Synthetic Biology, —, pp. 1–12

doi: —
Advance Access Publication Date: Day Month 2024
Paper

PAPER

Biosensors Characterization: Formal methods from the Perspective of Proteome Fractions

Nicolás A. Vaccari,[1] Dahlin Zevallos-Aliaga,[1] Tom Peeters[2] and Daniel G. Guerra[1,*]

[1]Laboratorio de Moléculas Individuales, Laboratorios de Investigación y Desarrollo, Facultad de Ciencias e Ingeniería, Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430, 15102, Lima, Peru and [2]Open BioLab Brussels, Erasmushogeschool Brussel, Quai de l'Industrie 170, 1070 Anderlecht, Brussels, Belgium
*Corresponding author: Daniel G. Guerra - daniel.guerra@upch.pe
FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Many studies characterize transcription factors and other regulatory elements to control the expression of heterologous genes in recombinant systems. However, most lack a formal approach to analyse the parameters and context-specific variations of these regulatory components. This study addresses this gap by establishing formal and convenient methods for characterising regulatory circuits. We model the bacterial cell as a collection of a small number of proteome fractions. Then, we derive the proteome fraction over time and obtain a general theorem describing its change as a function of its expression fraction, which represents a specific portion of the total biosynthesis flux of the cell. Formal deduction reveals that when the proteome fraction reaches a maximum, it becomes equivalent to its expression fraction. This equation enables the reliable measurement of the expression fraction through direct protein quantification. In addition, experimental data demonstrate a linear correlation between protein production rate and specific growth rate over a significant time period. This suggests a constant expression fraction within this window. The expression fractions estimated from the slopes of these intervals and those obtained from maximum protein amount points can both be independently fitted to a Hill function. In the case of an IPTG biosensor, in five cellular contexts, expression fractions determined by the maximum method and the slope method produced similar dose-response parameters. Additionally, by analysing two more biosensors, for mercury and cumate detection, we demonstrate that the slope method can be effectively applied to various systems, generating reliable Hill function parameters.

Key words: proteome fraction, regulation, biosensors

Introduction

Synthetic biology relies on the assumption that biological systems can be engineered by design, assembling parts and circuits like modules to create new systems of increasing complexity [1, 2]. However, regulatory circuits and even individual components often exhibit variable behaviours depending on the cellular context or culture conditions [3, 4, 5, 6]. In addition, research groups employ different techniques and data analysis methods when characterising these regulatory elements experimentally. This variation further hinders the ability to exchange parts in a modular approach.

Despite these challenges, the field has grown steadily. New biological parts are constantly being incorporated for novel engineered functions, leading to practical successes at different levels. For example, transcription factors and DNA or RNA sequences can link intracellular metabolites or extracellular signals to gene expression, controlling native or heterologous genes [7, 8, 9]. This strategy is applied in creating new biosynthetic pathways [10, 11], identifying optimal strains for the production of a molecule of interest[12, 13], and developing biosensors for various applications such as environmental monitoring, food safety, and disease diagnosis [14, 15, 16, 17]. Furthermore, the activity of specific pathways can be regulated through transcription factors or sensing RNA elements, thus introducing a dynamic control to optimise metabolic flux toward the desired biosynthesis [18, 19]. There are a large number of characterised transcription factors, evaluated by their dose-response, detection limit, sensitivity, and dynamic range [20, 21, 19, 22]. However, the lack of standardised procedures to analyse their ability to regulate reporter genes hinders progress in establishing robust libraries of modular tools. For example, there is no consensus on how to select culture time points that allow reliable measurement averaging between replicates or facilitate comparisons of output under different conditions, cellular contexts, and laboratories [23, 24].

This study addresses this challenge by proposing a theoretical framework for the characterisation of regulatory circuits. This framework is based on the perspective that the cell is composed of a few proteome fractions, a concept that was used to study bacterial metabolism and growth regulation [25, 26, 27, 28, 29, 30, 31, 32]. By mathematically defining the proteome fraction and deriving it over time, we introduce a theorem that links the evolution of a set of proteins to a

© The Author —. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

specific fraction of the cell's biosynthesis rate, which we call the expression fraction. We formally demonstrate that this expression fraction can be quantified by two alternative methods: directly by analysing a single point of peak accumulation of reporter protein; or through the estimation of a linear relationship between specific production and specific growth rates during a given interval. We tested the reciprocal consistency of these two methods experimentally by comparing the dose-response profile of different *Escherichia coli* strains induced by varying IPTG concentrations and two other biosensors based on transcription factors, for the detection of mercury and cumate.

We consider that the framework presented in this article helps to tackle well-known challenges that are still prevalent in synthetic biology today [3, 33]. Firstly, it offers readily reproducible methodologies for characterising genetic circuits and their components, relying on easily measurable data such as culture growth and reporter protein levels. Secondly, our approach simplifies the mathematical modelling of genetic regulation systems by masking complexity through the concept of proteome fractions, while still capturing the interplay between native and heterologous cellular functions. This facilitates a deeper understanding of the inherent properties of sensing circuits, differentiating them from characteristics that depend on the cell and culture context. Finally, we speculate on how to extend this analysis to investigate orthogonality and potential crosstalk between heterologous and native genetic components.

Methods

E. coli strains

MG1655 ΔendA ΔrecA (DE3) was a gift from Kristala Prather (Addgene code # 37854) [34]. BLR(DE3) and BL21(DE3) agar stabs were generously provided by Carlos Bustamante at UC Berkeley and Mauricio Baez at the Universidad de Chile, respectively. Cells were rendered competent by the $CaCl_2$ incubation protocol and transformed by 42 °C heat shock with pUC-T7GFP or simultaneously with pUC-T7GFP and pACYCDuet-1. Luria broth (LB) selection agar plates were supplemented with glucose 1% to reduce the basal expression of T7 and ampicillin to select cells transformed with the plasmid pUC-T7GFP, or ampicillin and chloramphenicol to select those with the plasmids pUC-T7GFP and pACYCDuet-1.

Plasmids

Sequences were designed using Benchling RRID:SCR_013955 [Biology Software] (2020). Retrieved from https://benchling.com. Twist Bioscience synthesised a gene block containing the following: T7 promoter, lac operator, Shine-Dalgarno ribosome binding sequence, a codon-optimised GFPmut3 coding sequence [35] fused with the LVA tag for rapid degradation (AANDENYLVA) [36], and the T7 terminator. The complete sequence is available in GenBank, accession number PQ015608. A high copy number GFP expression plasmid, pUC-T7GFP, was built by inserting this construct into the pUC57 vector through standard digestion and ligation reactions between the restriction sites NcoI and KpnI. To introduce additional expression of LacI, cells were transformed with the low-copy plasmid pACYCDuet-1, acquired from GenScript.

The mercury biosensor was constructed by inserting the *merR* gene, which encodes the MerR mercury sensitive transcription factor, and the promoter from transposon Tn501 from *Pseudomonas aeruginosa* plasmid pVS1 (GenBank: Z00027.1), containing the MerR DNA binding site, in the pUC57 vector. This enabled conditional expression of the *RFP* reporter gene in response to mercury. NEB® Stable *E. coli* competent cells were transformed with the resulting plasmid. The complete design, cloning procedures, and characterisation are described in [37].

A cumate biosensor was constructed based on a regulatory circuit adapted from Choi et al. [38]. Briefly, a constitutively expressed CymR protein represses a promoter that controls the expression of GFPmut3 fused with the LVA tag for rapid degradation (AANDENYLVA). Cumate-induced dissociation of CymR activates GFP expression. The circuit sequence was synthesised as a gBlock by Integrated DNA Technologies, Inc. and cloned at the NcoI and NotI sites of the pUC57 vector. The resulting plasmid was transformed into BL21(DE3), BLR(DE3) and MG1655(DE3) competent cells. The complete sequence is available in GenBank, accession number PQ010742.

Culture conditions

The Specific growth rates and fluorescence data presented here were obtained from 200 μL microcultures in M9 medium supplemented with 0.2% casamino acids and 0.24% glucose, prepared as follows: Freshly transformed colonies were handpicked to inoculate 10 mL of M9 medium supplemented with ampicillin, or ampicillin and chloramphenicol when pACYCDuet-1 was included. These cultures were incubated at 37 °C and 250 RPM overnight, their cell density was measured and harvested by centrifugation at 5000 RPM for 5 minutes. Immediately, cells were resuspended in fresh medium to the appropriate volume to obtain an optical density of 0.1 (approximately 2 x 10^7 cells·mL^{-1}) and were loaded in a 96 microtiter plate. At this point, reporter protein expression was induced in 195 μ L of culture by manually pipetting 5 μL of the appropriate inducer (IPTG, $HgBr_2$, cumate), reaching the final concentration indicated for each case, completing 200 μL in each well. These microcultures were incubated at 30 °C with oscillatory shaking at 220 RPM for 16 hours without the addition of antibiotics. The microtiter plate was covered with its plastic lid and the outer wells were filled with 200 μL of water so that a similar humidity surrounded all monitored cultures. This configuration resulted in a volume loss of less than 2. 5% due to evaporation after 16 hours.

Measurements

Culture density (OD_{600}) and fluorescence were monitored in a plate reader (Tecan Infinite 200 Pro). The production of GFP was registered with the following settings: 475 nm excitation and 516 nm emission, with the signal gain set at 50 out of 100. Similarly, the production of RFP was recorded using 585-nm excitation and 608-nm emission, with the signal gain set at 95. In both cases, fluorescence data were collected with 25 flashes, reading every 15 minutes.

Total protein quantification and visualisation

For this assay, two flasks (for each strain), containing 50 mL of transparent medium, were inoculated with the bacteria so that the final OD_{600} was 0.1. One of the flasks was induced with IPTG (0.5 mM) at the beginning of the culture. The cultures were then left to grow at 30 °C. Cultures were sampled at three time points: 0 h, 2 h and 4 h. For each of the samples taken during bacterial growth, two tests were carried out. Quantification of total proteins using the Bradford reagent and visualisation of total proteins using a polyacrylamide gel. For the total protein quantification essay, 2 mL of culture was taken and the bacteria were pelleted at 13500 rpm for 5 min. This pellet was then resuspended in 500 μL of 8M Urea and lysed by three cycles of freeze thawing for 15 min at -70°C and 15 min at 37 ° C. Finally, 10 μL of this suspension was mixed with 300 μL of Bradford reagent (Thermo Fisher Scientific), incubated 10 min at room temperature and measured at 595 nm for protein quantification. For visualisation of the total proteins essay, 2 mL of culture was pelleted, resuspended in 500 μL of 2% SDS and heated at 95 ° C for 10 min. Subsequently, this suspension was run on a 15% polyacrylamide gel.

Estimation of the heterologous fraction, specific growth rate and specific production rate rate

We considered fluorescence directly proportional to the mass of the reporter protein (GFP or RFP) and the optical density proportional to the total protein mass. Thus, the ratio of fluorescence to optical density, $\frac{flu(t)}{OD_{600}(t)}$, was measured to estimate the heterologous fraction, φ_H.

To calculate the specific growth rate at each time point $\mu(t)$, the optical density was measured at t and at the immediate next time point $t+\Delta t$. The difference between both measurements was divided by Δt and the optical density at the former point t:

$$\mu(t) = \frac{OD_{600}(t+\Delta t) - OD_{600}(t)}{\Delta t} \cdot \frac{1}{OD_{600}(t)} \tag{1}$$

The specific production rate of *GFP* (or *RFP*) was estimated by normalizing the slope of fluorescence accumulation over time. For GFPmut3, the fluorophore maturation time is 4.1 minutes[35], therefore, no maturation delay was introduced in the calculations as it is shorter than our measurement time window of 15 minutes. On the other hand, the merR system used *RFP* as a reporter which has a maturation time of 1 hour, meaning that the red fluorescence that was observed in any given time is proportional to the amount of *RFP* produced 4 measuring points earlier. We corrected this by shifting the fluorescence curve back in time by 4 points. Finally, we calculated ρ_H by measuring the fluorescence at time t and at next time point $t+\Delta t$. We divide the difference by time step Δt and optical density at time t:

$$\rho_H(t) = \frac{flu(t+\Delta t) - flu(t)}{\Delta t} \cdot \frac{1}{OD_{600}(t)} \tag{2}$$

Calculations

Data handling, processing, plotting, regression analysis, and modeling were performed using the R programming language. All linear regressions were performed using R's built-in `lm` and `nls` functions. The data set employed was uploaded in figshare (DOI: https://doi.org/10.6084/m9.figshare.26314723.v1) [39]. To determine the expression fractions using the maximum method, the time series data for φ_H were grouped by strain and induction condition (N=8 for BL21(DE3) with the T7 system, N=9 for the MerR and CymR systems, and N = 4 for all other cases). Subsequently, φ_H^{max} was identified for each series within each group and these values were averaged. The average φ_H^{max} was obtained at various inducer concentrations and then used in the Hill fitting process.

For the slope method, the expression fraction was determined as the slope obtained from a linear regression between the specific growth rate and the specific production rate during growth (μ ¿ 0). To establish the appropriate time window for this regression, all specific growth rate data for each strain under each induction condition were analysed. At each time point, a simple t test was used to assess whether the average value of the specific growth rate (N=8 for BL21(DE3) with the T7 system, N=9 for the MerR and CymR systems and N = 4 for all other cases) differed significantly from zero. If a significant difference was found ($\alpha = 0.05$), the time point was considered part of the growth phase and included in the regression analysis. Subsequently, the earliest data points were iteratively excluded one by one until the determination coefficient (R^2) reached a maximum value, which was greater than 0.9 in all cases. The slopes of the regressions obtained at various inducer concentrations were registered and used to fit the Hill function.

Dose-response profiles were used to build Hill functions for each biosensor. A scatterplot was created depicting the expression fractions data (obtained from slopes or φ_H^{max} values) versus the inducer concentration for each titration experiment. The `nls` function with the port algorithm was then used to fit a Hill function to this plot. Lower bounds of 0 were set for the values of H (maximum expression) and k_I (affinity constant), while a lower bound of 1 was set for the value of n (Hill coefficient). The weight of each data point was determined by the inverse of its variance, which assigned a higher weight to more precise measurements.

Results

DE3 system illustrates the challenges of biosensors characterisation

We began our investigation experimenting with the DE3 expression system in B and K-12 *E. coli* strains. This system combines a native element, the LacI repressor, and a heterologous one, the T7 RNA polymerase (T7RNAP)[40], and is widely applied in the laboratory-scale production of recombinant proteins [41]. In addition, the LacI repressor served as a model for biosensor studies [42], and both the LacI repressor and T7RNAP have been used as modular components in the construction of logic gates within synthetic regulatory circuits [43, 44, 45].

For our DE3 system experiments, we used a plasmid containing the GFP gene under the control of the T7 promoter and a LacI operator. Using this setting as an IPTG biosensor, we employed a two-stage culture protocol. The first stage was focused on the accumulation of biomass, then in a subsequent phase, starting at $OD_{600} = 0.1$, the expression was induced by IPTG addition. This configuration revealed variations in specific growth rates and GFP output both between different *E. coli* strains and between those with and without additional expression of the LacI repressor (introduced through the plasmid pACYCDuet), as shown in Figure 1.

Although differences between cultures are visible, analysing them requires applying objective criteria that are not self-evident. A common option to compare biosensors is to analyse their results at the peak of their specific growth rate, which does not necessarily align at any particular time of culture in the case of different strains. In our case, even though all cultures started at the same cellular density and were stimulated with IPTG at the same time, BLR(DE3) and MG1655(DE3) showed their maximum specific growth rates later than BL21(DE3) as shown in Figure 1. Consequently, if these points are selected, the production of GFP by BL21(DE3) and BLR(DE3) with additional LacI would be plotted just after 30 minutes of induction, when the fluorescence is still relatively low, while those of BLR(DE3) and MG1655(DE3) would be evaluated after longer accumulation (Figure 1 C). Although the peak of the specific growth rate is a common reference for metabolism and regulation studies, in our case analysing the protein yield at these time points would lead to unfair comparisons. This misrepresents the different strains and could imply wrong interpretations of the circuit's properties. For example, strain MG1655(DE3) would appear to be a producer similar to B strains simply because the latter are analysed after a shorter protein accumulation.

This case illustrates that analysing a culture's response to heterologous expression requires identifying a time point that offers biologically meaningful information, taking into account the induction strategy and the culture's evolution over time. When aiming for maximum yield or signal, experimenters often choose the stationary phase to analyse the accumulated protein. Alternatively, through trial and error, analysis may be shifted to an earlier time point, as is common in the case of recombinant protein production, where short induction times minimise the potentially undesirable consequences of heterologous overexpression [46, 47, 48], such as toxicity and protein misfolding. In the case of developing biosensors for analyte detection or for dynamically controlling engineered metabolic pathways, it is preferable to centre the analysis during exponential growth, near the peak of the specific growth rate, as this is when the anabolic capacity is at its maximum. In addition, studies of bacterial physiology use the growth peak as a reference because this is when bacterial metabolism reaches an

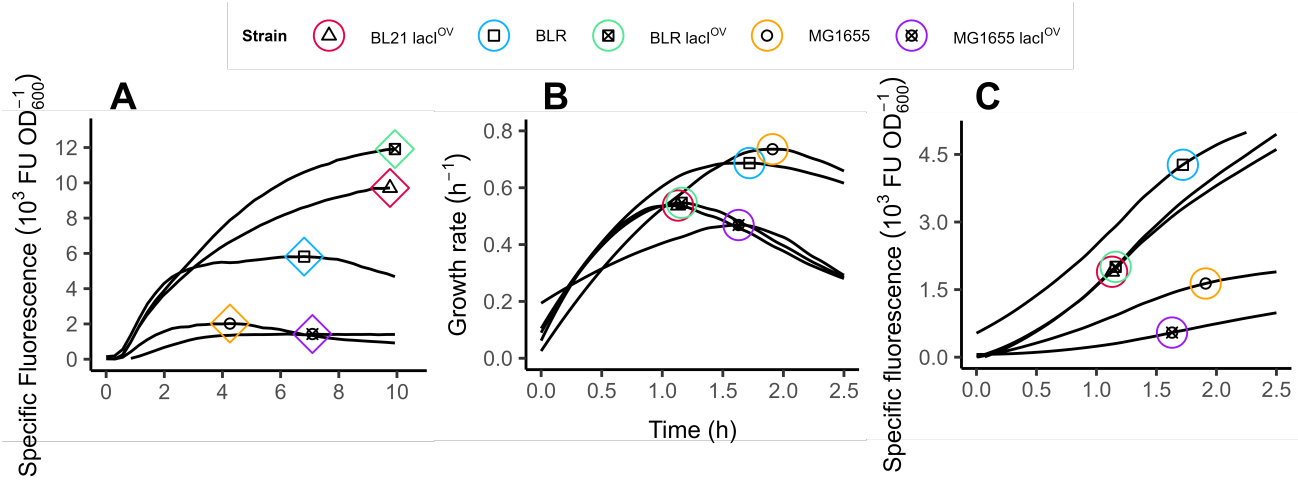


Figure 1. Analysing the heterologous proteome fraction φ_H at different time points. A) The specific fluorescence in time represents the accumulation of GFP over 10 hours of cultivation after induction by 1 mM IPTG. Diamonds indicate the point of maximum specific fluorescence, φ_H^{max}. B) Specific growth rate of the same systems; the circles indicate the point of maximum specific growth rate in each culture. C) The specific fluorescence in the same experiment, within the first 2.5 hours. ALT TEXT: images comparing the normalised fluorescence when measured at the moment of peak growth rate vs measuring its maximum value.

optimum within the constraints imposed by external conditions and internal regulations. However, as stated above, focussing on the time point of maximum specific growth rate might not be informative in the case of comparing different strains in scenarios such as two-stage cultures, consisting of a biomass accumulation phase and a consecutive production phase. Given this circumstance, a new criterion is needed to select measurement points to reliably study genetic circuits, such as biosensors, in a manner that enables accurate comparisons and reproducibility.

A theorem for the evolution of proteome fractions in time: the concept of expression fraction

We argue that when analysing the production of heterologous proteins in general, and whole cell biosensors in particular, the focus should shift from protein yield to biosynthesis flux allocation. We conceptualised the general biosynthesis rate divided into a set of rates (Figure 2), each producing a different type of protein, termed a proteome fraction (φ_i). Similarly as has been done in models previously reported [26, 28, 29, 25, 30, 27, 31], proteins will be grouped together as a unified proteome fraction if they share a similar biological function and are regulated by the same cues [25, 32]. In previous work, a proteome fraction that fulfils the core structural functions will remain constant (φ_Q), while the other two fractions are variable and perform nutrient assimilation ($\varphi_C(t)$) and biomass production ($\varphi_R(t)$), respectively. Finally, in the case of one set of heterologous proteins controlled by a single circuit, these will constitute a fraction called the heterologous fraction ($\varphi_H(t)$). The total anabolic flux will be divided into the synthesis of specific protein sets, in proportions denoted as the expression fractions, represented by $f_i(t)$, where i is the protein group to which it is dedicated (f_R, f_Q, f_H, f_C in figure 2D). In this way, the expression fraction f_i can be understood as the proportion of total ribosomal activity that synthesises the proteome fraction φ_i. The value of $f_i(t)$ will depend on the specific regulatory mechanism of the protein set and the conditions of the cell. The full conceptual and mathematical reasoning behind the expression fraction is detailed in the supplementary material.

With these definitions, solving the derivation of a proteome fraction over time (complete derivation in the Biosynthesis Allocation Theorem section of the supplementary material) produces an equation that links the change in a proteome fraction to its expression fraction and the total biosynthesis rate, $\rho(t)$, as follows:

$$\frac{d\varphi_i(t)}{dt} = \rho(t)\cdot[f_i(t) - \varphi_i(t)] \tag{3}$$

In the subsequent sections, we will explain how $f_i(t)$ can be determined from easily obtainable experimental data.

Finding the expression fraction and its regulation: The maximum method

A remarkable consequence emerging from equation 3 is that for any proteome fraction, when its differential over time equals zero ($\frac{d\varphi_i(t)}{dt} = 0$), the expression fraction and the proteome fraction must be equal $f_i(t) = \varphi_i(t)$ (assuming that the cell's anabolic rate is never zero as long as it is alive). Therefore, regardless of the growth behaviour of a culture, registering the heterologous fraction when it has reached a maximum should correctly represent the expression of the circuit. Data collected at this point under various conditions can be used reliably to build a function that represents the regulated expression of the circuit, for example, in response to the variation of the inducer concentration.

It is generally accepted that gene expression responds to the concentration of an inducer following a Hill function [49], a model commonly applied to biosensors [42]. Using a Hill equation for representing the regulation of the heterologous fraction as a function of the inducer I, we obtain the following equation (for more details see supplementary information):

$$\varphi_H^{max}(I) = f_H(I) = H\cdot\frac{I^n}{(H+1)\cdot I^n + {k_I}^n} \tag{4}$$

Here, H represents the maximum possible expression for the circuit, I is the molecular inducer concentration, k_I is related to the affinity for I, and n is the Hill coefficient that expresses the cooperativity of the response to I.

Now we will use the values of $f_i(t)$obtained from the maximum values of the heterologous fraction, φ_H^{max}, as described above, to test if they follow a dose-response profile in the shape of a typical Hill function. First, we verified by SDS PAGE analysis that even for the highest GFP overexpression, φ_H represented less than 15% of total biomass (Figure 5S); therefore, we infer that in most cases the coefficient $(H+1)$ can be approximated to 1, thus resulting in a typical Hill function.

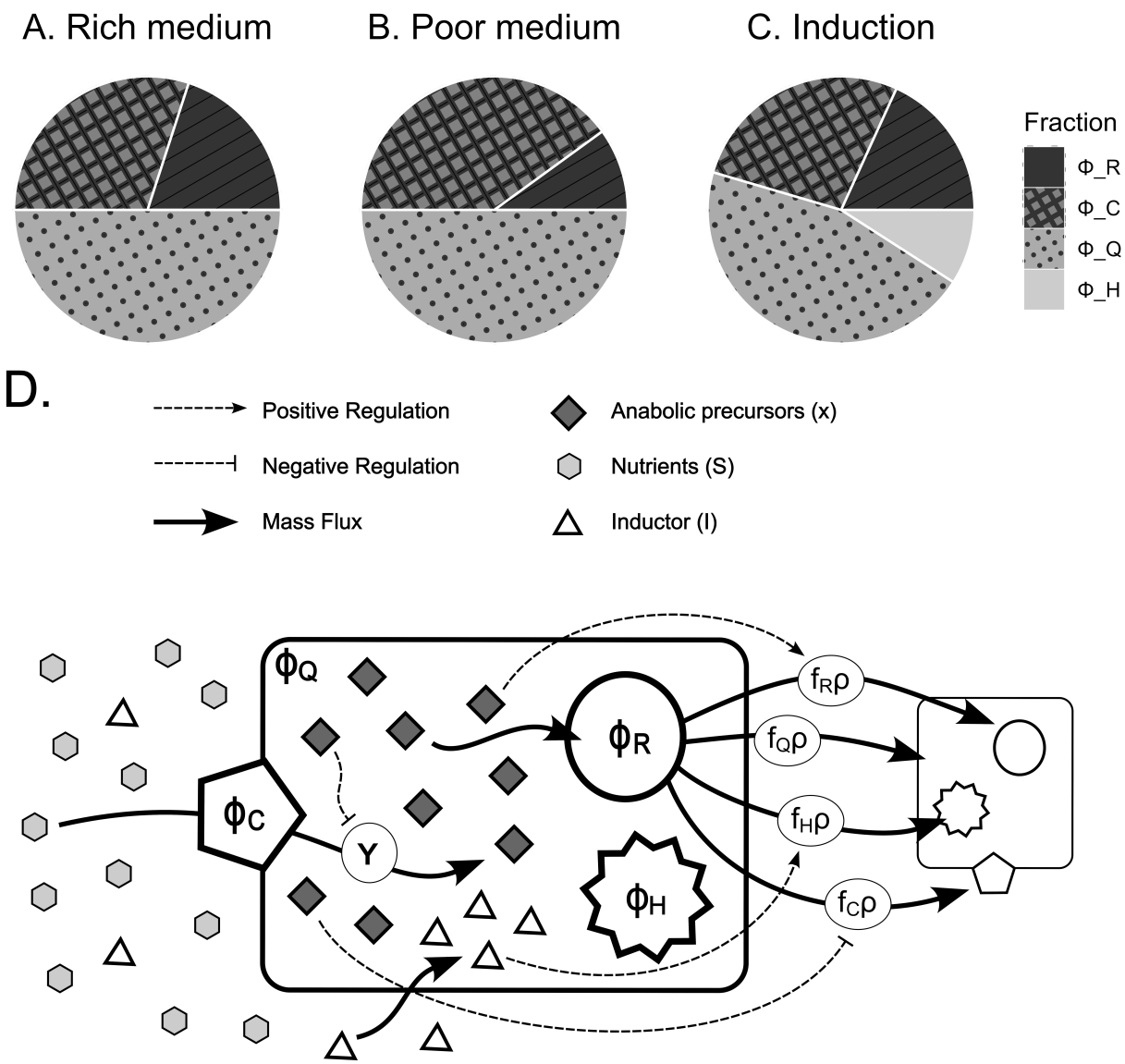


Figure 2. A model of proteome allocation and metabolic rate. Representation of the proteome fractioning of a cell: A) in a rich medium that allows the cell to maximize its specific growth rate through the activity of ribosomal fraction, φ_R; B) in a poor medium, where the cell maximizes its fixation of nutrients through the activity of the carbon fixator fraction, φ_C; C) producing a heterologous fraction, φ_H, causing a redistribution of the native proteome allocation. D) Representation of rates within bacteria. Biochemical transformations are represented by solid arrows. The light gray hexagons represent the nutrients $S(t)$, which are converted into anabolic precursors, $x(t)$, (dark gray diamonds) by the nutrient fixator, φ_C (pentagon) at a rate Y. The anabolic precursors are then used by the ribosomal fraction, φ_R (circle), at a rate ρ, to produce new biomass (represented by the small cell on the right side). Biosynthesis is divided into a multi-outlet branching of specific protein production rates (black arrows), each one with their respective expression fractions, $f_R(t)$, $f_C(t)$, $f_Q(t)$, $f_H(t)$. Dashed arrows and dashed blunt arrows represent positive and negative regulation, respectively. The concentration of anabolic precursors stimulates the specific rate of φ_R production and represses the specific rate of φ_C production; the specific rate of φ_Q production is constant and the specific rate of φ_H production is activated by its specific inducer (triangles), which diffuses from the extracellular space. ALT TEXT: Graphical representation of the proteome fractions and idealisation of the cell's activity as compound by a set of fluxes for carbon fixation and protein biosynthesis

Importantly, because in this kind of titration experiment, the curves represent the behaviour of the whole cell, k_I and n will not strictly correspond to the dissociation constant and the molecular cooperativity of the binding reaction between the transcription factor and the inducer.

For the DE3 expression system, the expression fractions, derived from experimental φ_H^{max} values, related to the inducer concentration according to a Hill function. We achieved excellent fits in five different cell contexts, including three distinct *E. coli* strains with and without an additional LacI repressor (Figure 3). The parameters obtained from these fits are listed in Table 1. In contrast, extracting normalised fluorescence values at the time of maximum specific growth rate did not yield a Hill function in any of the five cases analysed. This is likely because the inducer (IPTG) is not yet equilibrated with the regulatory machinery at this point (which will be shown by the non-linear regions

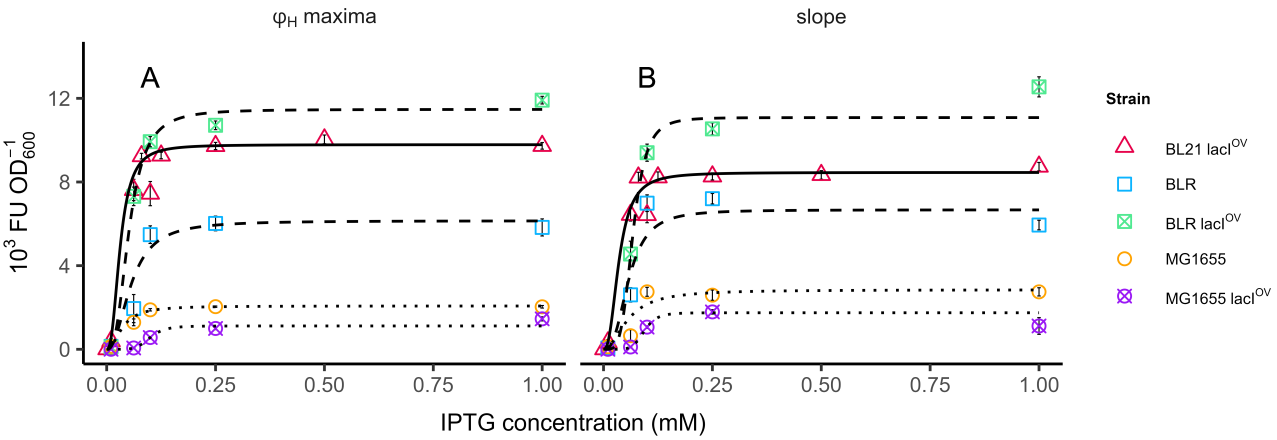


Figure 3. Expression fractions in response to the inducer, IPTG. A) Dose-reponse diagram made by plotting the maximum values of specific fluorescence (fluorescence over cellular density) obtained with the indicated IPTG concentration concentration in 5 different cellular contexts. B) Dose-reponse diagram made by plotting the slopes obtained in the linear regressions of the specific production vs specific growth graphs against its corresponding IPTG concentration. For both graphs the lines show the Hill function fitted for each data set. All strains bear the DE3 insert. Cells co-transformed with pACYCDuet for the additional expression of LacI are indicated as $lacI^{OV}$. (N = 8 for BL21(DE3), and N = 4 for all other cases)
ALT TEXT: Hill plots using slope method and maximum method for various strains with the DE3 system

in the plots of Figure 4). Therefore, observations made at the time of maximum specific growth rate may not always accurately represent the behaviour of the biosensor regulatory circuit.

Finding the expression fraction during growth: The slope method.

In the cases analysed, we observed that the maximum point for the heterologous fraction is reached close to or during the stationary phase. Because there are physiological changes in the onset of this phase, there is no guarantee that the biosynthesis fraction dedicated to our gene of interest will remain constant in other stages of the culture. Moreover, as stated previously, it is important for biosensor developers, especially in the field of metabolic engineering, to have a solid method to characterise heterologous expression during growth.

We aim to develop an effective analysis strategy by linking the expression fraction to the specific growth rate. By definition, the biosynthesis rate of the heterologous protein fraction, $\rho_H(t)$, is the product of its expression fraction, f_H and the total protein biosynthesis rate, $\rho(t)$:

$$\rho_H(t) = f_H(t) \cdot \rho(t) \tag{5}$$

During exponential growth, the protein-to-total cellular-mass ratio remains relatively constant [50]. Therefore, we can confidently assume that the total protein biosynthesis rate remains linear to the specific growth rate, which implies that we can substitute $\rho(t)$ with $\mu(t)$ without altering the proportionality of change. Since measuring the specific growth rate directly is simpler than determining the total biosynthesis rate of the cell, this approximation yields a convenient equation:

$$\rho_H(t) = f_H(t) \cdot \mu(t) \tag{6}$$

It follows that the f_H can be experimentally determined as equal to the slope in a plot of the specific production rate of the heterologous protein vs specific growth rate of the culture. As an approximation, we graphed the fluorescence specific production rate (ρ_H) against the specific growth rate (μ) at all time points in Figure 4 . Following induction, our initial data points from each culture show a curved line (grey dots) resembling a hook. This curved line transitions into a straighter segment (colored dots) that can be approximated by a linear relationship. We interpret this initial curvature as the time it takes for the inducer and the regulatory machinery to reach a steady state. The following straighter segment, on the other hand, represents a period with a constant slope or expression fraction (f_H) of the heterologous protein.

We used the resulting f_H values to generate a dose-response diagram and fit a Hill function, similar to the analysis performed with the values obtained from the φ_H^{max} method (Figure 3).

Another important step is to investigate whether the f_H values obtained from φ_H^{max} are similar to those obtained from the slopes. This would indicate that f_H remains constant throughout the interval in which the linear trend is maintained, including a significant portion of exponential growth and the beginning of the stationary phase. In Figure 4 (and Figure 1S) we present a visual comparison of the values f_H determined from the slope of the specific production versus the specific growth rate (dashed lines) and from the maximum heterologous fraction, φ_H^{max} (solid lines). The similarity of the results obtained through these two different methods in the case of the DE3 system suggests that the expression fractions are indeed constant within these intervals, including the beginning of the stationary phase, where the values of φ_H^{max} were identified.

Furthermore, it will be important to compare the Hill parameters (H, k_I, n) obtained by fitting the f_H values from both the maximum method and the slope method. Table 1 presents the Hill parameters obtained by fitting these values f_H as a function of inducer concentration in different expression systems.

Application on other biosensors (non T7RNAP)

The DE3 expression system depends on transcription driven by T7RNAP, which is suitable for various applications but somewhat restricted in terms of regulatory possibilities. In contrast, numerous transcription factors and other regulatory elements have been identified in relation to bacterial RNAP. To evaluate the broader applicability of our methods, we examined inducer titration experiments from two additional systems using transcription by *E. coli* RNAP.

First, a mercury biosensor was constructed based on the MerR transcription factor (complete characterisation in [37]). Upon ionic mercury titration, we observed that the reporter protein accumulated continuously, even during the stationary phase; therefore, we were unable to identify φ_H^{max}. However, the specific production vs specific growth rate graphs (Figure 5) produced linear shapes. A linear region was detected

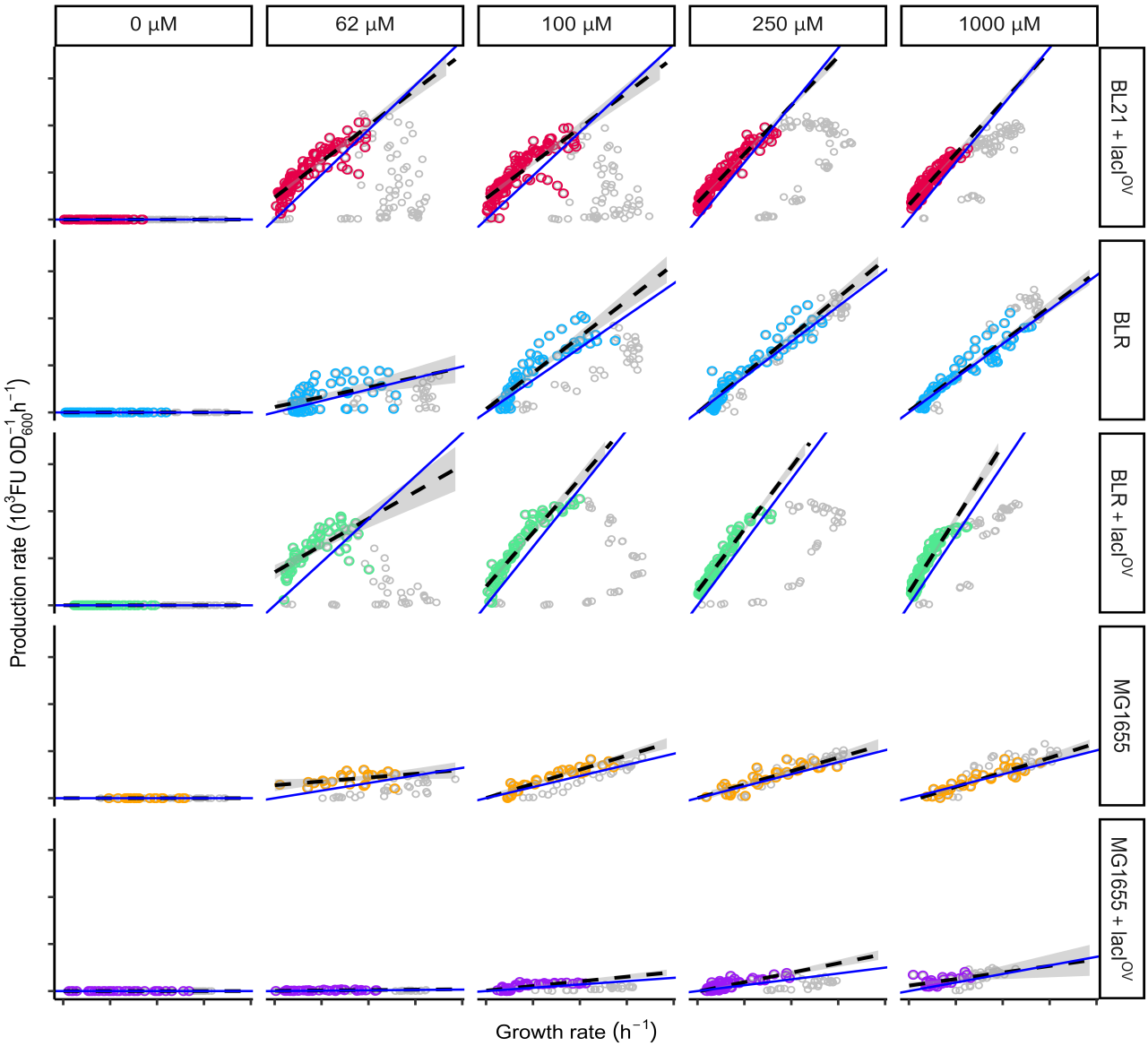


Figure 4. Expression fractions at different IPTG concentrations as obtained from two analysis methods. Experimental data taken after induction are directly plotted and then a time window is selected (colored dots) by excluding the earlier points (grey points) to be fitted to a linear regression (dashed line), maximizing R correlation coefficient as described in the main text. Each solid line represents a line function where the slope is set to be equal to the value of φ_H^{max} found at the indicated condition and strain. ALT TEXT: DE3 system - Scatter plots of specific production rate vs specific growth rate. In each plot, a selected interval was used to make a linear regression. The blue line represents the maximum specific fluorescence represented as a slope.

in the plots and was approximated to a single slope that intersected at zero. A dose-response plot was plotted using the expression fractions obtained from these slopes and was fitted to a Hill function (Figure 4S).

In addition, a cumate biosensor was built based on the constitutive expression of CymR repressor as seen in previous reports [38]. In the case of cultures induced with cumate, the reporter protein production decreased earlier than the arrest of growth, that is, before the onset of the stationary phase. Consequently, a φ_H^{max} was readily identified. Also, we observed that the specific production vs specific growth rate graphs exhibited a very short lapse of nonlinear behaviour, indicating that the intracellular concentration of cumate equilibrates relatively rapidly. The linear lapse in the relation between specific production and specific growth rates showed an intersection different from zero on the specific growth rate axis (Figure 5 A), thus representing the period when specific production ceased while growth continued ($\mu > 0$). Therefore, the regression in the cumate system must include the value of this intersect, thus implying that the equation for the specific production rate ρ is:

$$\rho_H(t) = f_H(t) \cdot [\mu(t) - b] \tag{7}$$

We anticipate that the intersection $f_H(t) \cdot b$ will be present in other biosensors when the maximum heterologous protein fraction (φ_H^{max}) occurs at a positive specific growth rate ($\mu > 0$). In systems like the CymR-cumate system, where the specific growth rate at the zero specific production rate intersection is significant (comparable to the slope value), it is no longer true that the maximum φ_H^{max} equals f_H. Therefore, the regulatory parameters extracted from the dose-response plots will differ when using the φ_H^{max} method compared to the slope method. For

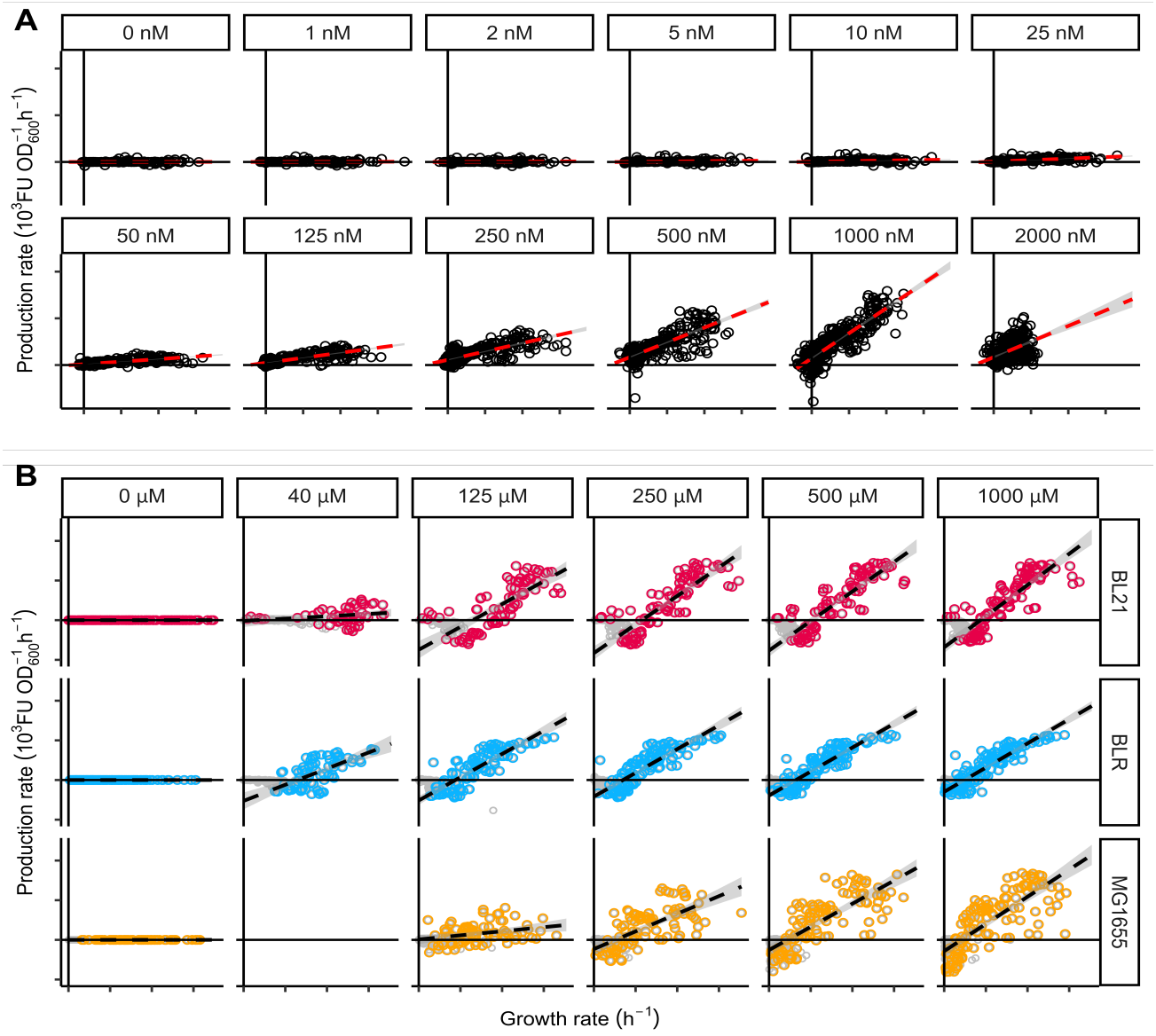


Figure 5. Expression fractions of biosensors at various inducer concentrations. A) Expression fractions of the MerR system in response to ionic mercury. Graphs show the RFP production upon induction with 0 to 2000 nM $HgBr_2$, as indicated, in NEB®Stable cells. Dark dots are experimental data fitted to a linear regression (dashed line) whereas open circles correspond to early time points that are excluded to maximize the correlation coefficient as indicated in the main text. B) Expression fractions at different cumate concentrations in the CymR-cumate system. Graphs show the GFP production upon induction with 0 to 1000 μM cumate, as indicated, in BL21(DE3), BLR(DE3) and MG1655(DE3) cultures transformed with the CymR-cumate system. Colored dots are experimental data fitted to a linear regression (dashed line) whereas grey dots correspond to early time points that are excluded to maximize R^2 as indicated in the main text. The expression fractions here determined were used to graph dose-response diagrams (Figure 1S and 2S in Supplementary Material) and were fitted to Hill functions (parameters in Table 1). ALT TEXT: MerR system and CymR system - Scatter plot of specific production rate vs specific growth rate. In each plot, a selected interval was used to make a linear regression.

illustrative purposes, Figure 2S presents Hill plots generated by both methods. However, the expression fractions obtained from the slope method should be considered reliable.

Discussion

This research tackles the challenge of standardising methods for characterising regulatory circuits and their modular components across various cellular contexts. The consistent behaviour of these parts is crucial for the long-term goal of designing predictable biological systems based on well-understood individual components. To this aim, we utilise conceptual tools from bacterial physiology to facilitate the characterisation of sensing circuits, particularly focussing on transcription factor-based biosensors.

The conceptualisation of the cell as a limited proteomic space introduced by Scott et al. [25] and expanded by subsequent kinetic models [32, 26] addresses the interdependence between the specific growth rate and the regulation of different sets of genes. Using the time derivative of the formal definition of a proteome fraction, the focus is shifted from protein yield to protein biosynthesis rate. Following the two alternative methods presented here, the experimental results readily produce the expression fraction, which represents a protein-specific

Sensor	Strain	$H(flu/OD_{600})$		$k_I(\mu M)$		n	
		Maxima	Slopes	Maxima	Slopes	Maxima	Slopes
IPTG	**BL21(DE3) + LacI^{OV}**	9784 ± 191	8457 ± 225	31.0 ± 3.7	35.6 ± 5.5	2.5 ± 0.3	2.5 ± 0.5
	BLR	6148 ± 496	6668 ± 651	55.8 ± 12.9	58.0 ± 14.9	2.2 ± 0.3	2.7 ± 0.5
	BLR + LacI^{OV}	11473 ± 319	11080 ± 563	51.7 ± 5.1	67.7 ± 6.9	2.8 ± 0.3	4.4 ± 1.8
	MG1655(DE3)	2072 ± 44	2846 ± 438	36.6 ± 4.1	56.3 ± 22.8	2.1 ± 0.2	2.0 ± 0.7
	MG1655(DE3) + LacI^{OV}	1116 ± 116	1747 ± 97	99.8 ± 4.5	93.6 ± 2.6	5.9 ± 1.4	6.6 ± 0.5
Cumate	**BL21(DE3)**	1170 ± 84	2965 ± 110	89.6 ± 13.4	79.7 ± 6.1	2.5 ± 0.2	4.3 ± 0.6
	BLR(DE3)	989 ± 45	2390 ± 33	50.8 ± 5.3	33.5 ± 1.8	2.3 ± 0.2	4.0 ± 0.2
	MG1655(DE3)	1539 ± 69	2575 ± 127	243 ± 17	208 ± 15.3	1.8 ± 0.2	3.2 ± 0.5
Hg^{2+}	**NEB® Stable**	-	11722 ± 3130	-	554 ± 281 (nM)	-	0.99 ± 0.1

Table 1. Hill function parameters describing the biosensors responses. The Hill function was fitted to the values of the expression fraction plotted as a function of inducer concentration (IPTG, cumate, or $HgBr_2$). The expression fractions were determined through two methods: Maxima, each fraction of expression is measured as equal to the heterologous protein fraction at its maximum value, $f_H = \varphi_H^{max}$. *Slopes*, each fraction of expression is determined by measuring the slope within the linear interval of the specific production rate vs specific growth rate plot, $f_H = \frac{\Delta\rho_H}{\Delta\mu}$.

production rate. By means of inducer titrations, the experimenters may obtain the expression fraction at various inducer concentrations, which are then summarised in a single set of Hill function parameters.

We first evaluated our approach by characterising the DE3 system in three distinct strains, considering both the presence and the absence of additional expression of the LacI repressor. Although the slope and maximum methods yield reasonably similar results, it is evident that these parameters are significantly influenced by the cellular context. For instance, the hooked shapes in the rate graphs in Figure 4 illustrate the time interval during which the expression fraction is not constant, presumably because the inducer has not yet reached equilibrium with the transcription factor. This lag period is shorter for higher concentrations of the inducer, as one would expect from a diffusion-dependent equilibration. Furthermore, it seems to be strain dependent and is associated with B strains exhibiting a higher apparent affinity for the inducer, as shown by a lower value of k_I than MG1655(DE3). These variations cannot be attributed to actual differences in the association reaction between the LacI protein and IPTG; therefore, these observations could be attributed to variations in IPTG permeability, considering the regulation of lactose permease (LacY) and the relative contribution of passive diffusion and active transport among these strains. A K-12 strain has been reported to depend on inducible LacY for the internalisation of IPTG [51, 52]. In agreement with this, when LacI overexpression was introduced in MG1655(DE3), k_I increased three or four times its value, which means that the apparent affinity of the system decreased, probably due to the reduced internalisation of IPTG caused by the decrease in LacY expression. This careful analysis of the rate plot and the Hill parameters in different contexts provides valuable observations specific to the cells where the system is tested. In future studies, we plan to evaluate and deepen into more biosensor cases, including the biosensors for mercury and cumate presented here, to see in what cases the Hill parameters are kept across various cells, and in what cases the variations are informative of strain-specific characteristics.

We evaluated the general applicability of the two analytical methods (maximum and slope) by analysing titration experiments from two additional biosensing systems. These systems detect cumate and mercury using CymR and MerR transcription factors, respectively. The maximal fractions method was not suitable for the MerR system because the reporter protein continuously accumulated, even after entering the stationary phase, preventing the detection of a clear maximum point φ_H^{max}. For the cumate biosensor, although the maximal fractions were readily identifiable, the presence of an intercept implies that φ_H^{max} does not equal the expression fraction. This discrepancy explains the inconsistencies between the Hill parameters obtained for the CymR system from the maximum and slope methods. Despite these limitations with the maximum method, the slope method remained a reliable approach for both the CymR and MerR systems. The extended linear regions observed in the rates plots indicated the suitability of these systems for analysis using the slope method. Furthermore, the brief non-linear phase suggests a faster equilibration with their specific inducers compared to the slower equilibration observed with IPTG.

In addition to analysing their applicability conditions, we also consider the limitations of each method that should be acknowledged for future uses. The maximum fraction method relies on a single data point, often obtained during the stationary phase. This approach does not provide information on the interval of time at which the expression fraction thus determined is valid. Furthermore, cellular functions are down-regulated during the stationary phase, potentially compromising the suitability of the method for assessing the biotechnological applications of a system despite a high apparent heterologous expression level. In the case of the slope method, it employs a linear regression in a time window of constant slope in the data, thereby overcoming this limitation of the maximum method. However, the ranges of linear correlations must be identified for each analysed culture condition, which involves a potential source of error. In this method, the selection of points for the regression analysis might be subjective or insufficient if dealing with a small number of data points. To address this limitation, we propose defining strict data requirements and implementing an algorithm for a more objective determination of the linear range. Other improvements will arise from better protein measurement accuracy. Here, we approximated the reporter proteome fraction, φ_H, simply normalising the fluorescence by OD_{600}. Better measurements can be made by complementary techniques, such as relating the registered fluorescence intensity to specific amounts of reporter protein through a calibration curve and MS-supported proteomic approaches.

In broad terms, this work presents methods and concepts designed to aid researchers in applying principles from advanced metabolic studies to characterise new regulatory elements and develop novel biosensor circuits. For example, a significant challenge in biocircuitry design is the lack of formal evaluation of orthogonality between heterologous gene expression and native functions [3, 53]. Existing methods often lack standardisation and rely on comprehensive and expensive techniques such as RNAseq and ribosome profiling [54, 55, 56, 57]. Alternatively, researchers employ specific, hypotheses-driven tests to identify interference between regulatory components [58, 59, 60, 61]. In contrast, the framework presented here offers a relatively simple approach to assess how the regulations of native functions, composed of expression fractions f_C (nutrient fixators) and f_R (biomass producers), are influenced by f_H (heterologous expression). For example, researchers can manipulate the carbon sources in the culture

to observe changes in f_C [25, 62, 32]. Alternatively, physical or biochemical stress mediated by ppGpp can be used to monitor changes in f_R [26, 63, 64]. These manipulations will activate distinct cellular regulatory mechanisms and, if a novel heterologous circuit affects these mechanisms, a dose-response relationship should emerge, with increased heterologous expression correlating with heightened interference. This approach could reveal interactions between native metabolic regulations or stress responses and the heterologous gene(s). In this way, the concepts of proteome fractions and expression fractions offer a more nuanced approach to evaluate reporter protein production compared to simple normalisation by culture density or growth rate. These methods remain cost-effective and convenient, potentially encouraging wider adoption of a systems-wide approach in the characterisation of biomolecular parts for synthetic biology and improving reproducibility across laboratories.

Author information

Corresponding author

Daniel G. Guerra - Laboratorio de Moléculas Individuales, Laboratorios de Investigación y Desarrollo, Facultad de Ciencias e Ingeniería, Universidad Peruana Cayetano Heredia, Lima 15102, Perú. https://orcid.org/0000-0002-2410-722X, daniel.guerra@upch.pe, +51 319-0000 ext. 233217.

Author contribution

Nicolás A. Vaccari Conceptualised and developed the model, designed and cloned the IPTG biosensor system, performed titration experiments, performed data analysis and wrote the article. **Dahlin Zevallos-Aliaga** - Designed and cloned the biosensors for mercury and cumate and performed their titration assays. **Tom Peeters** - Provided critical feedback on the design and testing of the mercury biosensor. **Daniel G. Guerra** - Served as project supervisor, providing guidance and critical feedback for experimental design, modelling, and data analysis, and wrote the article.

Notes

The authors declare that they have no competing financial interests.

Material availability statement

The sequences of the genetic circuits designed for this study are available on GenBank: IPTG biosensor (GenBank: PQ015608), and the cumate biosensor (GenBank: PQ010742). The sequence of the mercury biosensor is described in [37]. All plasmids used in this study will be shared upon request to the corresponding author.

Data availability statement

Datasets available on FigShare: https://doi.org/10.6084/m9.figshare.26314723.v1

Acknowledgements

This work was supported by CONCYTEC – PROCIENCIA - Contrato 170-2020. The IPTG and cumate biosensors were developed and characterised thanks to the support of Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica – FONDECYT / PROCIENCIA, through the program EF-041 - Proyectos de Investigación básica y aplicada. Grant number CONV-000114-2015-FONDECYT-DE. The development of the mercury biosensor was funded by VLIR-UOS through the South Initiative grant code PE2020SIN292B122.

Supplementary information

Supporting Experimental Information: Additional experimental results showing dose-response diagrams for the MerR and CymR systems, total protein analysis by SDS-PAGE for all strains containing the DE3 system expressing GFP, and total protein quantification for induced and uninduced bacteria. Supporting Mathematical Framework: Detailed definitions, step-by-step derivations, and algebraic arrangements.

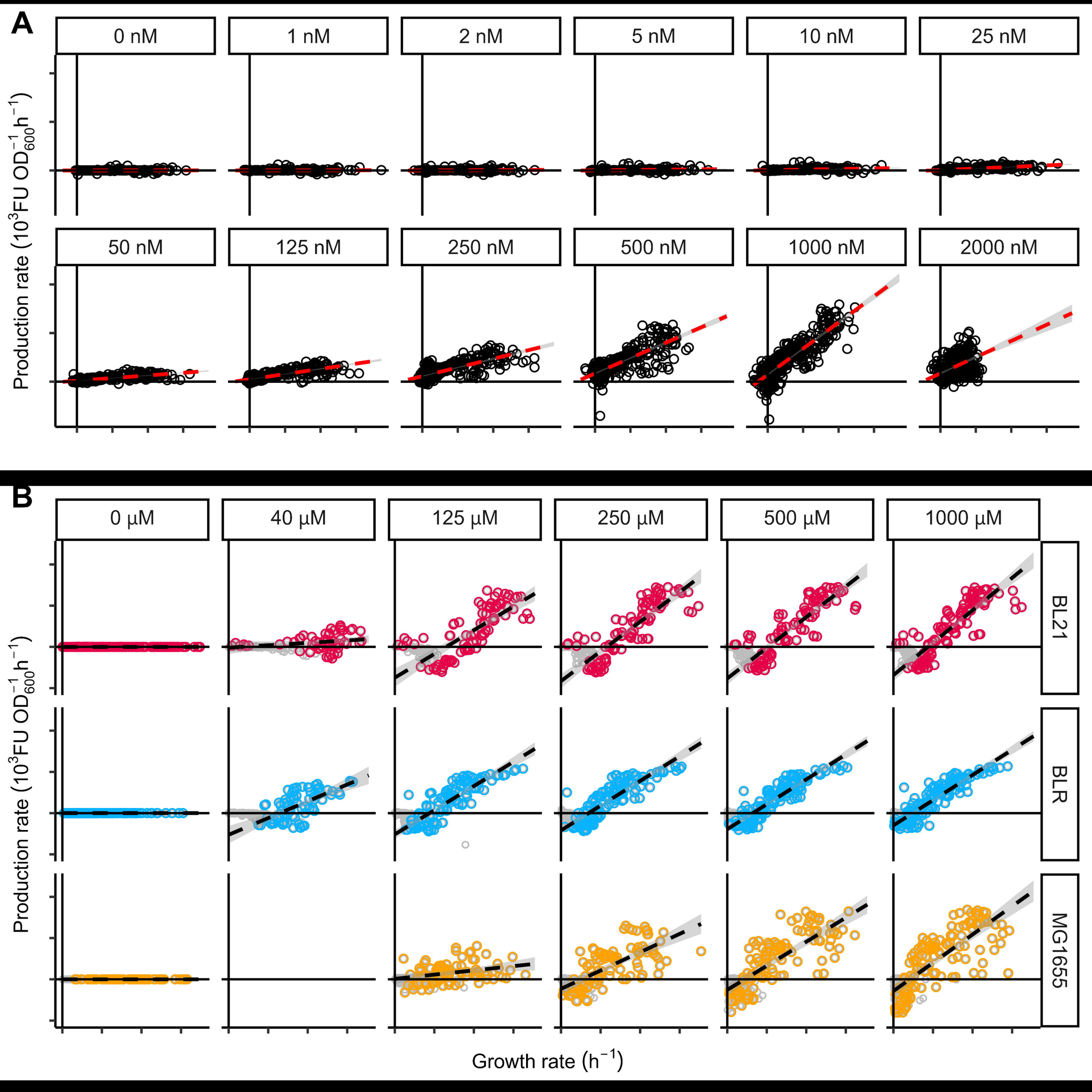
References

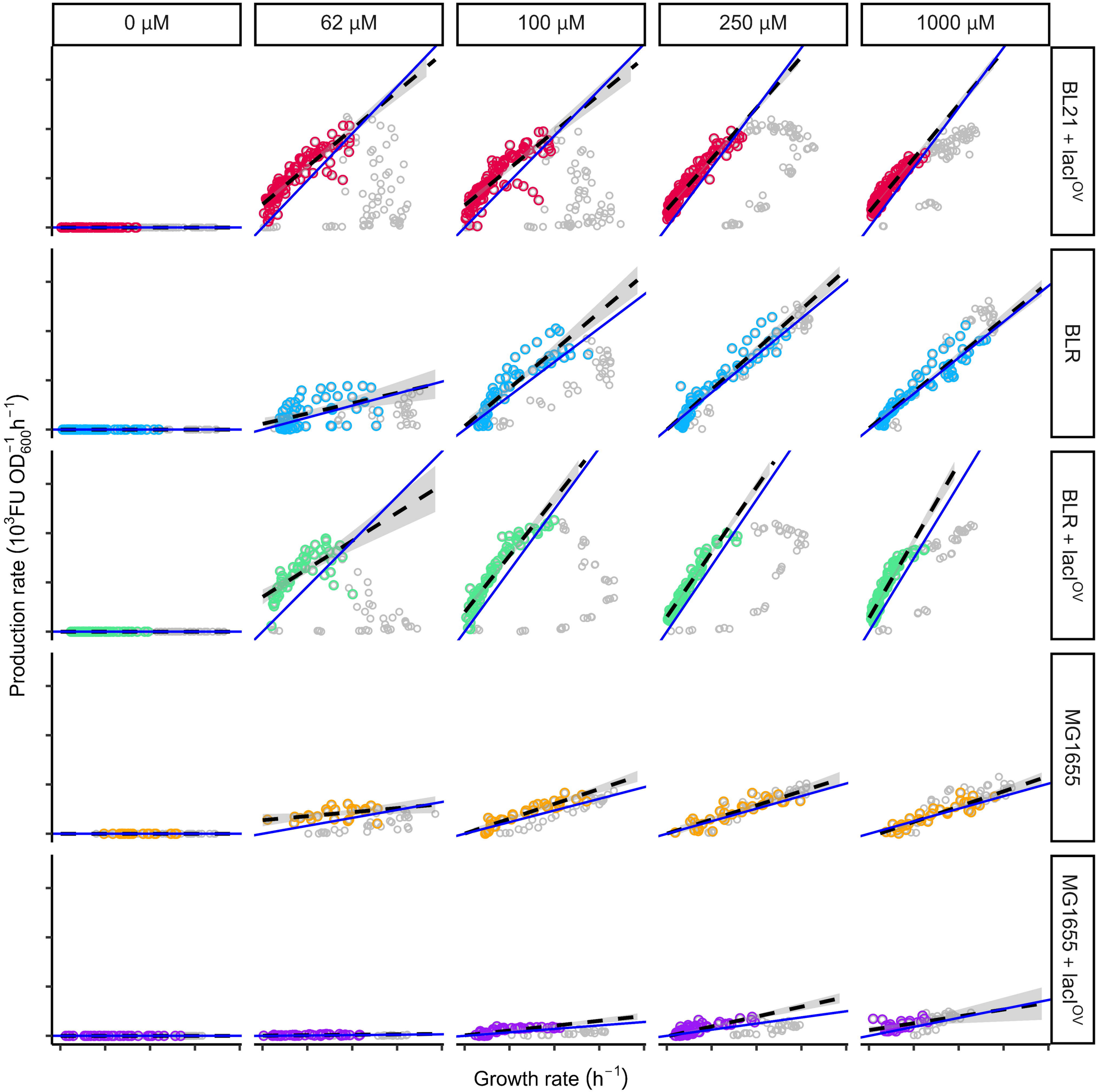
1. Steven A. Benner and A. Michael Sismour. Synthetic biology. *Nat Rev Genet*, 6(7):533–543, July 2005. Publisher: Nature Publishing Group.
2. Barry Canton, Anna Labno, and Drew Endy. Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol*, 26(7):787–793, July 2008.
3. Roberta Kwok. Five hard truths for synthetic biology. *Nature*, 463(7279):288–290, 2010.
4. Stefano Cardinale and Adam Paul Arkin. Contextualizing context for synthetic biology â€“ identifying causes of failure of synthetic biological systems. *Biotechnol J*, 7(7):856–866, July 2012.
5. Nymul Khan, Enoch Yeung, Yuliya Farris, Sarah J Fansler, and Hans C Bernstein. A broad-host-range event detector: expanding and quantifying performance between Escherichia coli and Pseudomonas species. *Synthetic Biology*, 5(1):ysaa002, January 2020.
6. Dennis Tin Chat Chan, Geoff S. Baldwin, and Hans C. Bernstein. Revealing the Host-Dependent Nature of an Engineered Genetic Inverter in Concordance with Physiology. *Biodes Res*, 5:0016, 2023.
7. Vincent Libis, Baudoin Delépine, and Jean-Loup Faulon. Sensing new chemicals with bacterial transcription factors. *Current Opinion in Microbiology*, 33:105–112, 2016.
8. Alyssa J. Chiang and Jeff Hasty. Design of synthetic bacterial biosensors. *Current Opinion in Microbiology*, 76:102380, 2023.
9. Lisa Tietze and Rahmi Lale. Importance of the 5' regulatory region to bacterial synthetic biology applications. *Microbial Biotechnology*, 14(6):2291–2315, 2021.
10. Jameson K Rogers, Noah D Taylor, and George M Church. Biosensor-based engineering of biosynthetic pathways. *Current Opinion in Biotechnology*, 42:84–91, December 2016.
11. Patarasuda Chaisupa and R. Clay Wright. State-of-the-art in engineering small molecule biosensors and their applications in metabolic engineering. *SLAS Technology*, 29(2):100113, April 2024.
12. Yuxi Teng, Jianli Zhang, Tian Jiang, Yusong Zou, Xinyu Gong, and Yajun Yan. Biosensor-enabled pathway optimization in metabolic engineering. *Current Opinion in Biotechnology*, 75:102696, 2022.
13. Melissa M Mitchler, Jessie M Garcia, Nichole E Montero, and Gavin J Williams. Transcription factor-based biosensors: a molecular-guided approach for natural product engineering. *Current Opinion in Biotechnology*, 69:172–181, 2021.
14. Shijing Chen, Xiaolin Chen, Hongfei Su, Mingzhang Guo, and Huilin Liu. Advances in synthetic-biology-based whole-cell biosensors: Principles, genetic modules, and applications in food safety. *International Journal of Molecular Sciences*, 24(9):7989, 2023.
15. Yueqian Chen, Yingying Guo, Yanwei Liu, Yuping Xiang, Guangliang Liu, Qinghua Zhang, Yongguang Yin, Yong Cai, and Guibin Jiang. Advances in bacterial whole-cell biosensors for the detection of bioavailable mercury: A review. *The Science of the Total*

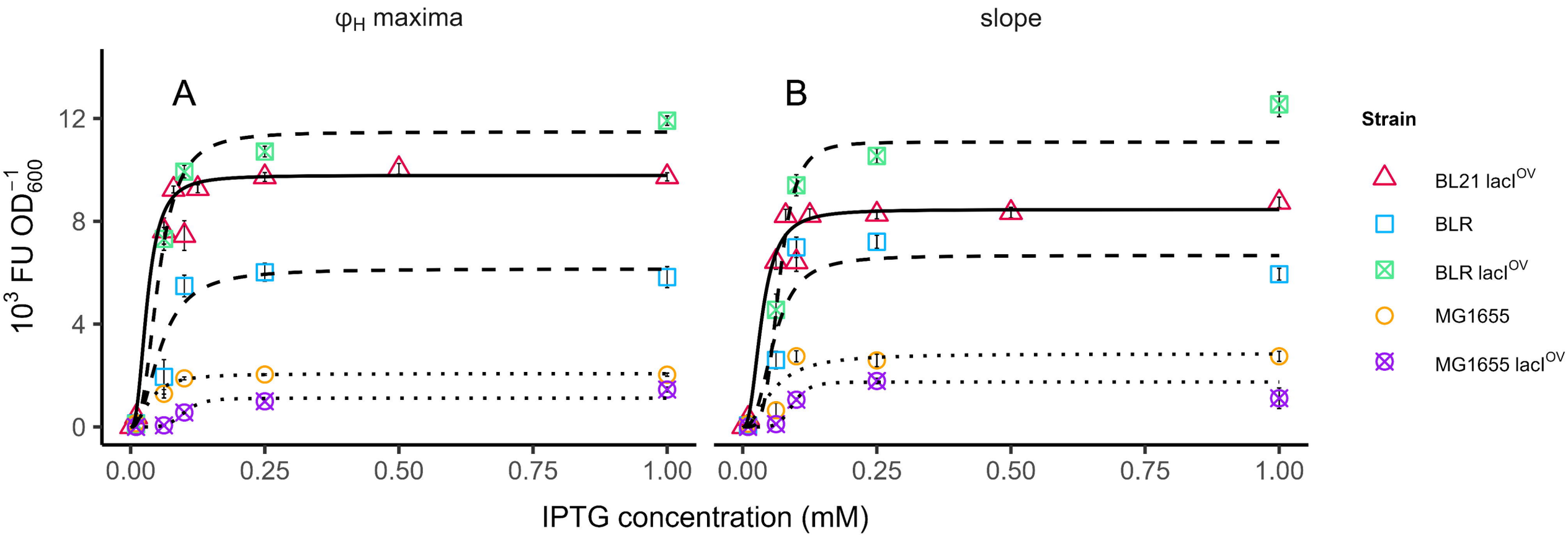
Environment, 868:161709, 2023.

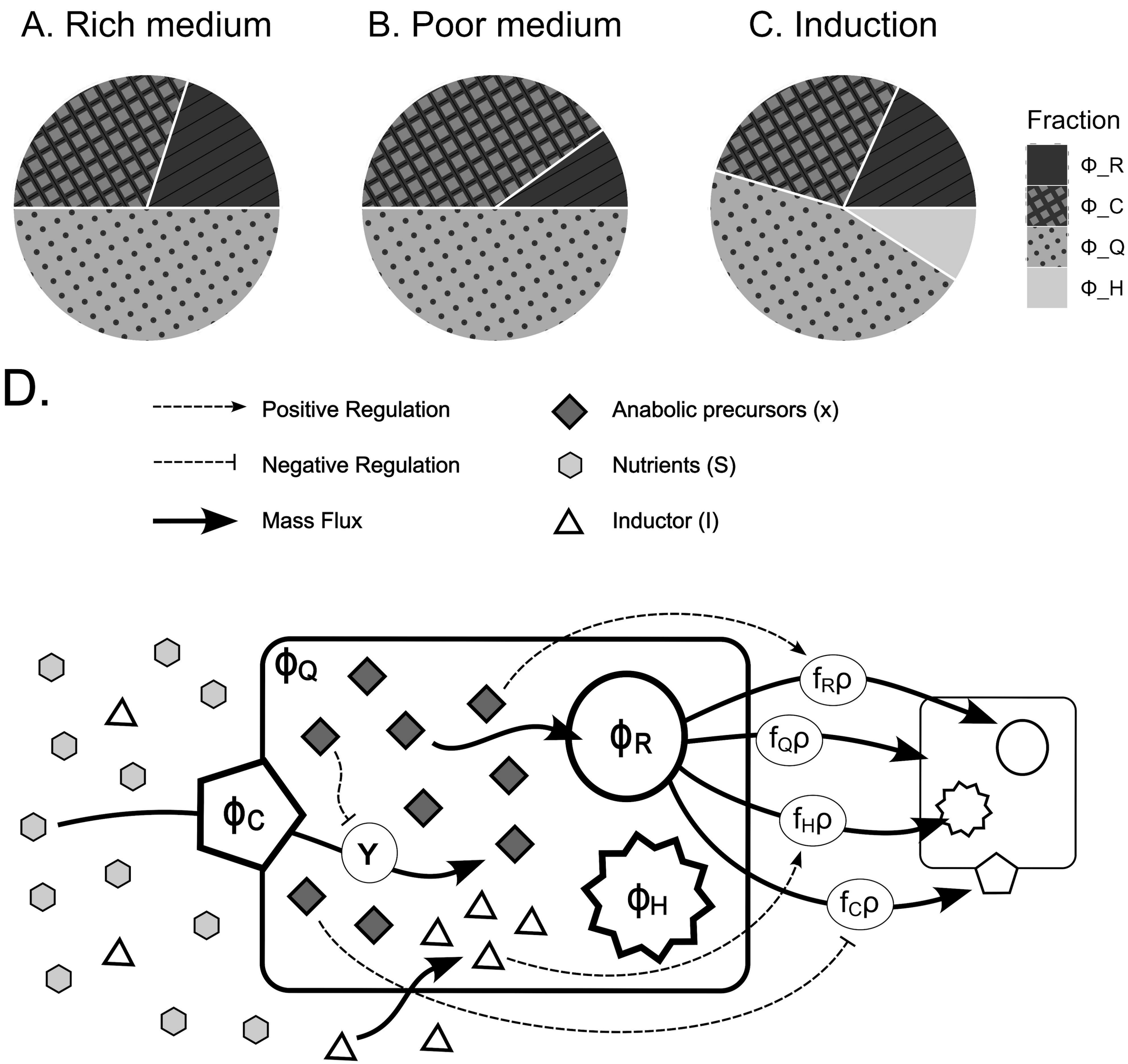
16. Qingyuan Gui, Tom Lawson, Suyan Shan, Lu Yan, and Yong Liu. The application of whole cell-based biosensors for use in environmental analysis and in medical diagnostics. *Sensors (Basel, Switzerland)*, 17(7):1623, 2017.
17. Baojun Wang and Cheemeng Tan. Special issue "synthetic biology for biosensing in health and environmental applications". *Biosensors*, 13(10):937, 2023.
18. Di Liu and Fuzhong Zhang. Metabolic feedback circuits provide rapid control of metabolite dynamics. *ACS synthetic biology*, 7(2):347–356, 2018.
19. Jonathan Tellechea-Luzardo, Martin T. Stiebritz, and Pablo Carbonell. Transcription factor-based biosensors for screening and dynamic regulation. *Frontiers in Bioengineering and Biotechnology*, 11:1118702, 2023.
20. Regina Mahr and Julia Frunzke. Transcription factor-based biosensors in biotechnology: current state and future prospects. *Applied Microbiology and Biotechnology*, 100(1):79–90, 2016.
21. Adam J. Meyer, Thomas H. Segall-Shapiro, Emerson Glassey, Jing Zhang, and Christopher A. Voigt. Escherichia coli "marionette" strains with 12 highly optimized small-molecule sensors. *Nature Chemical Biology*, 15(2):196–204, 2019.
22. Wenwen Yu, Xianhao Xu, Ke Jin, Yanfeng Liu, Jianghua Li, Guocheng Du, Xueqin Lv, and Long Liu. Genetically encoded biosensors for microbial synthetic biology: From conceptual frameworks to practical applications. *Biotechnology Advances*, 62:108077, January 2023.
23. A. Vignoni, Y. Boada, L. Boada Acosta, C. Andreu-Vilarroig, I. Alarcón, A. Requena, and J. Picó. Fluorescence calibration and color equivalence for quantitative synthetic biology. *IFAC-PapersOnLine*, 52(26):129–134, 2019.
24. Gonzalo Vidal, Carlos Vidal-Céspedes, Macarena Muñoz Silva, Carlos Castillo-Passi, Guillermo Yáñez Feliú, Fernán Federici, and Timothy J. Rudge. Accurate characterization of dynamic microbial gene expression and growth rate profiles. *Synth Biol (Oxf)*, 7(1):ysac020, 2022.
25. Matthew Scott, Carl W. Gunderson, Eduard M. Mateescu, Zhongge Zhang, and Terence Hwa. Interdependence of cell growth and gene expression: origins and consequences. *Science (New York, N.Y.)*, 330(6007):1099–1102, 2010.
26. David W. Erickson, Severin J. Schink, Vadim Patsalo, James R. Williamson, Ulrich Gerland, and Terence Hwa. A global resource allocation strategy governs growth transition kinetics of escherichia coli. *Nature*, 551(7678):119–123, 2017. Number: 7678.
27. Pooja Sharma, Parth Pratim Pandey, and Sanjay Jain. Modeling the cost and benefit of proteome regulation in a growing bacterial cell. *Physical Biology*, 15(4):046005, 2018. Number: 4.
28. Francis Mairet, Jean-Luc Gouzé, and Hidde De Jong. Optimal proteome allocation and the temperature dependence of microbial growth laws. *npj Systems Biology and Applications*, 7(1):14, 2021. Number: 1.
29. Matteo Mori, Chuankai Cheng, Brian R. Taylor, Hiroyuki Okano, and Terence Hwa. Functional decomposition of metabolism allows a system-level quantification of fluxes and protein allocation towards specific metabolic functions. *Nature Communications*, 14(1):4161, 2023. Number: 1.
30. Matthew Scott and Terence Hwa. Shaping bacterial gene expression by physiological and proteome allocation constraints. *Nature Reviews Microbiology*, 21(5):327–342, 2023. Number: 5.
31. Chenhao Wu, Matteo Mori, Miriam Abele, Amir Banaei-Esfahani, Zhongge Zhang, Hiroyuki Okano, Ruedi Aebersold, Christina Ludwig, and Terence Hwa. Enzyme expression kinetics by escherichia coli during transition from rich to minimal media depends on proteome reserves. *Nature Microbiology*, 8(2):347–359, 2023. Number: 2.
32. Benjamin D. Towbin, Yael Korem, Anat Bren, Shany Doron, Rotem Sorek, and Uri Alon. Optimality and sub-optimality in a bacterial growth law. *Nature Communications*, 8:14123, 2017.
33. Matthew W Lux, Elizabeth A Strychalski, and Gary J Vora. Advancing reproducibility can ease the 'hard truths' of synthetic biology. *Synth Biol (Oxf)*, 8(1):ysad014, October 2023.
34. Hsien-Chung Tseng, Catey L. Harwell, Collin H. Martin, and Kristala L. J. Prather. Biosynthesis of chiral 3-hydroxyvalerate from single propionate-unrelated carbon sources in metabolically engineered e. coli. *Microbial Cell Factories*, 9:96, 2010.
35. Enrique Balleza, J. Mark Kim, and Philippe Cluzel. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature Methods*, 15(1):47–51, 2018.
36. Jens Bo Andersen, Claus Sternberg, Lars Kongsbak Poulsen, Sara Petersen Bjørn, Michael Givskov, and Søren Molin. New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Applied and Environmental Microbiology*, 64(6):2240–2246, 1998.
37. Dahlin Zevallos-Aliaga, Stijn De Graeve, Pamela Obando-ChÃ¡vez, NicolÃ¡s A. Vaccari, Yue Gao, Tom Peeters, and Daniel G. Guerra. Highly Sensitive Whole-Cell Mercury Biosensors for Environmental Monitoring. *Biosensors*, 14(5):246, May 2024. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.
38. Young J. Choi, Lyne Morel, Teffanie Le François, Denis Bourque, Lucie Bourget, Denis Groleau, Bernard Massie, and Carlos B. Míguez. Novel, versatile, and tightly regulated expression system for Escherichia coli strains. *Applied and Environmental Microbiology*, 76(15):5058–5066, August 2010.
39. Nicolas A. Vaccari, Dahlin Zevallos, Tom Peeters, and Daniel G. Guerra. biosensors_de3_system.csv.
40. F. William Studier and Barbara A. Moffatt. Use of bacteriophage t7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, 189(1):113–130, 1986.
41. Kay Terpe. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, 72(2):211–222, 2006.
42. Ahmad A. Mannan, Di Liu, Fuzhong Zhang, and Diego A. Oyarzún. Fundamental design principles for transcription-factor-based metabolite biosensors. *ACS synthetic biology*, 6(10):1851–1859, 2017.
43. Baojun Wang, Richard I. Kitney, Nicolas Joly, and Martin Buck. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nature Communications*, 2(1):508, October 2011. Number: 1 Publisher: Nature Publishing Group.
44. Wenya Wang, Yuwenbin Li, Yaqiong Wang, Chen Shi, Chenmeng Li, Qiang Li, and Robert J. Linhardt. Bacteriophage T7 transcription system: an enabling tool in synthetic biology. *Biotechnology Advances*, 36(8):2129–2137, December 2018.
45. Drew M. DeLorenzo and Tae Seok Moon. Construction of Genetic Logic Gates Based on the T7 RNA Polymerase Expression System in Rhodococcus opacus PD630. *ACS synthetic biology*, 8(8):1921–1930, August 2019.
46. Olivier Borkowski, Francesca Ceroni, Guy-Bart Stan, and Tom Ellis. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current Opinion in Microbiology*, 33:123–130, 2016.
47. Frank Hoffmann and Ursula Rinas. Stress induced by recombinant protein production in escherichia coli. In *Physiological Stress Responses in Bioprocesses*, volume 89, pages 73–92. Springer Berlin Heidelberg, 2004. Series Title: Advances in Biochemical Engineering/Biotechnology.

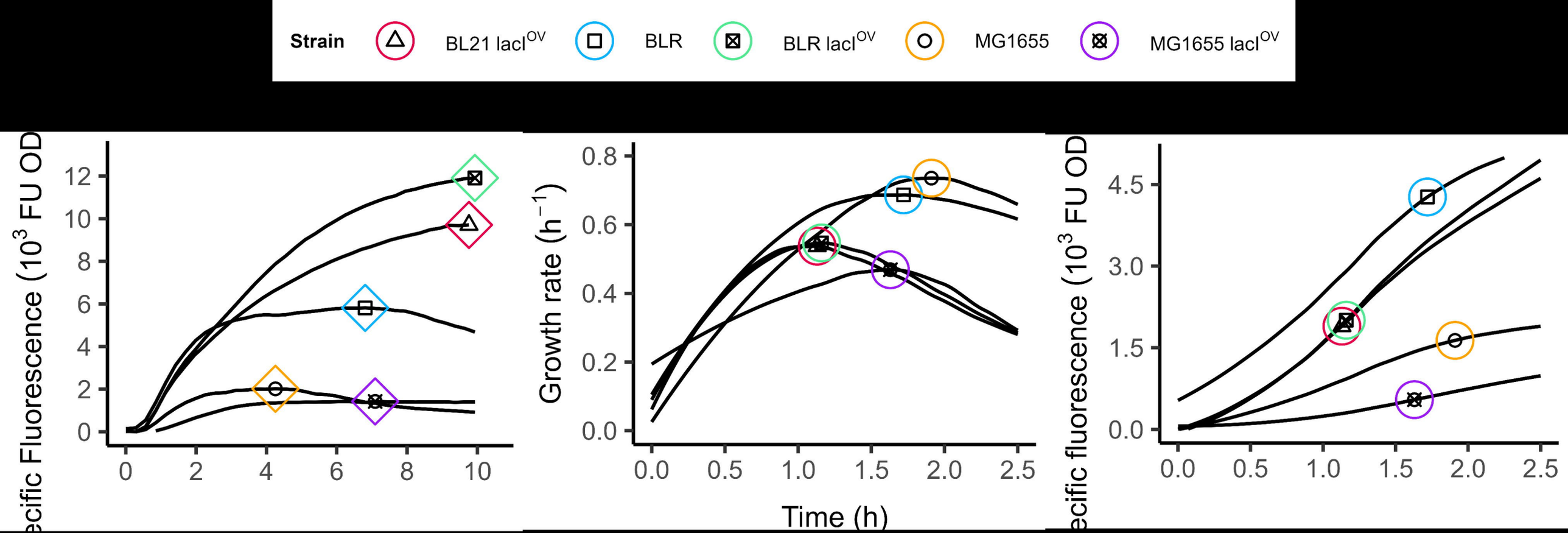
48. Girish H. Rajacharya, Ashima Sharma, and Syed Shams Yazdani. Proteomics and metabolic burden analysis to understand the impact of recombinant protein production in e. coli. *Scientific Reports*, 14(1):12271, 2024.
49. Uri Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall/CRC, 2006.
50. P. P. Dennis and H. Bremer. Macromolecular composition during steady-state growth of escherichia coli b-r. *Journal of Bacteriology*, 119(1):270–281, 1974.
51. Alfred FernÃ¡ndez-CastanÃ©, GlÃ²ria Caminal, and Josep LÃ³pez-SantÃn. Direct measurements of IPTG enable analysis of the induction behavior of E. coli in high cell density cultures. *Microbial Cell Factories*, 11:58, May 2012.
52. Alfred FernÃ¡ndez-CastanÃ©, Claire E. Vine, GlÃ²ria Caminal, and Josep LÃ³pez-SantÃn. Evidencing the role of lactose permease in IPTG uptake by *Escherichia coli* in fed-batch high cell density cultures. *Journal of Biotechnology*, 157(3):391–398, February 2012.
53. Victor de Lorenzo. Beware of metaphors: Chasses and orthogonality in synthetic biology. *Bioengineered Bugs*, 2(1):3–7, 2011. PMID: 21636982.
54. Thomas E Gorochowski, Amin Espah Borujeni, Yongjin Park, Alec Ak Nielsen, Jing Zhang, Bryan S Der, D Benjamin Gordon, and Christopher A Voigt. Genetic circuit characterization and debugging using RNA -seq. *Molecular Systems Biology*, 13(11):952, November 2017.
55. Qijun Liu, Jörg Schumacher, Xinyi Wan, Chunbo Lou, and Baojun Wang. Orthogonality and burdens of heterologous AND gate gene circuits in e. coli. *ACS Synthetic Biology*, 7(2):553–564, 2018. Publisher: American Chemical Society.
56. Amin Espah Borujeni, Jing Zhang, Hamid Doosthosseini, Alec A. K. Nielsen, and Christopher A. Voigt. Genetic circuit characterization by inferring RNA polymerase movement and ribosome usage. *Nat Commun*, 11(1):5001, October 2020.
57. Deepti Vipin, Zoya Ignatova, and Thomas E. Gorochowski. Characterizing Genetic Parts and Devices Using RNA Sequencing. *Methods in Molecular Biology*, 2229:175–187, 2021. Book Title: Synthetic Gene Circuits ISBN: 9781071610312 9781071610329 Place: New York, NY Publisher: Springer US.
58. Alec AK Nielsen, Thomas H Segall-Shapiro, and Christopher A Voigt. Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Current Opinion in Chemical Biology*, 17(6):878–892, 2013. Synthetic biology • Synthetic biomolecules.
59. Virgil A Rhodius, Thomas H Segall-Shapiro, Brian D Sharon, Amar Ghodasara, Ekaterina Orlova, Hannah Tabakh, David H Burkhardt, Kevin Clancy, Todd C Peterson, Carol A Gross, and Christopher A Voigt. Design of orthogonal genetic switches based on a crosstalk map of sigma's, anti-sigma's, and promoters. *Molecular Systems Biology*, 9(1):702, 2013.
60. Alan Costello and Ahmed H. Badran. Synthetic biological circuits within an orthogonal central dogma. *Trends in biotechnology*, 39(1):59–71, 2021.
61. Indra Bervoets, Maarten Van Brempt, Katleen Van Nerom, Bob Van Hove, Jo Maertens, Marjan De Mey, and Daniel Charlier. A sigma factor toolbox for orthogonal gene expression in escherichia coli. *Nucleic Acids Research*, 46(4):2133–2144, 2018.
62. Conghui You, Hiroyuki Okano, Sheng Hui, Zhongge Zhang, Minsu Kim, Carl W. Gunderson, Yi-Ping Wang, Peter Lenz, Dalai Yan, and Terence Hwa. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500(7462):301–306, 2013. Number: 7462 Publisher: Nature Publishing Group.
63. Chenhao Wu, Rohan Balakrishnan, Nathan Braniff, Matteo Mori, Gabriel Manzanarez, Zhongge Zhang, and Terence Hwa. Cellular perception of growth rate and the mechanistic origin of bacterial growth law. *Proceedings of the National Academy of Sciences*, 119(20):e2201585119, May 2022. Publisher: Proceedings of the National Academy of Sciences.
64. Francesca Ceroni, Alice Boo, Simone Furini, Thomas E. Gorochowski, Olivier Borkowski, Yaseen N. Ladak, Ali R. Awan, Charlie Gilbert, Guy-Bart Stan, and Tom Ellis. Burden-driven feedback control of gene expression. *Nature Methods*, 15(5):387–393, 2018.

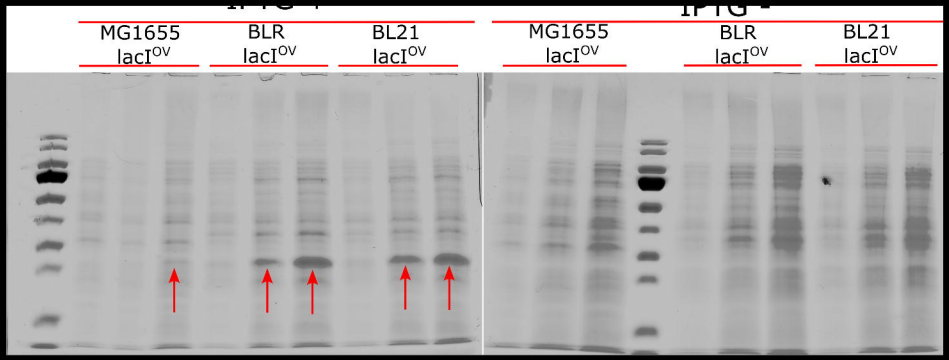

A
0 nM
1 nM
2 nM
5 nM
10 nM
25 nM
50 nM
125 nM
250 nM
500 nM
1000 nM
2000 nM
Production rate (10^3FU $OD_{600}^{-1}h^{-1}$)
B
0 μM
40 μM
125 μM
250 μM
500 μM
1000 μM
BL21
BLR
MG1655
Production rate (10^3FU $OD_{600}^{-1}h^{-1}$)
Growth rate (h^{-1})

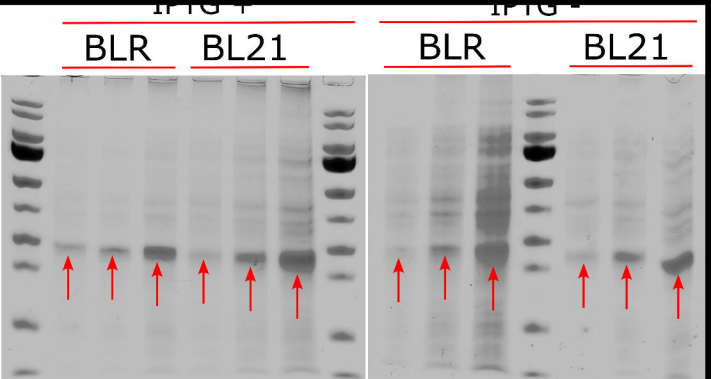

0 μM
62 μM
100 μM
250 μM
1000 μM
BL21 + lacI^{OV}
BLR
BLR + lacI^{OV}
MG1655
MG1655 + lacI^{OV}
Production rate (10^3FU $OD_{600}^{-1}h^{-1}$)
Growth rate (h^{-1})

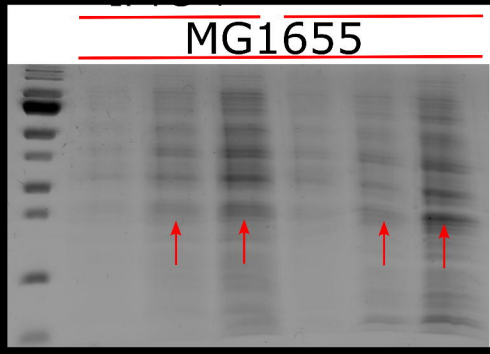

φH maxima
slope
A
B
12
8
4
0
10^3 FU OD^-1_600
0.00
0.25
0.50
0.75
1.00
IPTG concentration (mM)
Strain
BL21 lacI^OV
BLR
BLR lacI^OV
MG1655
MG1655 lacI^OV

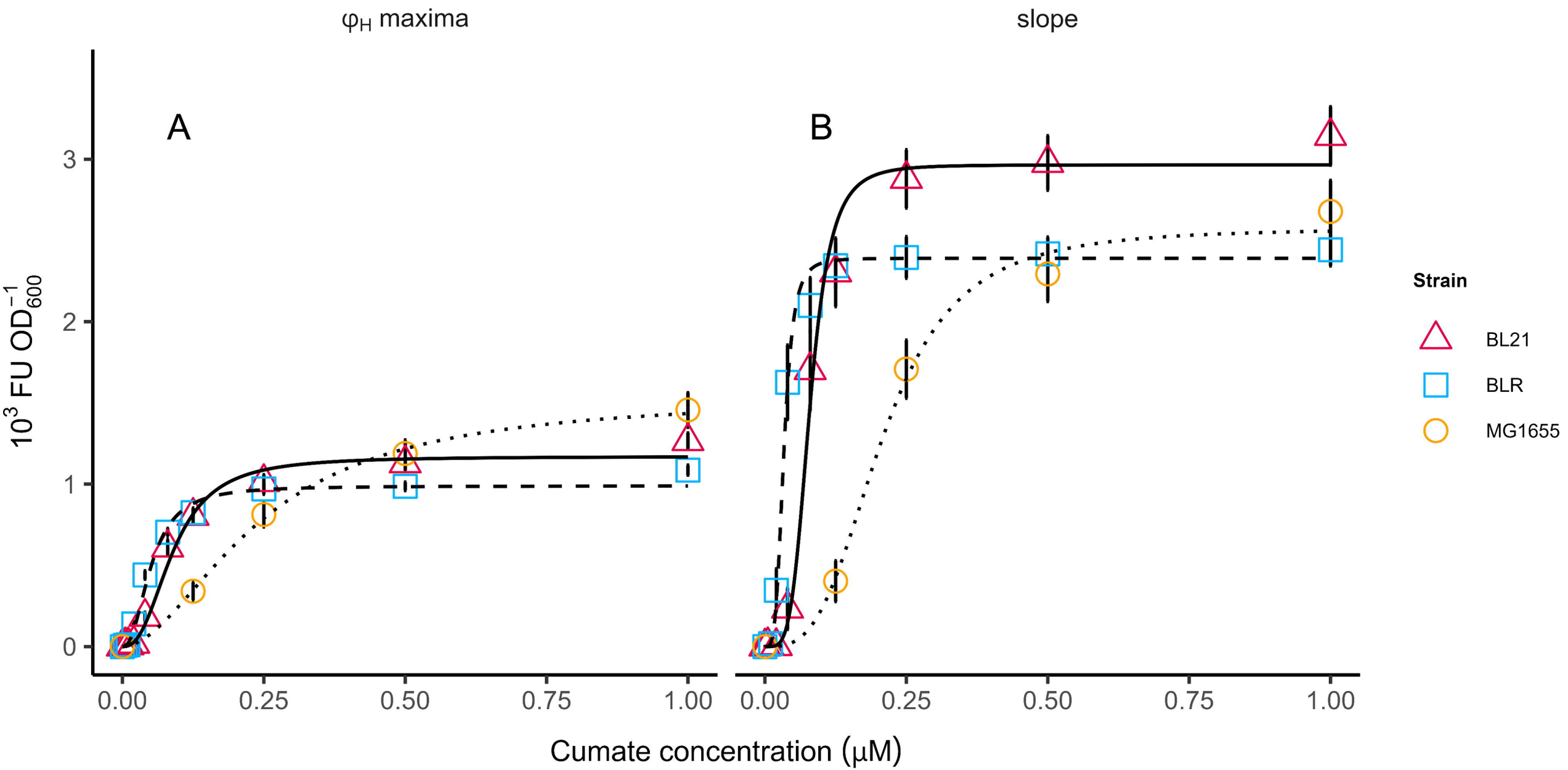

A. Rich medium
B. Poor medium
C. Induction
Fraction
Φ_R
Φ_C
Φ_Q
Φ_H
D.
Positive Regulation
Negative Regulation
Mass Flux
Anabolic precursors (x)
Nutrients (S)
Inductor (I)
ϕQ
ϕR
ϕC
ϕH
Y
fRρ
fQρ
fHρ
fCρ

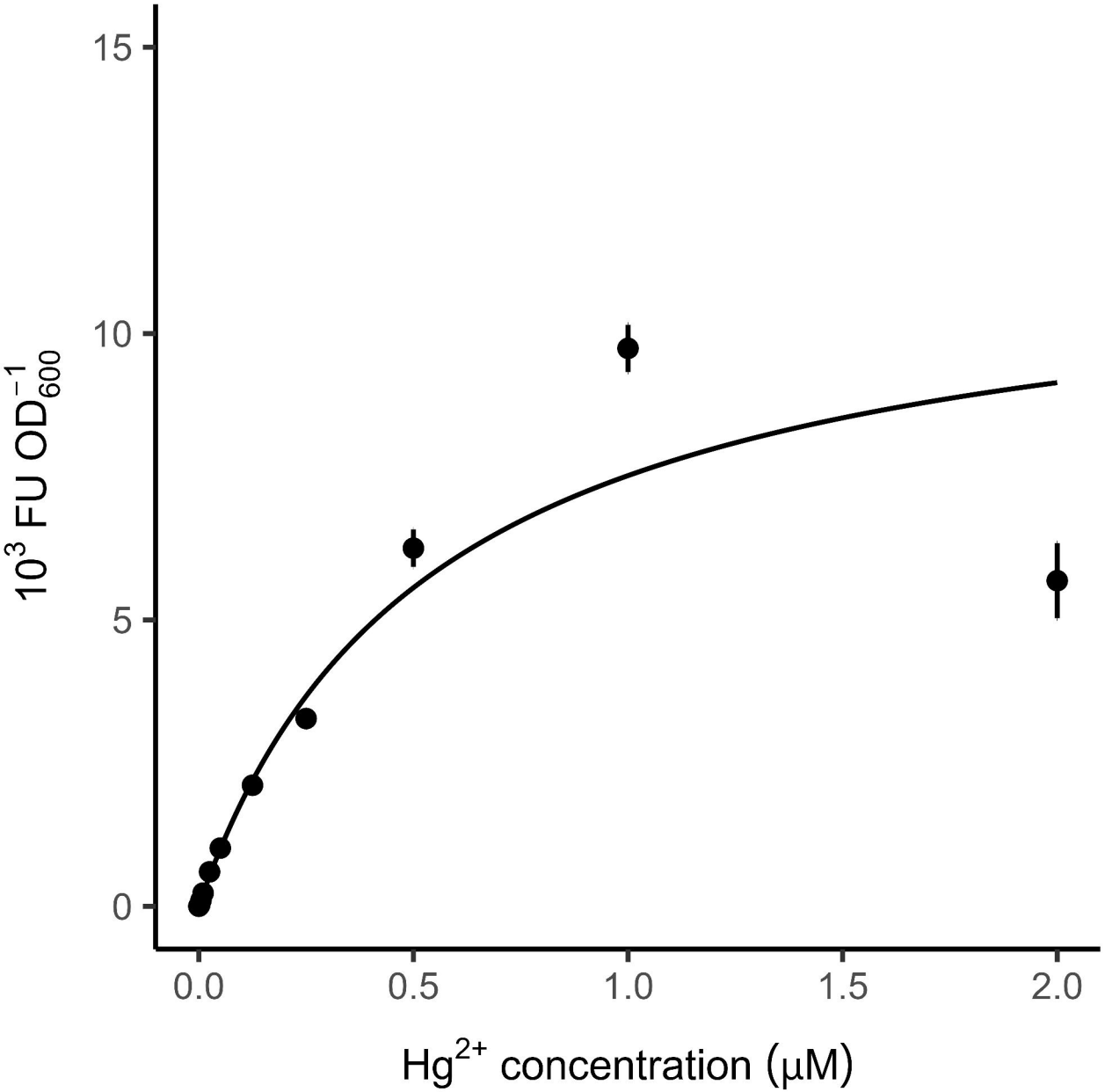

Strain
BL21 lacI^OV
BLR
BLR lacI^OV
MG1655
MG1655 lacI^OV
ecific Fluorescence (10^3 FU OD
0
2
4
6
8
10
12
Growth rate (h^−1)
0.0
0.2
0.4
0.6
0.8
0.5
1.0
1.5
2.0
2.5
Time (h)
ecific fluorescence (10^3 FU OD
0.0
1.5
3.0
4.5


MG1655
lacI^OV
BLR
lacI^OV
BL21
lacI^OV
MG1655
lacI^OV
BLR
lacI^OV
BL21
lacI^OV


BLR
BL21
BLR
BL21


MG1655


φH maxima
slope
A
B
3
2
1
0
10^3 FU OD^-1_600
0.00
0.25
0.50
0.75
1.00
Cumate concentration (μM)
Strain
BL21
BLR
MG1655


15
10
5
0
10^3 FU OD_{600}^{-1}
0.0
0.5
1.0
1.5
2.0
Hg^{2+} concentration (μM)