produce PAmCherry1, mNeonGreen and mEos3.2 (Online Supplement, Supplementary Figure S1).

Expression plasmids using *araC pBAD* do not support the same production titers of recombinant proteins as other expression plasmids (16, 20, 26) and are, therefore, not as widely used for this purpose (39). Since TIR^{EVOL} had amplified the maximal ON state of *araC pBAD*, we speculated that it would be useful for recombinant protein production. To explore this possibility, we compared production titers with pET28a and pET15b (39). Although these expression plasmids differ from pBAD/HisB in a number of ways (Figure 4A), they are amongst the most commonly used expression plasmids in circulation (39) and, therefore, serve as a useful benchmark. In the experiment, pBAD/HisB (containing *araC pBAD* and either TIR^{STD} or TIR^{EVOL}) was transformed into MC1061, grown to mid-exponential phase and induced with 0.2% (w/v) L-arabinose. MC1061 is deficient for arabinose catabolism owing to the *araD139* mutation, allowing sustained induction of protein expression. pET28a-sfGFP and pET15b-sfGFP were transformed into BL21(*DE3*), grown to mid-exponential phase and induced with 1 mM IPTG. After a standard 3-h induction period, cells were harvested and whole-cell fluorescence was quantified. We observed that pET28a and pET15b supported >10-fold higher production titers of sfGFP than pBAD/HisB (TIR^{STD}) but comparable levels to that observed with pBAD/HisB (TIR^{EVOL}) (Figure 4B). When the experiment was carried out with a 20-h induction period, pBAD/HisB (TIR^{STD}) could still not compete with pET28a and pET15b, whereas pBAD/HisB (TIR^{EVOL}) was again comparable with the pET plasmids (Figure 4C). Taken together, these data indicate that pBAD/HisB (which couples *araC pBAD* and TIR^{EVOL}) can compete with the pET system of expression plasmids, while still providing tight control of expression levels and titratability. However, it should be noted that comparative data were only generated in small-scale cultures, with one recombinant protein, and in one strain background. Further benchmarking needs to be carried out,

A

Expression plasmid	Size (bp)	Selection marker	Replicon (ORI)	Copy number
pET28a	5293	Kan^R (Tn903)	pBR322	Med (15-20)
pET15b	5708	Amp^R (Tn3.1)	pBR322	Med (15-20)
pBAD/HisB	4077	Amp^R (Tn3.12)	pBR322 (Δ*rop*)	High (≥20)

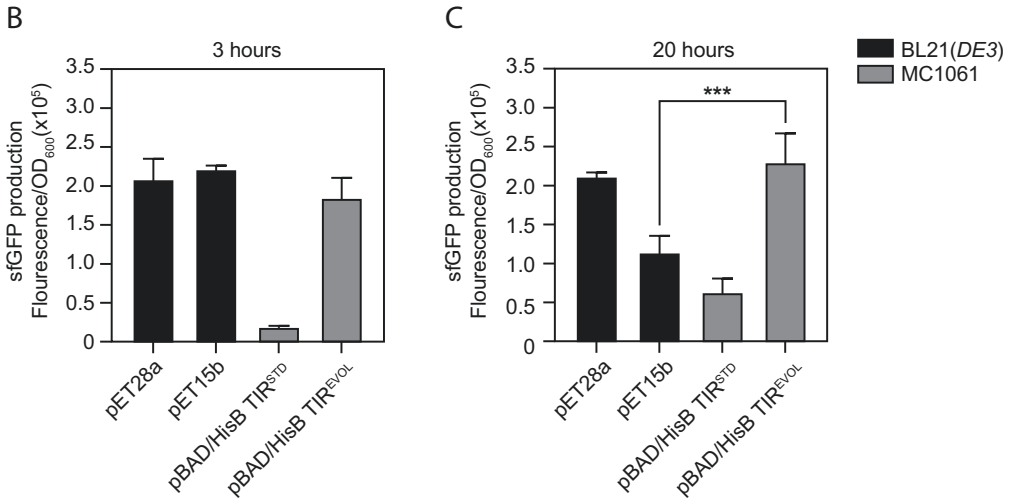


Figure 4. *araC pBAD* and TIR^{EVOL} cooperate effectively in recombinant protein production. (A) Salient features of the pBAD/HisB, pET28a and pET15b expression plasmids, which were used in this experiment. Comparative expression levels of sfGFP after (B) 3 h or (C) 20 h of induction. Fluorescence values were normalized by the OD_{600} of the culture. pET28a and pET15b were harbored in the BL21(*DE3*) strain and induced with 1 mM IPTG. pBAD/HisB-sfGFP (TIR^{STD} and TIR^{EVOL}) were harbored in the MC1061 strain and induced with 0.2% (w/v) L-arabinose. Data are presented as mean ± S.D. ($n = 3$). A statistically significant difference of $P < 0.001$ is denoted by *** (two-tailed Student's *t*-test).

as shown in (26), to determine how pBAD/HisB (TIR^{EVOL}) compares with other expression systems.

4. Discussion

The *araC pBAD* genetic sensor module has a notoriously narrow dynamic range (16, 20, 26), which limits its use in bioengineering applications. In the pBAD/HisB expression plasmid, *araC pBAD* is adjacent to a standardized *TIR* (TIR^{STD}) embedded in an N-terminal purification and detection tag. We used a synthetic evolution process to identify TIR^{EVOL} and demonstrated that it amplified the signal from the *araC pBAD* genetic sensor. The maximum ON signal of *araC pBAD* was increased by 5-fold when coupled with TIR^{EVOL} (i.e. compared to TIR^{STD}). The increased performance observed with *araC pBAD* did not affect titratability or other performance parameters such as background (or leakiness), operational range and sensitivity. Moreover, it is unlikely that they affected solubility of the recombinant protein, as this is an intrinsic property of the protein itself. Taken together, the study indicates that the signal from *araC pBAD* can be amplified by coupling it with TIR^{EVOL}.

At this point in time, bacterial expression plasmids containing *araC pBAD* are most commonly used for titrating the expression levels of recombinant proteins during genetic complementation studies or when balancing flux through metabolic pathways. They are rarely used for recombinant protein production as the maximum ON rate is low compared to expression plasmids containing other induction systems (16, 20, 26). We observed that, when transformed into the MC1061 strain, pBAD/HisB (containing *araC pBAD* and TIR^{EVOL}) could produce comparable levels of sfGFP to the pET/BL21(*DE3*) system. Given the high production yields we observed with pBAD/HisB (containing *araC pBAD* and TIR^{EVOL}), we anticipate that the community might want to utilize it for recombinant protein production (i.e. as an alternative to the pET/BL21(*DE3*) system). We anticipate that similar yields will be possible with other recombinant proteins, based on the fact that *araC pBAD* and TIR^{EVOL} increased production titers of other disparate coding sequences (e.g. PAmCherry1, mEos3.2 and mNG). However, we cannot exclude the possibility that some context-dependent differences in translation initiation may be observed when TIR^{EVOL} is used as a standardized genetic element, as the region encoding the N-terminal extension is only 111 nucleotides long and Woo et al. have reported that 290 nucleotides are required to buffer contextual effects (47). Nevertheless, we believe that pBAD/HisB (containing *araC pBAD* and TIR^{EVOL}) will be an attractive option for those looking to obtain high production yields of recombinant proteins.

Although the current study has focused on TIR^{EVOL}, it is worth noting that another six synthetically evolved *TIRs* were identified in this study (TIR^{02}, TIR^{08}, TIR^{11}, TIR^{21}, TIR^{23} and TIR^{41}). Our preliminary characterization indicated that these *TIRs* also increased the maximum ON rate and activation: repression ratio of *araC pBAD* (compared to TIR^{STD}). And as they exhibited different maximum ON rates, they can also be used to control titratability, for example, if multiple coding sequences need to be expressed in the same cell at different stoichiometries (i.e. by one concentration of L-arabinose). Thus, we can envision the various *TIRs* potentially being used to balance protein stoichiometries in synthetic metabolic pathways or when multi-subunit protein complexes need to be recombinantly expressed. These *TIRs* can, in principle, also be used to amplify laboratory evolved versions of *araC pBAD*, which detect and respond to D-arabinose (8), mevalonate (9), triacetic acid lactone (23), ectoine (24) and blue light (25).

Bioengineering applications require genetic sensor modules that have a broad dynamic range (48–52). One tried-and-tested approach to broaden dynamic range is to tune the translational efficiency of the output module by engineering the *TIR*. This has

been done previously by swapping Shine–Dalgarno sequences (53), using computational algorithms to design *TIRs de novo* (54–56), or directed/synthetic evolution processes (37–39, 44, 57, 58). A significant advantage of these approaches is that they do not change the sequence of the genetic sensor and, therefore, do not directly affect its sensing capabilities. The genetic sensor will continue to respond to the same concentrations of small molecule (input) and respond with the same transcriptional response (output). The optimized *TIR* simply functions as an 'amplifier' of the genetic sensor, by increasing the translational efficiency of the transcribed mRNA.

Supplementary data

Supplementary data are available at SYNBIO Online.

Material availability

Expression plasmids described in this study can be obtained from the nonprofit plasmid repository, Addgene.

Data availability

Unprocessed gel images and annotated plasmid sequences used are available in the Online Supplement. Microscopy images are deposited in The Cell Image Library (http://cellimagelibrary.org/groups/54692). All data described in this work are also freely available upon request at no cost and without restrictions.

Funding

Novo Nordisk Fonden [0071844 to D.O.D.]; Carl Trygger stiftelse [CTS 21:1637 to D.O.D.]; Australian Research Council [DP220101143].

Conflict of interest statement. The synthetic evolution process and the *TIRs* identified using it are patent protected (59). The patents are the property of CloneOpt AB, of which D.O.D. is a shareholder. All the other authors declare no competing interests.

References

1. Voigt,C.A. (2006) Genetic parts to program bacteria. *Curr. Opin. Biotechnol.*, **17**, 548–557.
2. Zhang,J., Jensen,M.K. and Keasling,J.D. (2015) Development of biosensors and their application in metabolic engineering. *Curr. Opin. Chem. Biol.*, **28**, 1–8.
3. Brophy,J.A.N. and Voigt,C.A. (2014) Principles of genetic circuit design. *Nat. Methods*, **11**, 508–520.
4. Hossain,G.S., Saini,M., Miyake,R., Ling,H. and Chang,M.W. (2020) Genetic biosensor design for natural product biosynthesis in microorganisms. *Trends Biotechnol.*, **38**, 797–810.
5. Rogers,J.K., Taylor,N.D. and Church,G.M. (2016) Biosensor-based engineering of biosynthetic pathways. *Curr. Opin. Biotechnol.*, **42**, 84–91.
6. Lee,S.K., Chou,H.H., Pfleger,B.F., Newman,J.D., Yoshikuni,Y. and Keasling,J.D. (2007) Directed evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters. *Appl. Environ. Microbiol.*, **73**, 5711–5715.
7. Vee Aune,T.E., Bakke,I., Drabløs,F., Lale,R., Brautaset,T. and Valla,S. (2010) Directed evolution of the transcription factor XylS for development of improved expression systems: directed evolution of XylS for improved expression systems. *Microb. Biotechnol.*, **3**, 38–47.
8. Tang,S.-Y., Fazelinia,H. and Cirino,P.C. (2008) AraC regulatory protein mutants with altered effector specificity. *J. Am. Chem. Soc.*, **130**, 5267–5271.
9. Tang,S.-Y. and Cirino,P.C. (2011) Design and application of a mevalonate-responsive regulatory protein. *Angew. Chem.*, **123**, 1116–1118.
10. Berg,L., Kucharova,V., Bakke,I., Valla,S. and Brautaset,T. (2012) Exploring the 5′-UTR DNA region as a target for optimizing recombinant gene expression from the strong and inducible Pm promoter in *Escherichia coli*. *J. Biotechnol.*, **158**, 224–230.
11. Bakke,I., Berg,L., Aune,T.E.V., Brautaset,T., Sletta,H., Tøndervik,A. and Valla,S. (2009) Random mutagenesis of the PM promoter as a powerful strategy for improvement of recombinant-gene expression. *Appl. Environ. Microbiol.*, **75**, 2002–2011.
12. Dabirian,Y., Li,X., Chen,Y., David,F., Nielsen,J. and Siewers,V. (2019) Expanding the dynamic range of a transcription factor-based biosensor in *Saccharomyces cerevisiae*. *ACS Synth. Biol.*, **8**, 1968–1975.
13. Mannan,A.A., Liu,D., Zhang,F. and Oyarzún,D.A. (2017) Fundamental design principles for transcription-factor-based metabolite biosensors. *ACS Synth. Biol.*, **6**, 1851–1859.
14. Ding,N., Yuan,Z., Zhang,X., Chen,J., Zhou,S. and Deng,Y. (2020) Programmable cross-ribosome-binding sites to fine-tune the dynamic range of transcription factor-based biosensor. *Nucleic Acids Res.*, **48**, 10602–10613.
15. Rogers,J.K., Guzman,C.D., Taylor,N.D., Raman,S., Anderson,K. and Church,G.M. (2015) Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Res.*, **43**, 7648–7660.
16. Meyer,A.J., Segall-Shapiro,T.H., Glassey,E., Zhang,J. and Voigt,C.A. (2019) Escherichia coli "Marionette" strains with 12 highly optimized small-molecule sensors. *Nat. Chem. Biol.*, **15**, 196–204.
17. Schleif,R. (2000) Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet.*, **16**, 559–565.
18. Schleif,R. (2010) AraC protein, regulation of the L-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMS Microbiol. Rev.*, **34**, 779–796.
19. Cagnon,C., Valverde,V. and Masson,J.-M. (1991) A new family of sugar-inducible expression vectors for *Escherichia coli*. *Protein Eng. Des. Sel.*, **4**, 843–847.
20. Guzman,L.M., Belin,D., Carson,M.J. and Beckwith,J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.*, **177**, 4121–4130.
21. Ceroni,F., Algar,R., Stan,G.-B. and Ellis,T. (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods*, **12**, 415–418.
22. Ceroni,F., Boo,A., Furini,S., Gorochowski,T.E., Borkowski,O., Ladak,Y.N., Awan,A.R., Gilbert,C., Stan,G.-B. and Ellis,T. (2018) Burden-driven feedback control of gene expression. *Nat. Methods*, **15**, 387–393.
23. Tang,S.-Y., Qian,S., Akinterinwa,O., Frei,C.S., Gredell,J.A. and Cirino,P.C. (2013) Screening for enhanced triacetic acid lactone production by recombinant *Escherichia coli* expressing a designed triacetic acid lactone reporter. *J. Am. Chem. Soc.*, **135**, 10099–10103.
24. Chen,W., Zhang,S., Jiang,P., Yao,J., He,Y., Chen,L., Gui,X., Dong,Z. and Tang,S.-Y. (2015) Design of an ectoine-responsive AraC mutant and its application in metabolic engineering of ectoine biosynthesis. *Metab. Eng.*, **30**, 149–155.
25. Romano,E., Baumschlager,A., Akmeriç,E.B., Palanisamy,N., Houmani,M., Schmidt,G., Öztürk,M.A., Ernst,L., Khammash,M. and Di Ventura,B. (2021) Engineering AraC to make it responsive to light instead of arabinose. *Nat. Chem. Biol.*, **17**, 817–827.

26. Balzer,S., Kucharova,V., Megerle,J., Lale,R., Brautaset,T. and Valla,S. (2013) A comparative analysis of the properties of regulated promoter systems commonly used for recombinant gene expression in *Escherichia coli*. *Microb. Cell Factories*, **12**, 26.
27. Reeve,B., Hargest,T., Gilbert,C. and Ellis,T. (2014) Predicting translation initiation rates for designing synthetic biology. *Front. Bioeng. Biotechnol.*, **2**, 1–6.
28. McCarthy,J.E. and Gualerzi,C. (1990) Translational control of prokaryotic gene expression. *Trends Genet. TIG*, **6**, 78–85.
29. Laursen,B.S., Sørensen,H.P., Mortensen,K.K. and Sperling-Petersen,H.U. (2005) Initiation of protein synthesis in bacteria. *Microbiol. Mol. Biol. Rev.*, **69**, 101–123.
30. Kozak,M. (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene*, **361**, 13–37.
31. Gualerzi,C.O. and Pon,C.L. (1990) Initiation of mRNA translation in prokaryotes. *Biochemistry*, **29**, 5881–5889.
32. Subach,F.V., Patterson,G.H., Manley,S., Gillette,J.M., Lippincott-Schwartz,J. and Verkhusha,V.V. (2009) Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nat. Methods*, **6**, 153–159.
33. Pédelacq,J.-D., Cabantous,S., Tran,T., Terwilliger,T.C. and Waldo,G.S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.*, **24**, 79–88.
34. Shaner,N.C., Lambert,G.G., Chammas,A., Ni,Y., Cranfill,P.J., Baird,M.A., Sell,B.R., Allen,J.R., Day,R.N., Israelsson,M. *et al.* (2013) A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat. Methods*, **10**, 407–409.
35. Zhang,M., Chang,H., Zhang,Y., Yu,J., Wu,L., Ji,W., Chen,J., Liu,B., Lu,J., Liu,Y. *et al.* (2012) Rational design of true monomeric and bright photoactivatable fluorescent proteins. *Nat. Methods*, **9**, 727–729.
36. Watson,J.F. and García-Nafría,J. (2019) *In vivo* DNA assembly using common laboratory bacteria: A re-emerging tool to simplify molecular cloning. *J. Biol. Chem.*, **294**, 15271–15281.
37. Mirzadeh,K., Martínez,V., Toddo,S., Guntur,S., Herrgård,M.J., Elofsson,A., Nørholm,M.H.H. and Daley,D.O. (2015) Enhanced protein production in *Escherichia coli* by optimization of cloning scars at the vector-coding sequence junction. *ACS Synth. Biol.*, **4**, 959–965.
38. Mirzadeh,K., Toddo,S., Nørholm,M.H.H. and Daley,D.O. (2016) Codon optimizing for increased membrane protein production: a minimalist approach. *Methods Mol. Biol. Clifton NJ*, **1432**, 53–61.
39. Shilling,P.J., Mirzadeh,K., Cumming,A.J., Widesheim,M., Köck,Z. and Daley,D.O. (2020) Improved designs for pET expression plasmids increase protein production yield in *Escherichia coli*. *Commun. Biol.*, **3**, 214.
40. Liu,H. and Naismith,J.H. (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.*, **8**, 91.
41. *pBAD/His A, B, and C pBAD/*Myc-*His A, B, and C*. Vectors for Dose-Dependent Expression of Recombinant Proteins Containing N- or C-Terminal 6×His Tags in *E. coli*. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/pbad_man.pdf (1 July 2021, date last accessed).
42. Daley,D.O., Rapp,M., Granseth,E., Melén,K., Drew,D. and von Heijne,G. (2005) Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science*, **308**, 1321–1323.
43. Postma,M. and Goedhart,J. (2019) PlotsOfData—a web app for visualizing data together with their summaries. *PLOS Biol.*, **17**, e3000202.
44. Cumming,A.J., Khananisho,D., Harris,R., Bayer,C.N., Nørholm, M.H.H., Jamshidi,S., Ilag,L.L. and Daley,D.O. (2022) Antibiotic-efficient genetic cassette for the TEM-1 β-lactamase that improves plasmid performance. *ACS Synth. Biol.*, **11**, 241–253.
45. Siegele,D.A. and Hu,J.C. (1997) Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc. Natl. Acad. Sci.*, **94**, 8168–8172.
46. Khlebnikov,A., Risa,O., Skaug,T., Carrier,T.A. and Keasling,J.D. (2000) Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture. *J. Bacteriol.*, **182**, 7029–7034.
47. Yoo,S.M., Jung,S.-W., Yeom,J., Lee,S.Y. and Na,D. (2020) Tunable gene expression system independent of downstream coding sequence. *ACS Synth. Biol.*, **9**, 2998–3007.
48. Zhang,F., Carothers,J.M. and Keasling,J.D. (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.*, **30**, 354–359.
49. Nguyen,N.H., Kim,J.-R. and Park,S. (2018) Application of transcription factor-based 3-hydroxypropionic acid biosensor. *Biotechnol. Bioprocess Eng.*, **23**, 564–572.
50. Skjoedt,M.L., Snoek,T., Kildegaard,K.R., Arsovska,D., Eichenberger,M., Goedecke,T.J., Rajkumar,A.S., Zhang,J., Kristensen,M., Lehka,B.J. *et al.* (2016) Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. *Nat. Chem. Biol.*, **12**, 951–958.
51. Cheng,F., Tang,X.-L. and Kardashliev,T. (2018) Transcription factor-based biosensors in high-throughput screening: advances and applications. *Biotechnol. J.*, **13**, 1700648.
52. Kasey,C.M., Zerrad,M., Li,Y., Cropp,T.A. and Williams,G.J. (2018) Development of transcription factor-based designer macrolide biosensors for metabolic engineering and synthetic biology. *ACS Synth. Biol.*, **7**, 227–239.
53. Levin-Karp,A., Barenholz,U., Bareia,T., Dayagi,M., Zelcbuch,L., Antonovsky,N., Noor,E. and Milo,R. (2013) Quantifying translational coupling in *E. coli* synthetic operons using RBS modulation and fluorescent reporters. *ACS Synth. Biol.*, **2**, 327–336.
54. Salis,H.M., Mirsky,E.A. and Voigt,C.A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.*, **27**, 946–950.
55. Seo,S.W., Yang,J.-S., Kim,I., Yang,J., Min,B.E., Kim,S. and Jung,G.Y. (2013) Predictive design of mRNA translation initiation region to control prokaryotic translation efficiency. *Metab. Eng.*, **15**, 67–74.
56. Na,D. and Lee,D. (2010) RBSDesigner: software for designing synthetic ribosome binding sites that yields a desired level of protein expression. *Bioinforma. Oxf. Engl.*, **26**, 2633–2634.
57. Mirzadeh,K., Shilling,P.J., Elfageih,R., Cumming,A.J., Cui,H.L., Rennig,M., Nørholm,M.H.H. and Daley,D.O. (2020) Increased production of periplasmic proteins in *Escherichia coli* by directed evolution of the translation initiation region. *Microb. Cell Factories*, **19**, 85.
58. Rennig,M., Martinez,V., Mirzadeh,K., Dunas,F., Röjsäter,B., Daley,D.O. and Nørholm,M.H.H. (2018) TARSyn: tunable antibiotic resistance devices enabling bacterial synthetic evolution and protein production. *ACS Synth. Biol.*, **7**, 432–442.
59. Daley,D., Mirzadeh,K., Toddo,S. and Guntur,S. (2016) Selective optimisation of a ribosome binding site for protein production. EP3234146; US10,696,963.