

# A coarse-grained bacterial cell model for resource-aware analysis and design of synthetic gene circuits

---

Received: 15 May 2023

---

Accepted: 27 February 2024

---

Published online: 04 March 2024

---

 Check for updates

---

Kirill Sechkar <sup>1</sup>, Harrison Steel <sup>1</sup>, Giansimone Perrino <sup>2,3</sup>  & Guy-Bart Stan <sup>2,3</sup> 

Within a cell, synthetic and native genes compete for expression machinery, influencing cellular process dynamics through resource couplings. Models that simplify competitive resource binding kinetics can guide the design of strategies for countering these couplings. However, in bacteria resource availability and cell growth rate are interlinked, which complicates resource-aware biocircuit design. Capturing this interdependence requires coarse-grained bacterial cell models that balance accurate representation of metabolic regulation against simplicity and interpretability. We propose a coarse-grained *E. coli* cell model that combines the ease of simplified resource coupling analysis with appreciation of bacterial growth regulation mechanisms and the processes relevant for biocircuit design. Reliably capturing known growth phenomena, it provides a unifying explanation to disparate empirical relations between growth and synthetic gene expression. Considering a biomolecular controller that makes cell-wide ribosome availability robust to perturbations, we showcase our model's usefulness in numerically prototyping biocircuits and deriving analytical relations for design guidance.

The engineering of novel biological systems with useful applications, known as synthetic biology, promises to address global challenges by revolutionizing healthcare, agriculture, and manufacturing<sup>1</sup>. While any engineered system's behavior must be predictable, its forecasting for synthetic biology designs is complicated by the interconnectedness of biological processes in living cells<sup>2–4</sup>. Indeed, even the gene circuit components that do not engage in direct biological interactions may exhibit interdependence caused by indirect couplings via shared cellular resources. Specifically, resource competition couplings arise when heterologous genes introduced into a host cell compete among themselves and with native genes for the same finite pools of resources that enable gene expression. Due to this competition, a synthetic gene circuit's performance may not be reliably predicted based on the characterization of its constituent modules in isolation<sup>5,6</sup>. Moreover, the redirection of resources from native gene expression, known as

gene expression burden, hinders cell growth and biomass accumulation<sup>3,4</sup>. Resource couplings therefore present a key challenge in the engineering of bacterial, fungal, and mammalian cells<sup>7–9</sup>.

To mitigate resource couplings, synthetic gene expression can be kept low enough to have no significant effect on cellular homeostasis<sup>10</sup>. Alternatively, feedback loops can improve a circuit's robustness to perturbations such as resource couplings<sup>7</sup>. Moreover, synthetic genes can be expressed via orthogonal molecular machinery which is unused in native gene expression and forms a separate resource pool, thereby reducing crosstalk between engineered circuits and native processes<sup>11</sup>. Nonetheless, the low-expression strategy is unsuitable for applications requiring high protein production, while the synthesis of orthogonal machinery or regulator proteins enabling feedback control can itself burden the cells<sup>10</sup>. Mathematical modeling of resource competition represents a promising approach to resource-aware design of

---

<sup>1</sup>Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK. <sup>2</sup>Department of Bioengineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK. <sup>3</sup>Imperial College Centre of Excellence in Synthetic Biology, Imperial College London, South Kensington Campus, London SW7 2AZ, UK. ✉ e-mail: [g.perrino@imperial.ac.uk](mailto:g.perrino@imperial.ac.uk); [g.stan@imperial.ac.uk](mailto:g.stan@imperial.ac.uk)

biocircuits and the development of more sophisticated strategies for countering resource couplings.

One can explicitly model all stages of a substrate binding and unbinding the shared resource for which it competes with other substrate species. However, gene expression models can also be simplified to ensure better interpretability and lower computational complexity. The entire resource-dependent process can be described with an effective rate constant that reflects the concentrations of all competing substrates (Fig. 1a), assuming that very fast association and dissociation make a resource-substrate complex's concentration change very little on the time scale of other reactions in the cell<sup>9,12,13</sup>. Such simplified models allow to easily determine a genetic module's sensitivity to resource couplings<sup>6</sup> and to optimize design parameters like gene dosage and ribosome-binding site (RBS) strength, achieving desired outputs despite unwanted couplings<sup>12,14</sup>. Modeling insights can also help design circuits that efficiently mitigate unwanted couplings<sup>7,13</sup> or even leverage resource competition as a gene regulation mechanism to achieve other objectives, such as plasmid copy number-independent synthetic gene expression in mammalian cells<sup>15</sup>.

The analysis of resource couplings in bacteria is further complicated by the interplay between synthetic gene expression and cell growth rate. Bacterial cells are fast-growing, so dilution due to cell division significantly influences gene expression dynamics<sup>16</sup>. Accordingly, changes in the cell's growth rate caused by gene expression burden can qualitatively alter a gene circuit's behavior<sup>16–18</sup>. Moreover, the size of a bacterial cell's pool of gene expression machinery—specifically, ribosomes—is also related to its growth rate<sup>19,20</sup>. This interdependence, described by the experimentally observed “bacterial growth laws”<sup>21</sup> (Fig. 1b), arises because cells optimize their gene expression to maximize the steady-state growth rate in given environmental conditions<sup>22–25</sup>. These phenomena limit the predictive power of simple resource competition models, which assume constant growth rate and resource availability<sup>6,9,10</sup>. The development of bacterial resource-aware biocircuits thus calls for resource-aware models that consider synthetic circuits within the context of the host cell and account for the impact of burden on cell growth.

Since allocating resources between different genes to achieve maximum growth rate can be considered an optimization problem, solving it allows to approximately predict gene expression in certain conditions<sup>23,26,27</sup>. However, in reality living cells do not behave as ideal growth rate optimizers, but rather implement near-optimal gene regulation strategies via biological reactions<sup>23</sup>. Although some past studies have reproduced the bacterial growth laws by assuming

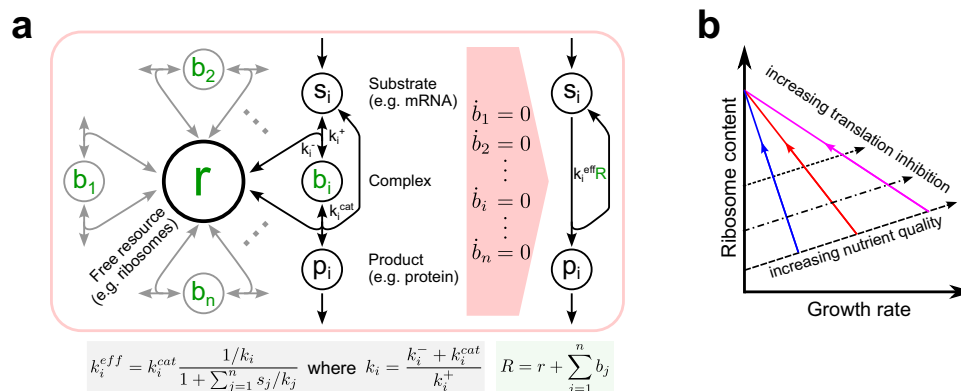
constitutive ribosome expression<sup>28,29</sup>, evidence increasingly suggests that near-optimal bacterial resource allocation is powered by the regulation of ribosomal genes' transcription by guanosine tetraphosphate (ppGpp), an alarmone molecule whose concentration lets the cell perceive its growth rate<sup>30,31</sup>.

Understanding how biochemical signals reflect cell growth rate and enable the optimization of gene expression requires mechanistic cell models, which incorporate the trade-offs faced by living cells, such as the finiteness of the cell's mass and its pool of gene expression machinery, energy, protein synthesis precursors, and other resources<sup>28,32</sup>. A range of such models with different levels of granularity exists, from whole-cell models<sup>33</sup> considering all known cellular processes to low-dimensional models, which consider the cell as a simple self-replicating machine that produces protein biomass in a single-step reaction. In the latter case, the cell's proteins with similar function and expression dynamics are grouped into several coarse-grained classes, and resource allocation modeling amounts to considering the ratios of different protein classes' mass fractions in the overall biomass<sup>22,23,26,34–37</sup>. Modeling frameworks that lie between these two extremes explicitly consider some aspects of gene expression and metabolic regulation—typically those believed most relevant for the phenomena they aim to explain—and adopt a simplified view of the others<sup>27–29,38–40</sup>.

The investigation of a host cell's interactions with synthetic gene circuits requires a balance between model realism and minimum complexity. Indeed, abstracting some gene expression steps or cellular signaling pathways carries the risk of neglecting important ways in which a synthetic device can influence the cell. Conversely, a finer view of gene expression and metabolic regulation, as well as the incorporation of biochemical interactions that are only significant in certain culture environments, may yield an overly detailed model with many unknown or unidentifiable parameters<sup>41</sup>. Excessive model complexity can also hinder informative biocircuit analysis and complicate the understanding of core biochemical processes defining the cell's state<sup>28</sup>.

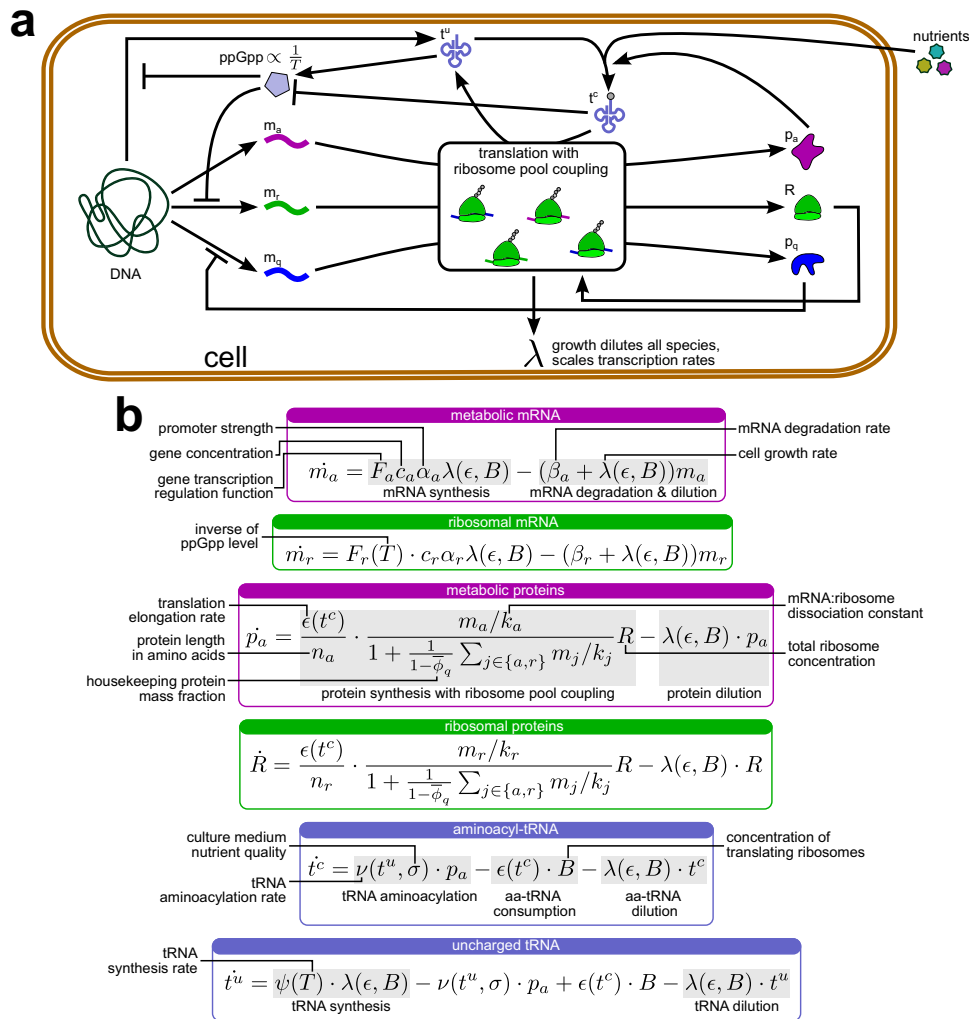
In this study, we aimed to enable easy yet reliable resource- and cellular context-aware design of synthetic circuits by developing a coarse-grained mechanistic model that would:

- (1) Based on physiologically meaningful parameters, allow to model the expression of a biocircuit's genes, considering their interactions via both indirect resource couplings and direct mechanisms commonly employed in synthetic biology.
- (2) Account for the context of the host cell and its interaction with synthetic genes, incorporating the key regulatory pathways



**Fig. 1 | Models and empirical laws for cellular resource availability.** **a** Instead of explicitly modeling the binding and unbinding of all competing substrates to the free resource (left), one can define the rate of a compound's synthesis as a product of the total abundance of the resource  $R$ , which can be free ( $r$ ) or bound to a substrate molecule ( $b_j$ ), and the effective rate constant  $k_i^{eff}$  (right). The effective rate constant depends on all competing species' concentrations, affinities to the

resource, and product synthesis rates<sup>9</sup>. **b** First (dashed line) and second (solid lines) bacterial growth laws relate growth rate to ribosome content in different conditions. Formulated by Scott et al.<sup>21</sup>, they postulate that the cell's ribosome content increases linearly with the growth rate as the culture medium's nutrient quality improves, but this relationship becomes inverse when translation is inhibited by an antibiotic.



**Fig. 2 | Schematic and definition of the proposed coarse-grained resource-aware cell model. a** Schematic of our host cell model. mRNAs are transcribed from genomic DNA and translated to make proteins. Competitive ribosome binding is not modeled explicitly—instead, effective rate constants relate each gene’s translation rate to the concentrations of protein precursors, ribosomes and all mRNAs in the cell. Ribosomes enable translation, whereas metabolic proteins catalyze nutrient import and tRNA aminoacylation. Native homeostatic regulation maintains housekeeping protein expression levels constant. The synthesis of tRNAs and

ribosomal mRNAs is repressed by ppGpp, whose concentration reflects the reciprocal of the ratio of concentrations of charged and uncharged tRNA concentrations. The overall rate of protein synthesis defines the growth rate, since dilution due to growth must keep the cell’s total protein mass constant. **b** The model’s ordinary differential equations (ODEs) and the physiological meaning of their terms. To avoid clutter, parameters and terms for different genes but with the same meaning are annotated only once.

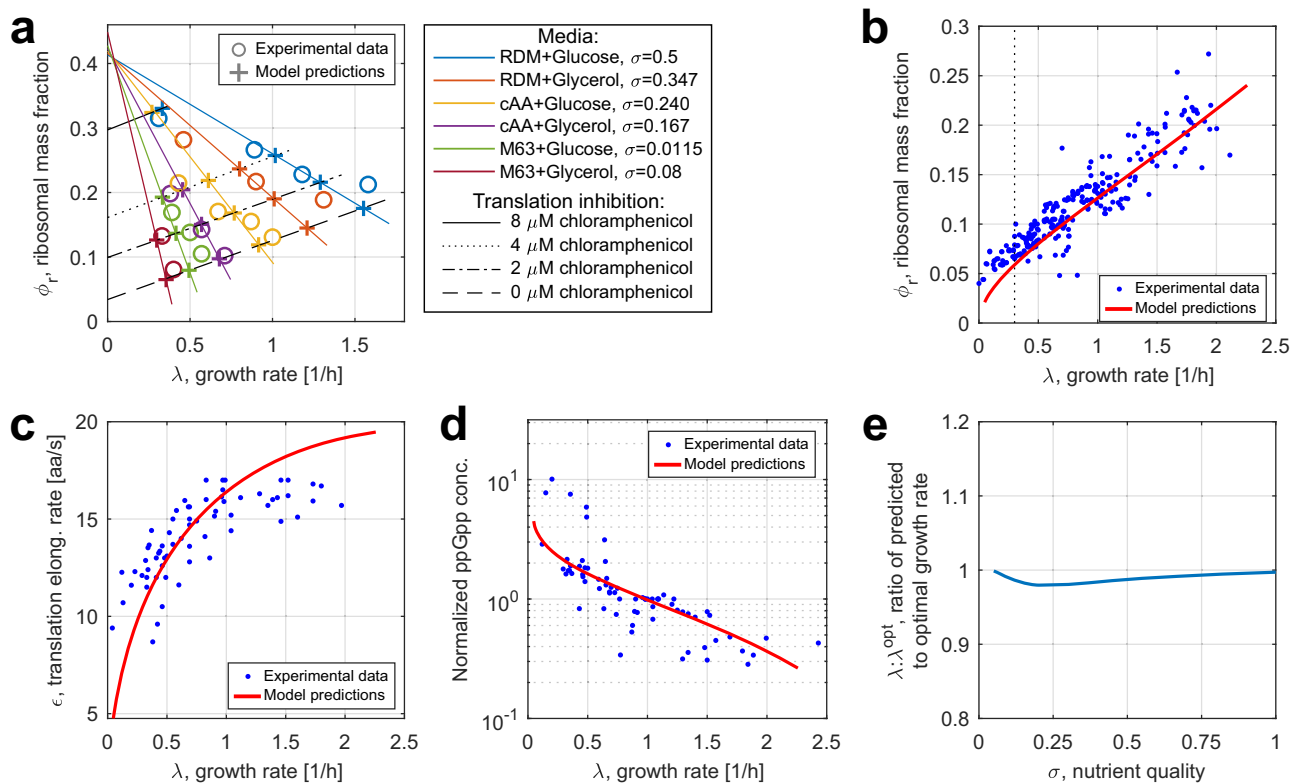
- presently understood to underlie cellular resource allocation and growth control, particularly ppGpp signaling.
- (3) Be minimally complex, enabling the derivation of informative analytical relations which can guide the choice of a biocircuit’s design parameters to achieve a desired behavior.

Defining such a model and parameterizing it for *E. coli*, we show that it reproduces experimentally observed bacterial growth phenomena, as well as empirical relations between the burden-dependent reduction in growth rate and different quantities characterizing heterologous gene expression in the cell. We also use our model to numerically reproduce the experimentally observed effects of resource competition on the behavior of self-activating gene circuits. Finally, we showcase our model’s usefulness for the development of resource-aware biocircuits by leveraging it to propose and analyze a biomolecular controller for mitigating gene expression burden. By maintaining near-constant ribosome availability at a cell-wide level, it reduces the effect of indirect couplings via the shared ribosome pool.

## Results

### A resource-aware cell model predicts growth phenomena

Our model, depicted in Fig. 2, distinguishes three bacterial gene classes, respectively labeled *r*, *a*, and *q*: ribosomal, metabolic, and housekeeping. The metabolic gene class enables tRNA aminoacylation and includes both aminoacyl-tRNA synthetases themselves and the enzymes enabling import and conversion of the culture medium’s nutrients into protein precursors. Other non-ribosomal genes belong to the housekeeping class, whose expression is regulated to be independent of culturing conditions<sup>19,20,26</sup>. Following a coarse-grained approach, all metabolic genes are treated as a single lumped gene; the same strategy is applied to ribosomal genes. Meanwhile, the abundance of housekeeping proteins in the cell is assumed constant—under a wide range of conditions, their share in the cell’s protein mass is fixed at  $\bar{\phi}_q \approx 0.59$ <sup>20</sup>. Hence, we avoid modeling their expression explicitly (Supplementary Note S1.5). Besides mRNA and protein expression, we model the concentrations of uncharged and charged (i.e., aminoacylated) tRNA molecules in the cell, since they play a key role in determining translation rates



**Fig. 3 | Growth phenomena predicted by our model.** **a** Experimental measurements by Scott et al.<sup>19</sup> used in parameter fitting (circles) and model predictions (crosses) for steady-state ribosomal mass fractions and growth rates. The first and the second growth laws are illustrated by varying culture media and extents of translation inhibition, respectively. Second growth law fits do not include the two predicted points for translation inhibition with 8  $\mu\text{M}$  of chloramphenicol due to them diverging from the linear growth law trend. Chloramphenicol action was modeled as outlined in Supplementary Note S2.1. **b–d** Comparison of model predictions (obtained by simulating Eqs. (15)–(20)) for the cell in steady state with experimental data from previous studies<sup>26</sup>. In **(b)**, the dotted vertical line denotes the  $\bar{\lambda} = 0.3$  threshold, left of which model predictions for  $\bar{\phi}_r$  significantly diverge

from the experimental data. In **(d)**, the ppGpp levels are normalized to the reference value for which  $\bar{\lambda} \approx 1 \text{ h}^{-1}$ . **e** Ratio of the steady-state growth rate  $\bar{\lambda}$  predicted by our model to the optimal growth rate  $\bar{\lambda}^{\text{opt}}$  for  $\sigma$  varied from 0.05 to 1. For each medium nutrient quality  $\sigma$  considered, we did the following. First, we recorded the steady-state cell growth rate  $\bar{\lambda}$  (here, the bar notation indicates a variable's steady-state value) predicted by our model. According to it,  $T$ , the inverse of ppGpp concentration, reflects the charged and uncharged tRNA levels as detailed by Eq. (28). Then, we assumed that ppGpp concentration in the cell is instead constant, ran simulations for different fixed values of  $T$  and identified the maximum steady-state cell growth rate  $\bar{\lambda}^{\text{opt}}$  across all considered values. Source data are provided as a Source Data file.

and resource allocation between gene classes<sup>31</sup>. Our modeling choices and derivations are detailed in “Methods” and Supplementary Note S1.

To validate model predictions against experimental data, we parameterized it for *E. coli*, the most studied bacterial model organism<sup>19,26,31,42</sup> and one of the most popular host organisms in synthetic biology<sup>43</sup>. Most parameter values, displayed in Supplementary Table 1, were taken from literature. The rest were fitted to experimental measurements of growth rate and ribosomal mass fraction of *E. coli* subjected to different concentrations of the translation-inhibiting antibiotic chloramphenicol in various growth media<sup>19,26</sup> as described in “Methods” and Supplementary Note S2.

As demonstrated by Fig. 3a, our model's steady-state predictions for different culture conditions are generally consistent with empirical bacterial growth laws illustrated in Fig. 1b<sup>19,21</sup>. Indeed, when nutrient quality improves and chloramphenicol levels remain unchanged, the cell's ribosome content increases linearly with the growth rate, obeying the first bacterial growth law. Moreover, augmenting translation elongation inhibition for the same nutrient quality produces an inverse proportionality between the ribosome content and growth rate for chloramphenicol levels of up to 4  $\mu\text{M}$ . We also confirmed the consistency of our model predictions for the cell's ribosome content, translation elongation rate, and ppGpp level with experimental results. To this end, we varied the nutrient quality factor  $\sigma$  without inhibiting translation and compared our model's steady-state predictions with

experimental data from the 37 different experimental studies compiled by Chure and Cremer<sup>26</sup> (Fig. 3b–d). As for the cell's dynamic behavior, in Supplementary Note S1.6 we compared experimental data with the predictions of our model's extension that allows to simulate nutrient upshift scenarios<sup>34,36</sup>.

Figure 3b supports the linear-like bacterial growth law dependence between ribosome content and growth rate, although for growth rates below  $\lambda \approx 0.3 \text{ h}^{-1}$  our model predictions take a nonlinear downturn, diverging from experimental measurements. Moreover, for *E. coli* subjected to 8  $\mu\text{M}$  of chloramphenicol, predictions do not strictly follow the linear trends. Notably, these discordances arise in highly unfavorable conditions—that is, strong inhibition of translation or very poor nutrient quality of the medium. Hence, to some extent they may be attributed to experimental errors, since it is difficult to ensure that measurements in such conditions pertain to cells in steady state<sup>26</sup>. Hostile environments can also trigger the cell's stress response mechanisms unconsidered by our model, such as ribosome inactivation, preserving the bacterium's ability to synthesize proteins in adverse conditions<sup>42</sup>. Furthermore, since mRNA synthesis rate depends on the cell's growth rate<sup>44</sup>, low-nutrient culture media reduce the overall concentration of transcripts in the cell. This leads to mRNA scarcity, rather than translational resource allocation that our model focuses on, becoming the key determinant of the growth rate<sup>27</sup>. Together, these factors limit our model's predictive power in unfavorable conditions.



Finally, in Fig. 3e we confirmed whether ppGpp regulation in our model reproduces near-optimal control of steady-state growth rates in a wide range of culturing conditions in line with prior modeling and experimental studies<sup>23,24,26</sup>. As shown in Fig. 3e, for  $0.05 \leq \sigma \leq 1$  the growth rate forecast by our ppGpp regulation model was always very close (within 2.3%) to the optimal value, which supports the notion of near-optimal control of resource allocation.

### Including heterologous gene expression in the cell model

Our cell model can be extended to consider heterologous gene expression and simulate gene circuit dynamics while accounting for resource couplings and the effects of heterologous gene expression on the host's metabolism that can influence the biocircuit's performance, such as growth rate changes. Besides native genes, let there be a set of  $L$  heterologous genes  $X = \{x_1, \dots, x_L\}$ . To model their expression, we add to Eqs. (15)–(20) a pair of Ordinary Differential Equations (ODEs) for each heterologous gene, which describes its transcription and translation (Supplementary Note S3.1). For a synthetic gene  $x_i$ , they are:

$$\dot{m}_{x_i} = F_{x_i} c_{x_i} \alpha_{x_i} \lambda(\epsilon, B) - (\beta_{x_i} + \lambda(\epsilon, B)) m_{x_i} \quad (1)$$

$$\dot{p}_{x_i} = \frac{\epsilon(t^c)}{n_{x_i}} \cdot \frac{m_{x_i}/k_{x_i}}{1 + \frac{1}{1-\bar{\phi}_q} \sum_{j \in \{a,r\} \cup X} m_j/k_j} R - \lambda(\epsilon, B) \cdot p_{x_i} \quad (2)$$

where all parameters have a similar definition to those of the parameters describing the native genes, while the transcription regulation function  $F_{x_i}$  is circuit- and gene-specific. Likewise to the original host cell model, we assume that transcriptional resource couplings are insignificant, as are the toxicity and active degradation of heterologous proteins, although the latter assumption can be lifted by making minor modifications as outlined in Supplementary Note S3.2. Meanwhile, the constant housekeeping protein mass fraction  $\bar{\phi}_q \approx 0.59$  remains the same despite synthetic gene expression<sup>19–21</sup>. Consequently, the model's original ODEs for native mRNA and tRNA levels are not altered by the addition of synthetic genes. Conversely, Eqs. (17) and (18) describe protein concentrations and thus include terms for translation. Therefore, they must be amended to consider additional synthetic mRNAs competing for ribosomes:

$$\dot{p}_a = \frac{\epsilon(t^c)}{n_a} \cdot \frac{m_a/k_a}{1 + \frac{1}{1-\bar{\phi}_q} \sum_{j \in \{a,r\} \cup X} m_j/k_j} R - \lambda(\epsilon, B) \cdot p_a \quad (3)$$

$$\dot{R} = \frac{\epsilon(t^c)}{n_r} \cdot \frac{m_r/k_r}{1 + \frac{1}{1-\bar{\phi}_q} \sum_{j \in \{a,r\} \cup X} m_j/k_j} R - \lambda(\epsilon, B) \cdot R \quad (4)$$

Combined, all extensions yield Model VI in Supplementary Note S3.1, which describes the cell and the synthetic circuitry it hosts. All synthetic genes' expression in this work was modeled using these ODEs, with the generic Supplementary Eqs. (75) and (76) for synthetic gene expression substituted by the specific ODEs for each circuit, which we explicitly provide in Supplementary Note S4.

To demonstrate how our modeling framework captures the implications of resource competition for synthetic gene expression, we recreated two experimentally documented cases of self-activating gene circuits' behavior being qualitatively altered by resource competition and the dependence of host cell growth on synthetic gene expression. First, we considered the "winner-takes-all" phenomenon<sup>5</sup>, the ODEs for which are provided in Supplementary Note S4.2.1. Alone in the host cell, a self-activating synthetic gene can act as a bistable switch with a high- and a low-expression stable steady states<sup>45</sup>. In the winner-takes-all scenario, two such switches in the same cell interact via the shared resource pool (Fig. 4a). Hence, if one switch

("the winner") reaches its high-expression equilibrium first, increased resource competition can prevent the other switch from ever reaching the corresponding high-expression steady state (Fig. 4b and Supplementary Fig. 4b–d).

Second, we reproduced the emergence of bistability due to host-circuit interactions<sup>46,47</sup> using the equations in Supplementary Note S4.3.1. Typically, self-activating genes act as bistable switches only if positive feedback is cooperative, i.e., a protein activates gene expression more efficiently if multiple copies of it are present. A heterologous T7 RNA polymerase (RNAP) transcribing its own gene exhibits non-cooperative self-activation (Fig. 4c), so its expression is expected to be monostable. However, by slowing the host cell's growth rate and thus protein and mRNA dilution, burden introduces an additional feedback loop which can confer bistability to this circuit despite its non-cooperativity as shown in Fig. 4d, e.

Nonetheless, simulations and bifurcation analysis<sup>48</sup> reveal that the winner-takes-all effect fades as the self-activating genes' resource demand is reduced (Supplementary Fig. 4e), whereas changing T7 RNAP's toxicity can render the non-cooperative self-activator monostable (Supplementary Fig. 5). Our model can therefore be used to determine whether given circuit design parameters can give rise to resource competition phenomena of interest.

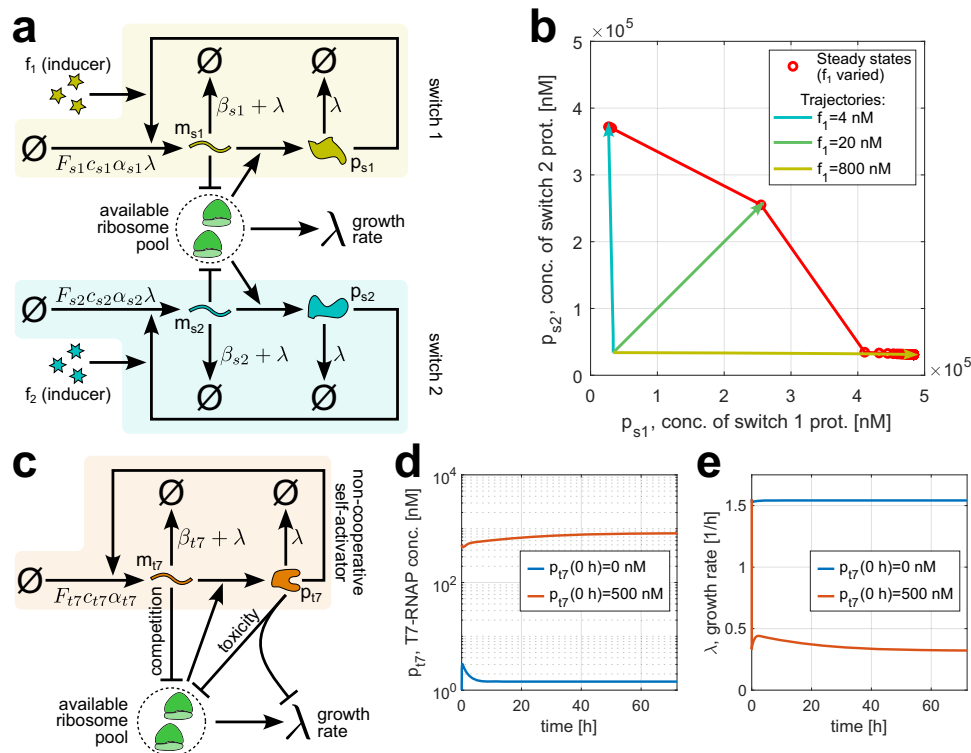
### Analytical predictions reveal how burden affects the cell

Besides enabling numerical prediction of circuit behavior, with some simplifying assumptions our model allows to derive analytical relations capturing the effect of different parameters on different variables' steady-state values. Namely, we assume that heterologous gene expression burden has very little effect on the cell's steady-state translation elongation rate and ribosomal gene transcription regulation function. While no experimental studies to date have directly evaluated these values' burden-dependence, in Supplementary Note S3.4 we show that this assertion directly follows from our model's definition, as well as confirm it numerically. This allows to postulate, regardless of which synthetic genes are present, that  $F_r \approx \bar{F}_r^{NB}$  and  $\epsilon \approx \bar{\epsilon}^{NB}$ , as well as  $k_i \approx \bar{k}_i^{NB}$  because the mRNA-ribosome dissociation constants are defined as functions of  $\bar{\epsilon}^{NB}$ . The <sup>NB</sup> ("no burden") index denotes steady-state values in absence of any synthetic gene expression, i.e.,  $X = \emptyset$ . While the formulae derived in this section require knowing the values of  $\bar{\epsilon}^{NB}$ ,  $\bar{F}_r^{NB}$ , and  $\bar{k}_i^{NB}$ , they can easily be retrieved by simulating the host cell model without synthetic genes for a given  $\sigma$ . Additionally, we assume that all mRNA molecules in the system decay at roughly the same rate, i.e.,  $\beta_i \approx \beta_j, \forall i, j \in \{a, \beta\} \cup X$ . Importantly, while the transcripts of individual native *E. coli* genes may have very different degradation rates<sup>49</sup>, the coarse-grained nature of our model means that this assumption only concerns the average degradation rates across the native gene classes, each of them spanning many genes.

Resource coupling analysis is commonly facilitated by lumping a gene's parameters into a coefficient which quantifies the gene's ability to seize expression resources and its own susceptibility to competition from other genes<sup>6,12,29</sup>. For a synthetic gene  $x_i \in X$ , we define it as the "translational burden"  $\xi_{x_i}$  shown in Eq. (5), where  $\bar{F}_{x_i}$  is the steady-state value of gene  $x_i$ 's transcription regulation function:

$$\xi_{x_i} = \frac{\bar{F}_{x_i} c_{x_i} \alpha_{x_i}}{\bar{k}_{x_i}^{NB}} \quad (5)$$

As revealed by the derivations in Supplementary Note S3.5, our translational burden factor combines the advantages of several existing lumped resource competition quantifiers. By considering the host cell's context, similarly to the "resource recruitment strengths"  $J$  defined by Santos-Navarro et al.<sup>29</sup>, we can use  $\xi_{x_i}$  values to capture the competition between synthetic and native genes. Namely, the steady-



**Fig. 4 | The model captures how resource competition and host-circuit interactions qualitatively alter circuit behavior.** **a** In the winner-takes-all scenario, the cell hosts two bistable switches, each consisting of a protein that, upon being allosterically modulated by an inducer molecule, cooperatively acts as a transcription factor activating its own expression. Adding the corresponding inducer to the medium causes a switch to move toward a high-expression equilibrium. For higher inducer levels, this happens on a shorter timescale (Supplementary Fig. 4a)<sup>5</sup>. **b** Phase plane diagram showing the two-switch system’s behavior upon simultaneous addition of inducer 1 and inducer 2 to the culture medium. Thin lines with arrowheads: the system trajectories for different concentrations of inducer 1 being added. Circles joined by a line: system’s final steady states as the concentration of inducer 1 increases from 4 nM to 800 nM with a step of 4 nM. In all cases, the concentration of inducer 2 being added is 20 nM. In line with winner-takes-all

behavior, the switch with a lower inducer concentration (slower activation) is prevented from reaching a high-expression equilibrium. Co-activation is achieved when the timescales of activation are identical for both switches, i.e.,  $f_1 = f_2 = 20 \text{ nM}$ <sup>5</sup>. **c** A heterologous T7 RNAP transcribing its own gene exhibits non-cooperative self-activation. However, bistability arises due to the host cell growth rate’s susceptibility to resource competition and synthetic protein toxicity. Note that here the gene transcription rate does not scale with cell growth due to transcription being enabled by heterologous machinery<sup>46</sup>. **d, e** Depending on the initial condition, RNA polymerase concentration and cell growth rate converge to different steady-state values, indicatively of bistability. For both trajectories shown, the initial condition for heterologous mRNA levels is  $m_{t7}(0 \text{ h}) = 0$ . The parameters and ODEs used in simulations are given in Supplementary Notes S4.2.1 and S4.3.1 for (b) and (d, e), respectively. Source data are provided as a Source Data file.

state mass fraction of heterologous protein in the cell as a function of  $\xi$  can be estimated as:

$$\bar{\phi}_X(\xi) \approx \frac{\xi}{\xi + \left(\sum_{j \in \{a,r\}} \bar{F}_j^{NB} c_j \alpha_j / \bar{k}_j^{NB}\right)} \text{ where } \xi = \sum_{x_i \in X} \xi_{x_i} \quad (6)$$

However, unlike resource recruitment strengths and similarly to the “resource demand coefficients”  $Q$  used by McBride and Del Vecchio<sup>6</sup>,  $\xi_{x_i}$  is independent of the cell growth rate. Hence, the growth rate itself can be analytically estimated from synthetic gene parameters according to Eq. (7):

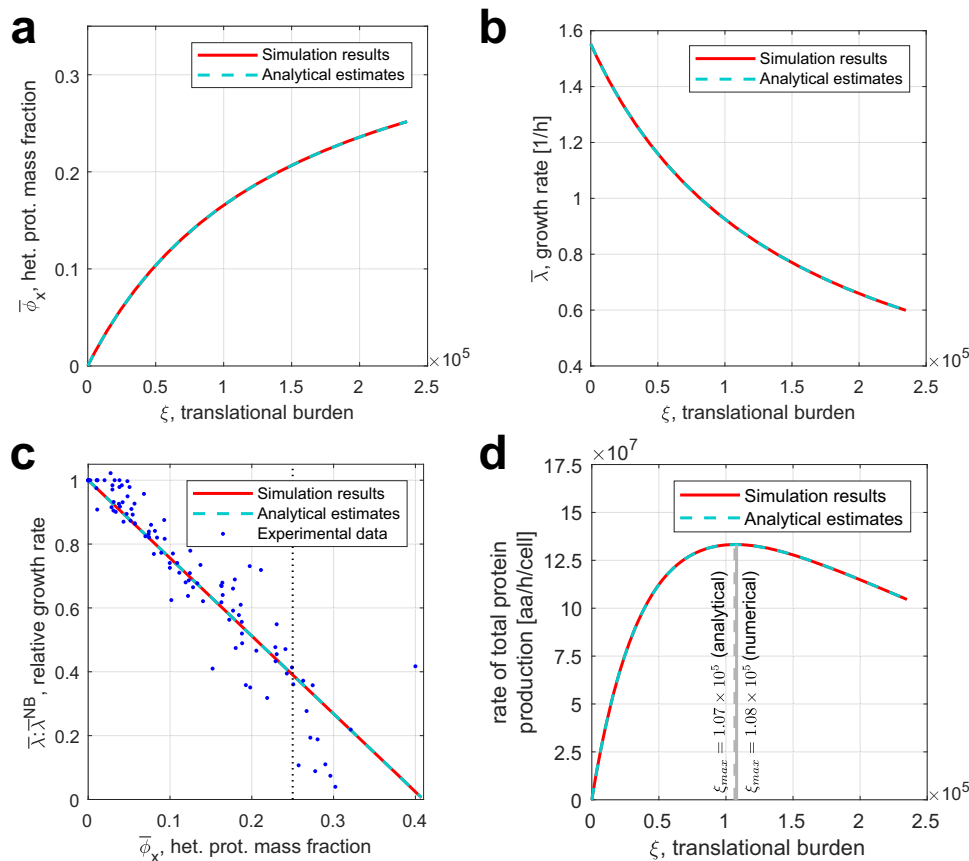
$$\bar{\lambda}(\xi) \approx \frac{\bar{e}^{NB} (1 - \bar{\phi}_q)}{M} \cdot \frac{\bar{F}_r^{NB} c_r \alpha_r / \bar{k}_r^{NB}}{\xi + \left(\bar{F}_r^{NB} c_r \alpha_r / \bar{k}_r^{NB} + c_a \alpha_a / \bar{k}_a^{NB}\right)} \quad (7)$$

A Hill relationship akin to Eq. (7) is sometimes used in resource-aware models that abstract the host’s native gene expression yet aim to capture cell growth rate’s burden-dependence<sup>17,18</sup>. However, these relations involve scaling factors defined arbitrarily for each synthetic circuit, such as the “metabolic burden threshold” in ref. 18. Conversely, our expression only involves the physiological parameters of native and synthetic genes.

Furthermore, the change in the cell growth rate relative to  $\bar{\lambda}^{NB}$  – that is,  $\bar{\lambda}(\xi = 0)$  – can be related to the total mass fraction of all heterologous proteins in the cell  $\bar{\phi}_X$  (Supplementary Note S3.6):

$$\frac{\bar{\lambda}}{\bar{\lambda}^{NB}} \approx 1 - \frac{\bar{\phi}_X}{1 - \bar{\phi}_q} \quad (8)$$

Several different empirical relations, such as Hill or linear dependencies, have been suggested to link gene expression burden to the reduction in cell growth rate<sup>6,19</sup>. Therefore, previous works have considered different formulae as part of their gene expression models in order to ensure that the modeling outcomes stay valid regardless of the assumed burden-growth dependency<sup>50</sup>. However, Eqs. (7) and (8) hint that these relations may not be mutually exclusive, but rather apply to different quantities describing heterologous gene expression. Namely, Eq. (7) relates growth rate to the translational burden – that is, a measure of the synthetic genes’ resource demand, determined by the circuit’s design parameters – giving rise to a Hill relation akin to that used in the work of McBride et al.<sup>6</sup>. If growth rate is linked to the heterologous protein yield, which is affected by resource couplings, a linear dependency given by Eq. (8) emerges, matching the observations of Scott et al.<sup>19</sup>. Our model therefore provides a possible unifying framework to explain different empirical laws describing the dependence of cell growth rates on synthetic gene expression.



**Fig. 5 | Numerical and analytical predictions for heterologous gene expression and its effects on the host cell.** **a, b** Hill activation and repression functions respectively relate translational burden to the steady-state heterologous protein mass fraction  $\bar{\phi}_X$  and cell growth rate  $\bar{\lambda}$ . **c** In the steady state,  $\bar{\phi}_X$ , the mass fraction of heterologous proteins in the cell, is linearly related to  $\bar{\lambda} : \bar{\lambda}^{NB}$ , the ratio of the corresponding growth rate to that without burden. The dotted vertical line denotes

the  $\bar{\phi}_X = 0.25$  threshold, right of which model predictions become unreliable. **d** The rate of total protein production per cell as a function of the translational burden.  $\xi_{\max}$  (vertical lines) is the burden value maximizing heterologous protein synthesis. All simulations assume  $\sigma = 0.5$ ; the parameters and ODEs used in simulations are given in Supplementary Note S3.1. Source data are provided as a Source Data file.

To compare our analytical estimates with the model's numerical predictions for different extents of translational burden, we simulated the expression of a constitutive heterologous gene by an *E. coli* cell (parameterized in Supplementary Table 4). The burden  $\xi$  exerted on the host by the synthetic gene was varied by sweeping through different values of the gene's DNA concentration from 1 nM to 1100 nM, producing the results plotted in Fig. 5a–c. Our approximate analytical predictions of Eqs. (6)–(8) closely follow the steady-state values obtained by numerical simulation (Fig. 5a–c, respectively), vindicating the assumptions made in the course of our analytical derivations. In Fig. 5c, both the analytical and numerical predictions are largely concordant with experimental data compiled by Chure and Cremer<sup>26</sup>, although the real and predicted measurements diverge when the heterologous protein overexpression stress is very high ( $\bar{\phi}_X > 0.25$ ). The reasons for this are likely similar to the general reasons why our model's predictions do not match experimental data in highly unfavorable conditions—namely, measurement errors<sup>26</sup>, ribosome inactivation<sup>42</sup>, and the dependence of slowly dividing cells' growth rate on their mRNA content<sup>27</sup>.

Analytical relations derived using our model and simplifying assumptions can facilitate the design of heterologous gene expression systems. For example, consider a population of *N. E. coli* cells expressing a constitutive heterologous protein of interest  $p_{poi}$  and dying at a constant rate  $\delta$ , modeled by Eq. (9):

$$\dot{N} = \lambda N - \delta N \quad (9)$$

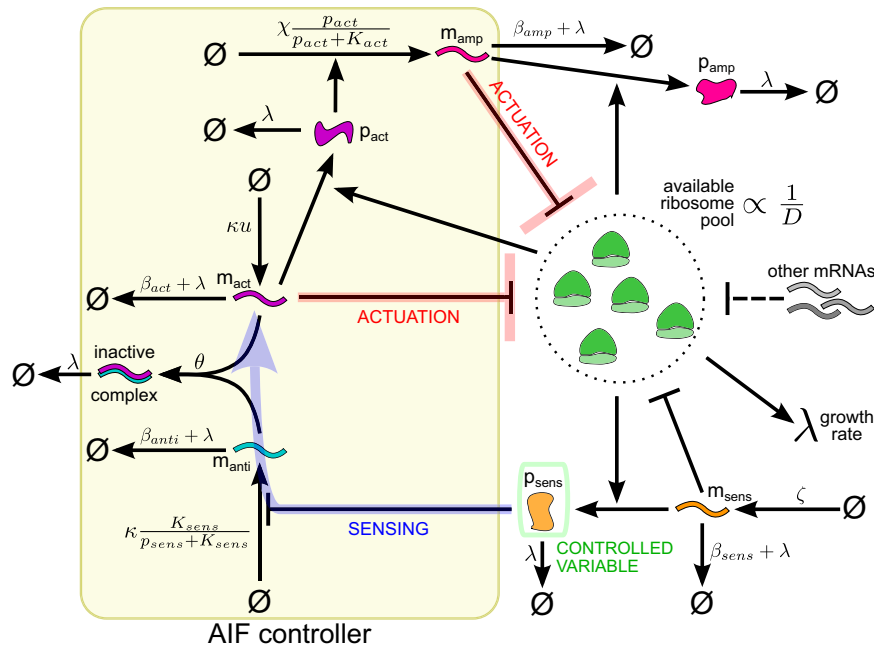
We can analytically find the optimal translational burden  $\xi_{\max}$  that maximizes the production rate  $\mu$  of the protein of interest by the cell population (Supplementary Note S3.7). This yields an analytical optimality condition based on the gene of interest's design parameters, namely its DNA copy number  $c_{poi}$ , promoter strength  $\alpha_{poi}$ , and apparent mRNA-ribosome dissociation constant  $\bar{k}_{poi}^{NB}$  (reflecting the RBS strength):

$$\frac{c_{poi} \alpha_{poi}}{\bar{k}_{poi}^{NB}} = \xi_{\max} = \frac{1 - \delta / \bar{\lambda}^{NB}}{1 + \delta / \bar{\lambda}^{NB}} \cdot \sum_{j \in \{a, r\}} \frac{\bar{F}_j c_j \alpha_j}{\bar{k}_j^{NB}} \quad (10)$$

In Fig. 5d, we plot the total protein production rate (calculated according to Supplementary Eq. (107) in Supplementary Note S3.7) as a function of the gene expression burden  $\xi$ , assuming that  $\delta = 0.25 \text{ h}^{-1}$ . Notably,  $\xi_{\max}$  yielded by Eq. (10) lies within 0.93% of the numerically found optimal value.

### Mitigating cell-wide resource couplings by integral feedback

To demonstrate how our model can facilitate resource-aware circuit development, we used it to design and analyze a biomolecular controller that mitigates resource couplings via the shared ribosome pool, reducing the impact of gene expression burden. Usually, robustness to couplings is achieved for a single variable or a limited set of synthetic genes that draw resources from an orthogonal pool, whose size is regulated by a biomolecular controller<sup>7</sup>. Conversely, our design reduces fluctuations in resource availability at the whole-cell level. Our cell



**Fig. 6 | Schematic of the integral biomolecular controller mitigating resource couplings.** If some disturbance increases the number of mRNAs competing for ribosomes, the concentration of the constitutive protein  $p_{sens}$  falls, which activates the production of annihilator RNA  $m_{anti}$ . Hence, the concentration of the actuator

mRNA  $m_{act}$  decreases; so does the expression of the amplifier mRNA  $m_{amp}$ , regulated by the actuator protein  $p_{act}$ . The ribosomes previously sequestered by the amplifier and actuator mRNAs are therefore released, increasing the resource availability back to its original level.

model is particularly suited for studying such controllers: by considering the expression of not just synthetic but all genes in the cell, it captures cell-wide resource couplings that our controller seeks to minimize. The analysis of these couplings is facilitated by the cell model's effective rate constant framework.

The proposed circuit, illustrated in Fig. 6 and modeled by Supplementary Eqs. (121)–(127), employs the Antithetic Integral Feedback (AIF) motif, which can maintain a physiological variable of interest at a desired setpoint value<sup>51,52</sup>. It comprises an annihilator and an actuator species, where the annihilator's synthesis rate depends on the controlled variable and the actuator species is produced at a constant rate, which sets the desired reference value. Since one annihilator and one actuator molecule can react to disable each other, the concentration of remaining non-annihilated actuator molecules reflects the integral of the error between the controlled variable and the reference. By influencing the rest of the system, the actuator molecules ensure robust perfect adaptation (RPA) to disturbances<sup>51,52</sup>. To control ribosome availability in the cell, we implement this motif using RNA logic, where the actuator is a protein-encoding mRNA and the annihilator is a small RNA (sRNA) that binds it to form a rapidly degraded complex<sup>53,54</sup>. As it does not involve protein synthesis, this RNA-based implementation is neither affected nor disrupted by translational resource couplings that we seek to mitigate, whereas transcriptional couplings in bacteria are largely insignificant<sup>6,14,55</sup>.

Besides the AIF motif's “actuator” and “annihilator” genes *act* and *anti*, the circuit comprises two more genes: the “sensor” *sens* and the “amplifier” *amp*. Ribosomal competition affects the level of the constitutively expressed transcription factor that is encoded by the sensor gene and regulates the annihilator sRNA's transcription. Actuation happens by changing concentrations of the transcripts competing for ribosomes. While the actuator mRNA sequesters some ribosomes, the protein encoded by it also regulates the transcription of an additional amplifier gene. This amplification is needed because heterologous mRNA concentration can significantly affect the ribosomal competition landscape only if it is large enough to be comparable to the

abundance of native transcripts. However, a substantial amount of actuator mRNAs is sequestered by the annihilator. Hence, achieving such a large concentration of heterologous transcripts by only expressing the actuator would require very high transcription rates. We therefore task another mRNA, which is not annihilated by any sRNA, with the bulk of the control action.

The variable of interest, whose fluctuations our controller aims to minimize, is the sensor protein's concentration  $p_{sens}$ , since the cell's resource availability is known to be captured by constitutive gene expression levels<sup>36</sup>. However, our model (with the simplifying assumptions outlined in the previous section) also allows to analytically retrieve the values of cell-wide variables that affect the expression of all genes in the cell. Namely, we can find the setpoint for the “resource competition denominator”:

$$D = 1 + \frac{1}{1 - \bar{\phi}_{q \in (a,r) \cup X}} \sum m_j/k_j, \quad (11)$$

In Eqs. (2)–(4), this value is the denominator of the effective protein synthesis rate constant, which adjusts the corresponding translation rate in view of competition from other transcripts. In a given culture medium and for a given steady-state value  $\bar{p}_{sens}$ , its setpoint value is given by Eq. (12). The meaning of the circuit's parameters found in this and other equations regarding our integral controller is illustrated in Fig. 6 and explained in Supplementary Tables 8 and 9.

$$\bar{D} = 1 + \frac{\bar{\lambda}\bar{\zeta}}{(\bar{\lambda} + \beta_{sens})\bar{k}_{sens}^{NB}} \cdot \left( \frac{n_{sens}}{M} \cdot K_{sens} \cdot \frac{1-u}{u} \right)^{-1} \quad (12)$$

Likewise, the steady-state growth rate maintained by the controller is given by:

$$\bar{\lambda} \approx \frac{\bar{\epsilon}^{NB}}{M} \cdot \bar{F}_r^{NB} \cdot c_r \alpha_r \cdot \frac{n_{sens} \bar{k}_{sens}^{NB}}{n_r k_r^{NB}} \cdot \frac{K_{sens}}{c_{sens} \alpha_{sens}} \cdot \frac{1-u}{u} \quad (13)$$



Besides setpoints for different variables, we can analytically estimate the controller's operation range—that is, the range of disturbance magnitudes which our design can negate to maintain a near-constant ribosome availability (Supplementary Note S4.4.2). If the cell is endowed with  $L$  synthetic genes  $\{x_1, x_2, \dots, x_L\}$  besides those of the controller, the disturbance caused by expressing them is mitigable only if:

$$\sum_{x_i \in \{x_1, x_2, \dots, x_L\}} \frac{\bar{F}_{x_i}^{NB} c_{x_i} \alpha_{x_i} \bar{\lambda}}{\bar{k}_{x_i}^{NB} (\bar{\lambda} + \beta_{x_i})} \leq \frac{c_{\text{sens}} \alpha_{\text{sens}} (1 - \bar{\phi}_q) M \bar{\lambda}}{K_{\text{sens}} n_{\text{sens}} \bar{k}_{\text{sens}}^{NB} (\bar{\lambda} + \beta_{\text{sens}})} \cdot \frac{1 - u}{u} - \sum_{j \in \{a, r, \text{sens}\}} \frac{\bar{F}_j^{NB} c_j \alpha_j \bar{\lambda}}{\bar{k}_j^{NB} (\bar{\lambda} + \beta_j)} \quad (14)$$

To test the proposed controller's performance and the accuracy of our estimates, we used our cell model to simulate how the integral controller reacts to the appearance of an additional mRNA species competing for the shared ribosome pool (Fig. 7a), plotting the outcome in Fig. 7b–f. Figure 7b, c shows that the amplifier mRNA levels fall in response to a step disturbance, restoring the original extent of resource competition. Consequently, the adaptation error—that is, the difference between the variables' values before and after disturbance—decreased almost twofold compared to the open-loop case where the annihilator, the actuator, and the amplifier are not expressed. This reduction of adaptation error is observed for various disturbing gene concentrations within the calculated operation range (Fig. 7g) and stays consistent in presence of parameter uncertainty, stochasticity of gene expression, and time-variant disturbance (Supplementary Figs. 6 and 7). Besides the abundance of transcripts competing for ribosomes, resource availability depends on the culture medium's nutrient quality. As Fig. 8 shows, the controller successfully reduces the adaptation error upon the induction of a disturbing gene in different media, as well as when nutrient quality and the competing mRNA's abundance are varied together.

Nevertheless, the observed adaptation errors in Fig. 7d–f are non-zero, which can be explained by “leakiness”. This phenomenon arises in AIF controllers when the actuator's and the annihilator's degradation and dilution rates are non-negligible compared to their rate of mutual elimination, which prevents the controlled variable from achieving RPA. Moreover, our derivations of the analytical estimates for  $\bar{\lambda}$  and  $\bar{D}$  neglect leakiness (Supplementary Note S4.4.2). Alongside the assumptions of constant  $\epsilon$  and  $F_r$  that we made in our derivations, this has likely contributed to the observed discrepancies between the numerically obtained and analytically estimated values for these variables. However, although leakiness can arbitrarily deteriorate the performance of AIF controllers, the observed adaptation errors remain relatively low, while changing our circuit's parameters can further improve the controller's performance. Figure 7g shows how increasing the amplifier gain  $\chi$  (i.e., the maximum amplifier mRNA production rate) lowers the adaptation error even as the disturbance's magnitude rises. Increasing the actuator and the annihilator RNAs' transcription rate  $\kappa$  has been shown to reduce the adaptation error of the controlled variable's mean value<sup>33</sup>—although in some cases this can lead to instability<sup>57</sup> (Supplementary Fig. 8).

A promising application for our controller is enforcing the assumption of modularity in synthetic biology, which is typically invalidated by resource couplings<sup>12</sup>. For instance, the shape of a gene's induction curve can be altered if competing synthetic genes are expressed in parallel<sup>6</sup>. However, our controller renders the induction curve more robust to different disturbances (Fig. 9). Notably, the inducible module, the disturbing gene, and the constitutive sensor gene in this case interact only via the shared resource pool and not directly. The inducible module's improved robustness to perturbations therefore demonstrates that our controlled variable—i.e., the

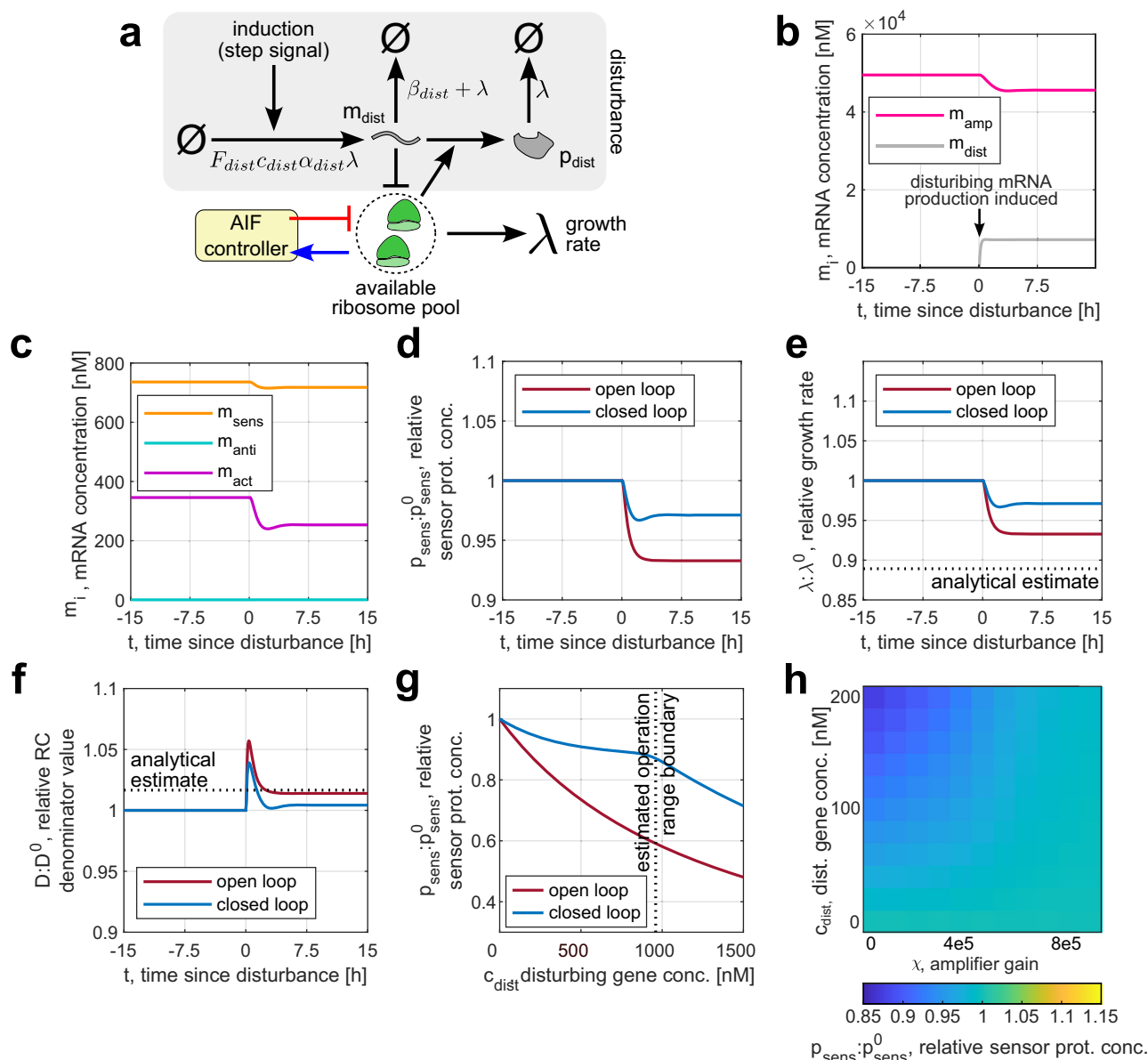
sensor gene's expression—indeed reflects resource availability at the level of the entire host cell, and that by keeping it constant our integral controller mitigates fluctuations in burden. Besides maintaining modularity, our controller could slow the loss of synthetic genes by engineered cells. A loss of heterologous gene expression to mutation normally means that the growth rate is less impaired by gene expression burden, so mutants outcompete the original engineered cells<sup>58</sup>. Conversely, as long as our controller itself remains functional, it keeps resource availability roughly constant regardless of whether any other synthetic genes are expressed. Without the reduction of translational burden, mutants lack the competitive advantage over the engineered cells.

Notably, our controller produces many transcripts that sequester ribosomes to release them when disturbances increase resource demand (Fig. 6), which burdens the cells expressing it. This reduction of resource availability and cell growth in exchange for control over cellular variables and robustness to perturbations is a common trade-off of growth rate controllers<sup>59,60</sup> and feedback circuits that mitigate fluctuations in resource demand<sup>11,61</sup>. Meanwhile, non-burdensome mitigation of cell-wide resource availability changes is currently restricted to feedforward architectures where the disturbing and the controller gene must be co-regulated<sup>62</sup>. Conversely, by sensing ribosome availability itself via the sensor gene's expression level, our feedback controller has no such requirement, mitigating perturbations irrespective of their cause.

## Discussion

The influence of resource competition on synthetic circuit performance, both through indirect interactions between different synthetic genes and through the burden imposed on the host cell, can be seen from several perspectives. One approach exclusively considers synthetic genes, applying effective rate frameworks to obtain a simple, easily interpretable model that predicts a circuit's behavior and the impact of changing its design parameters<sup>6,9,12</sup>. In this case, the interplay between heterologous gene expression and the host cell's state is either neglected<sup>12</sup> or abstracted and captured by simple phenomenological relations, which involve circuit-specific parameters without clear physiological meaning and do not mechanistically explain the effects of burden<sup>18</sup>. On the other hand, models of the entire bacterial cell reveal how resources are allocated between the host's own genes and those of the synthetic circuit<sup>28,38</sup>. However, highly complex cell models involving many variables and parameters obscure the key underlying processes and do not allow analytical derivations, requiring extensive numerical simulations to understand the effect of a given design choice. Meanwhile, excessive coarse-graining may leave out the interactions that are crucial for understanding bacterial resource allocation and designing gene circuits.

Our modeling framework combines the strengths of simplified resource competition analysis frameworks and cell models, balancing coarse-graining with accurate representation of gene expression and regulation. To this end, we apply an established resource-aware modeling framework, based on effective reaction rate constants<sup>9,12</sup>, to a coarse-grained model which relates protein synthesis to cell growth via the finite proteome trade-off<sup>28</sup> and captures the dynamics of near-optimal resource allocation in bacteria by incorporating the principles of flux-parity regulation of gene expression<sup>26</sup>. The host cell model, spanning only six variables, can be easily augmented with ODEs mirroring Eqs. (1) and (2) for each heterologous gene to describe an arbitrary synthetic gene circuit. Resource- and host-aware simulations of a circuit's behavior can then be performed using our Matlab implementation of the model found at [https://github.com/KSechkar/rc\\_e\\_coli](https://github.com/KSechkar/rc_e_coli)<sup>63</sup>, enabling rapid and cheap prototyping of resource-aware biomolecular controllers in silico<sup>64</sup>. Besides numerical analyses, our framework's simplicity allows to obtain a range of analytical relations between a synthetic circuit's design parameters and the host cell's



**Fig. 7 | Simulation of the integral controller's response to disturbance.**

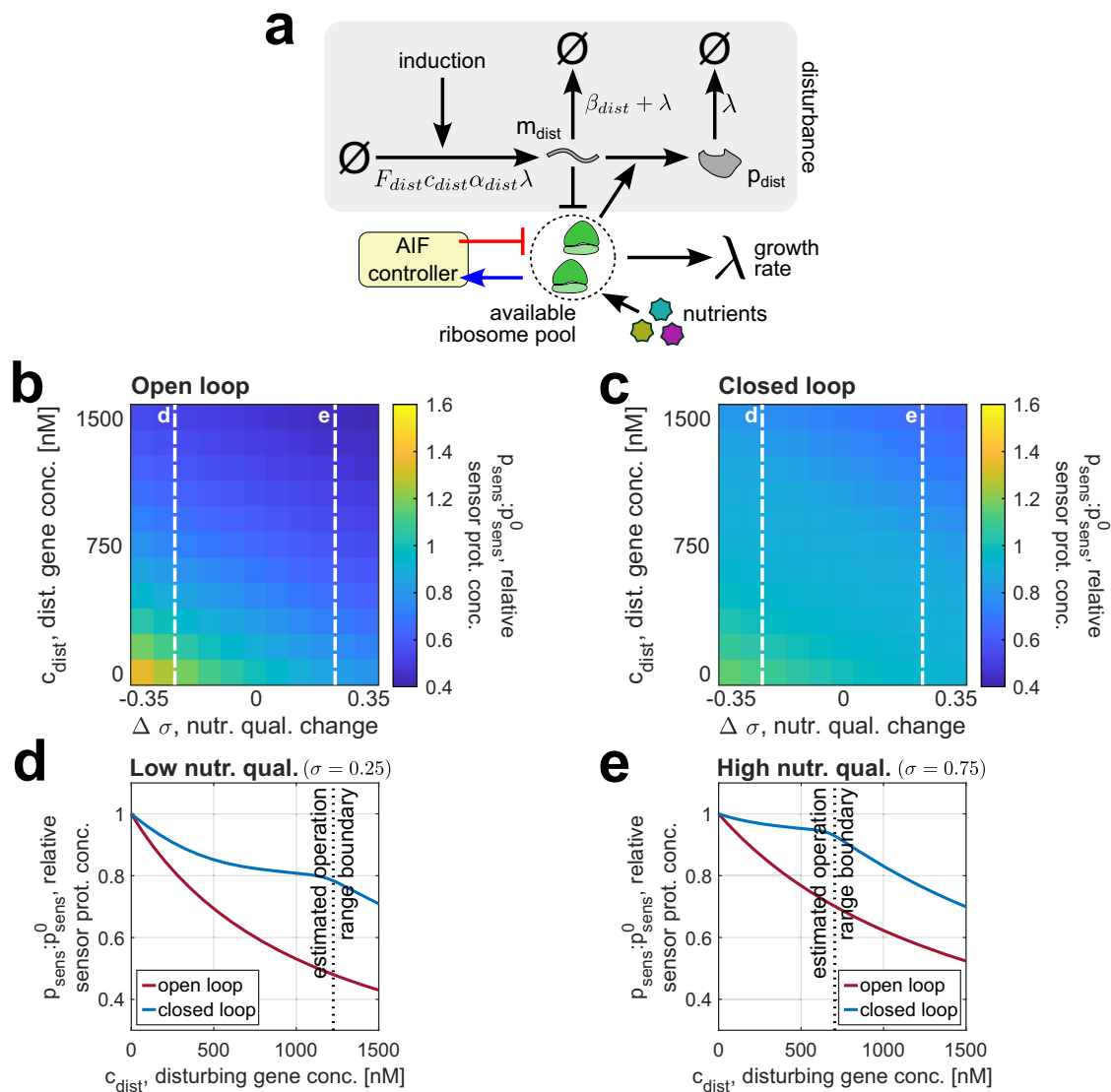
**a** Resource availability in the cell is disturbed by inducing the expression of an additional disturbing synthetic gene  $dist$ . The resulting change in burden is counteracted by the controller. **b, c** Evolution of simulated mRNA concentrations over time upon the induction of a disturbing gene with the parameters given in Supplementary Table 9. **d–f** Simulated trajectories for the sensor protein's concentration  $p_{sens}$  (the controlled variable), the cell's growth rate  $\lambda$ , and the resource competition denominator  $D$  upon the induction of a disturbing gene. For comparison, the response of an open-loop system (i.e., in absence of the actuator, the annihilator, and the amplifier genes) is plotted on the same axes. All variables are plotted relative to their steady-state values. **g** Dependence of the steady-state value of  $p_{sens}$ , relative to that for an undisturbed cell, on the magnitude of disturbance—

that is, the concentration of the disturbing gene's DNA in the cell. Open-loop responses are plotted on the same axes for comparison. Outside of the calculated operation range, the controller no longer mitigates resource couplings, so the closed-loop value of  $p_{sens}$  starts falling with increasing  $c_{sens}$  at the same rate as the open-loop value. **h** The steady-state value of  $p_{sens}$  maintained by the controller for various values of the amplifier gain  $\chi$  and the disturbing gene's DNA concentration  $c_{dist}$ . The adaptation error increases as the disturbance rises but can be mitigated by increasing the amplifier gain. Unless specified otherwise, simulation parameters for all panels are given in Supplementary Table 9, and the ODEs used to simulate the controller are Supplementary Eqs. (L21)–(L29) in Supplementary Note S4.4.1. Source data are provided as a Source Data file.

growth rate, reproducing the empirical relations commonly used to capture burden in synthetic gene expression models<sup>6,17–19</sup>. Consequently, our model not only yields the forms of these relations which are rooted in physiological parameters, but also provides them with a unifying framework.

Comprising only a few variables, our model nonetheless escapes the major limitations of using very low-dimensional biomass-centric frameworks, such as the original flux-parity regulation model<sup>26</sup>, in resource-aware circuit analysis. Namely, rather than considering molecules' fractions in the bacterial population's total biomass, our

model predicts their cellular concentrations, which are the quantities governing the biological interactions commonly leveraged in synthetic biology (e.g., transcription factor binding to DNA). Furthermore, as demonstrated by the different curve shapes in Fig. 5b, c, circuit design parameters' effect on the host cell is not directly reflected by the relationship between the cell growth rate and the synthetic protein mass fraction, but can be captured by our model. Unlike models that treat protein production as a single step<sup>23,26,29</sup>, our framework also allows to model RNA-based regulation, which, by avoiding the most burdensome and resource competition-dependent step of



**Fig. 8 | Dependence of the controller's performance on the culture medium.**

**a** Resource availability in the cell can be perturbed both by introducing additional genes competing for ribosomes and changing the culture medium's nutrient quality, both of which our controller is designed to counteract. **b, c** Dependence of the steady-state value of  $p_{\text{sens}}$  in an open- and closed-loop system, relative to that for an undisturbed cell in the default medium (i.e.,  $c_{\text{dist}} = 0, \sigma = 0.5$ ), on the disturbing gene's concentration and the change in the medium's nutrient quality. The dashed white lines mark the nutrient qualities considered in **(d)** and **(e)**. **d, e** Dependence of the steady-state value of  $p_{\text{sens}}$ , relative to that for an

undisturbed cell, on the disturbing gene's concentration for media with low ( $\sigma = 0.25$ , standing for  $\lambda \approx 0.566 \text{ h}^{-1}$  in the closed-loop case) and high ( $\sigma = 0.75$ , standing for  $\lambda \approx 1.462 \text{ h}^{-1}$  in the closed-loop case) nutrient qualities. Note that the undisturbed sensor protein concentration  $p_{\text{sens}}^0$  is calculated for  $c_{\text{dist}} = 0$  and the current nutrient quality—that is,  $\sigma = 0.25$  or  $\sigma = 0.75$ , as opposed to **(b, c)**, where it was defined for  $\sigma = 0.5$  in all cases. Unless specified otherwise, the simulation parameters for all panels are given in Supplementary Table 9, and the ODEs used to simulate the controller are Supplementary Eqs.(121)–(129) in Supplementary Note S4.4.1. Source data are provided as a Source Data file.

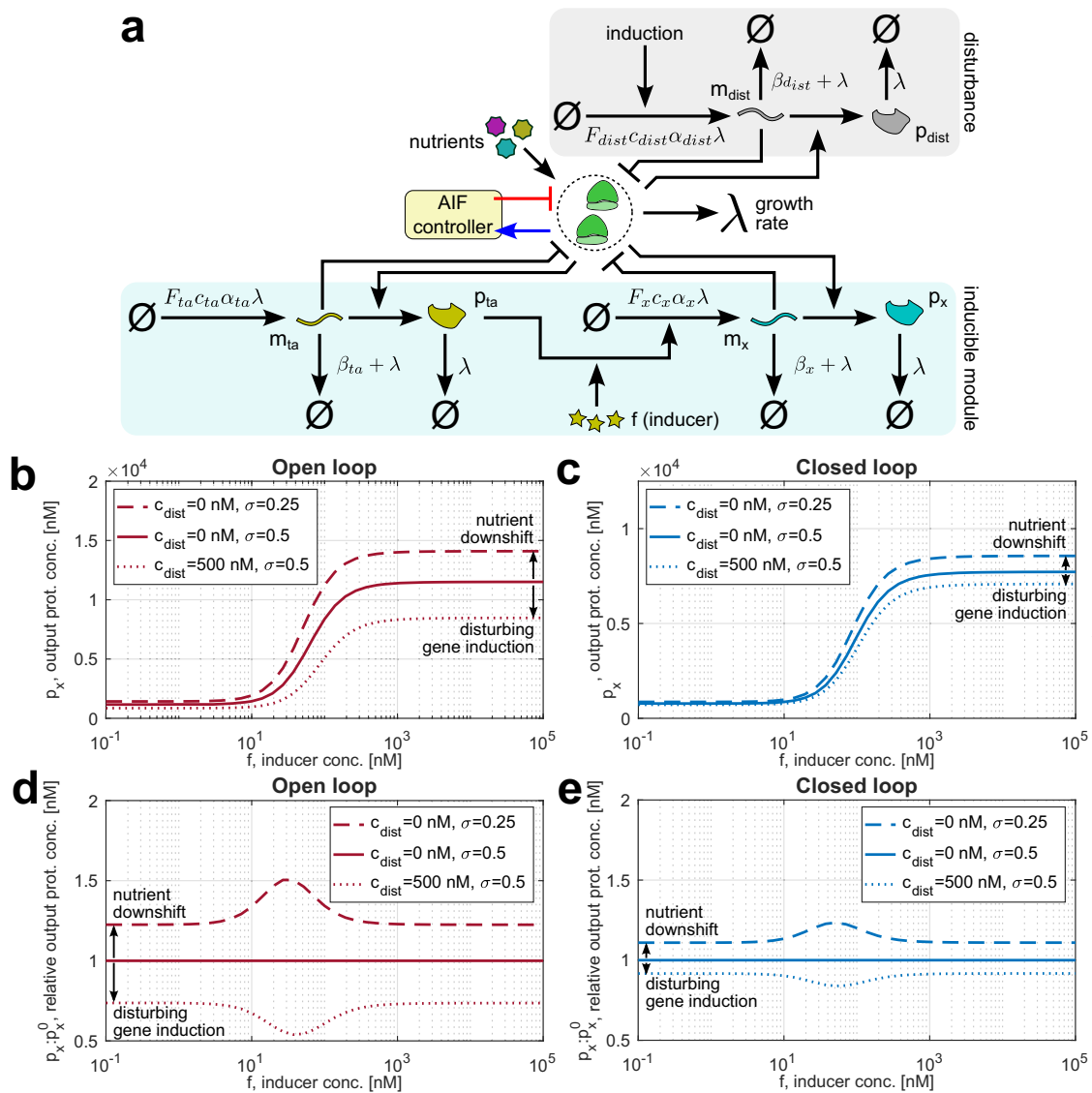
translation<sup>56</sup>, can be particularly useful in resource-aware circuit development.

We showcase our approach's advantages by designing and analyzing an integral controller that leverages RNA-based logic to mitigate fluctuations of the host cell's ribosome availability. The proposed controller acts at a cell-wide level, making our host-aware modeling framework more suited for its analysis than simpler resource competition models that neglect the cellular context. Our model enables numerical simulations of the controller. Furthermore, while the circuit infers resource availability via the proxy of the sensor gene's expression, the cell growth rate and extent of resource competition corresponding to the controller's setpoint, as well as our design's operation range, can be readily estimated using analytical relations.

Our model's reliability is supported by the concordance of its predictions with published experimental data. Although our model

describes *E. coli* cells, the finite proteome trade-off is believed to be common for most growing bacterial cells<sup>28</sup>, while several other bacteria, such as *S. coelicolor*, exhibit behavior consistent with flux-parity regulation of gene expression<sup>26,65</sup>. This makes our framework potentially applicable to other species upon re-parameterization. The model can also be extended to account for the aspects of cell physiology which are currently omitted but may be particularly relevant for certain applications. For instance, transcriptional resource couplings may become important in the cases of very high heterologous mRNA production. They can be captured by adopting the effective rate constant approach to simplify competitive RNA polymerase binding dynamics similarly to how we considered ribosome pool couplings<sup>12</sup>.

In conclusion, our work presents a case for combining the insights from resource competition analysis with the appreciation of resource allocation in bacterial cells, which makes it an important step toward



**Fig. 9 | The enforcement of modularity by our integral controller.** **a** Our integral controller mitigates the effect of competing synthetic genes and changes of the culture medium on the performance of an inducible genetic module, which consists of a transcription activation factor  $ta$  and the output gene  $x$  regulated by it. **b, c** The module of interest’s induction curve (steady-state output protein concentrations plotted against inducer levels in the medium) in open- and closed-loop setups with and without different disturbances. **d, e** Open- and closed-loop steady-

state concentrations of the output protein  $p_x$  relative to their values for a given inducer concentration  $f$  in the default culture medium and without disturbing gene expression. Unless specified otherwise, simulation parameters for all panels are given in Supplementary Table 9, and the ODEs used to simulate the controller are Supplementary Eqs. (121)–(134) in Supplementary Note S4.4.1. Source data are provided as a Source Data file.

easy design of reliable resource coupling-based controllers. Our model’s predictive power can be further enhanced by considering more types of couplings via different resource pools and incorporating the results of future experiments characterizing cell physiology. By providing numerical and analytical tools for developing resource-aware circuits with an appreciation for their interactions with the host, we hope to encourage further exploration of bacterial resource-aware controller designs and open up avenues for holistic analysis of bio-circuit performance in the context of the host cell.

## Methods

### Model definition

Our mechanistic cell model is given by Eqs. (15)–(20), where  $m_a$  and  $m_r$  are respectively the concentrations of metabolic and ribosomal mRNAs,  $p_a$  is the metabolic protein concentration,  $R$  is the cell’s total ribosome count, and  $t^c$  and  $t^u$  are the concentrations of aminoacylated

and uncharged tRNAs. The modeling assumptions giving rise to these equations are summarized in Table 1 and explained below:

$$\dot{m}_a = F_a c_a \alpha_a \lambda(\epsilon, B) - (\beta_a + \lambda(\epsilon, B)) m_a \quad (15)$$

$$\dot{m}_r = F_r(T) \cdot c_r \alpha_r \lambda(\epsilon, B) - (\beta_r + \lambda(\epsilon, B)) m_r \quad (16)$$

$$\dot{p}_a = \frac{\epsilon(t^c)}{n_a} \cdot \frac{m_a/k_a}{1 + \frac{1}{1-\phi_q} \sum_{j \in \{a,r\}} m_j/k_j} R - \lambda(\epsilon, B) \cdot p_a \quad (17)$$

$$\dot{R} = \frac{\epsilon(t^c)}{n_r} \cdot \frac{m_r/k_r}{1 + \frac{1}{1-\phi_q} \sum_{j \in \{a,r\}} m_j/k_j} R - \lambda(\epsilon, B) \cdot R \quad (18)$$

**Table 1 | Summary of modeling assumptions and equation terms affected by them**

Assumption	Equations	Relevant terms <sup>a,b</sup>
<b>mRNA synthesis</b>		
Negligible transcriptional couplings	(15), (16), (21)	$F_i \cdot c_i \cdot \alpha_i \cdot \lambda$
Constitutive metabolic gene expression	(15)	$F_a \equiv 1$
mRNA synthesis rate proportional to cell growth rate	(15), (16), (21)	$F_i \cdot c_i \cdot \alpha_i \cdot \lambda$
<b>Protein synthesis</b>		
Translational couplings captured by effective rate constants	(17), (18)	$\frac{\epsilon(t^c)}{n_i} \cdot \frac{m_j/k_j}{1 + \frac{1}{1-\phi_q} \sum_{j \in \{a,r\}} m_j/k_j} R$
Housekeeping protein mass fraction fixed	(17), (18), (22)	$\phi_q \equiv \bar{\phi}_q \equiv 0.59$
Housekeeping gene expression not modeled explicitly	(17), (18), (22)	$\frac{m_j/k_j}{1 + \frac{1}{1-\phi_q} \sum_{j \in \{a,r\}} m_j/k_j}$
<b>Degradation and dilution of species</b>		
mRNA degradation rate comparable to cell growth rate	(15), (16)	$-(\lambda + \beta_i)m_i$
Protein degradation rate negligible compared to cell growth rate	(17), (18)	$-\lambda p_i$
tRNA degradation rate negligible compared to cell growth rate	(19), (20)	$-\lambda t^c, -\lambda t^u$
<b>Cell growth rate regulation</b>		
Cell growth rate maintains constant protein mass per unit of cell volume	(25)	$\lambda(\epsilon, B) = \frac{\epsilon B}{M}$
<b>Translation elongation</b>		
Translation elongation rate governed by Michaelis-Menten kinetics	(26)	$\epsilon(t^c) = \epsilon_{\max} \frac{t^c}{t^c + K_\epsilon}$
<b>Protein precursor synthesis</b>		
tRNA aminoacylation rate governed by Michaelis-Menten kinetics	(27)	$\nu(t^u, \sigma) = \nu_{\max} \sigma \frac{t^u}{t^u + K_\nu}$
Culture medium's nutrient quality captured by the nutrient quality factor $\sigma$	(27)	$\nu(t^u, \sigma) = \nu_{\max} \sigma \frac{t^u}{t^u + K_\nu}$
<b>Flux-parity resource allocation</b>		
ppGpp level reflects the ratio of charged and uncharged tRNA concentrations	(28)	$T = \frac{t^c}{t^u} \propto \frac{1}{[\text{ppGpp}]}$
Ribosome synthesis regulated by ppGpp	(29)	$F_r(T) = \frac{T}{T + \tau}$
tRNA and ribosome synthesis co-regulated	(30)	$\psi(T) = \psi_{\max} F_r(T)$

<sup>a</sup>Generic index  $i$  means that the term pertains to both  $a$  and  $r$  genes.

<sup>b</sup>Where not all factors of a term reflect the assumption, the relevant part is given in bold.

$$t^c = \nu(t^u, \sigma) \cdot p_a - \epsilon(t^c) \cdot B - \lambda(\epsilon, B) \cdot t^c \quad (19)$$

$$t^u = \psi(T) \cdot \lambda(\epsilon, B) - \nu(t^u, \sigma) \cdot p_a + \epsilon(t^c) \cdot B - \lambda(\epsilon, B) \cdot t^u \quad (20)$$

**mRNA synthesis.** Gene expression and growth in bacteria are primarily affected by translational, rather than transcriptional, resource availability<sup>6,14,55</sup>. For simplicity, we therefore do not consider competition for transcriptional resources (e.g., RNA polymerases), so the rate of mRNA synthesis in Eqs. (15) and (16) is given by:

$$F_i \cdot c_i \cdot \alpha_i \cdot \lambda \quad (21)$$

where  $\alpha_i$  is the corresponding gene's promoter strength and  $c_i$  is its DNA's concentration, which for the cell's native genes is assumed to be 1 nM as a convention (since the volume of an *E. coli* cell is around  $1 \mu\text{m}^3$ , this is equivalent to one gene copy per cell<sup>12</sup>).  $F_i$  is the dimensionless transcription regulation function. For the ribosomal genes,  $F_r$  captures the effects of ppGpp regulation and will be defined later in this section. The metabolic gene class as a whole is commonly treated as

constitutively expressed<sup>19,21,22</sup>. Likewise, the expression of aminoacyl-tRNA synthetases—which we assume to be part of the metabolic gene class, despite them sometimes not belonging to it in some other models<sup>19,28,38</sup>—exhibits little to no dependence on ppGpp levels<sup>66</sup>. Hence, we use  $F_a \equiv 1$  across all culturing conditions. Finally, Eq. (21) includes the cell's growth rate  $\lambda$ , as RNA production rates across the bacterial genome increase linearly with the growth rate, presumably due to changes in the availability of the  $\sigma^{70}$  factor, which is responsible for the transcription of most genes in exponentially growing bacteria<sup>44,67</sup>. Consequently, since the transcription rate is measured in nM of mRNA synthesized per hour, while the units of  $c_i$  and  $\lambda$  are respectively nM and  $\text{h}^{-1}$ , the promoter strength  $\alpha_i$  is dimensionless. In Supplementary Note S2.3, we also discuss how the value of  $\alpha_i$  captures the possibility of a single mRNA molecule being simultaneously translated by multiple ribosomes<sup>12,68</sup>.

**Protein synthesis.** To define protein synthesis rates in Eqs. (17) and (18), we adopt an approach similar to that used in refs. 9, 12. Hence, in Supplementary Notes S1.2 and S1.3 we derive the effective translation rate constants:

$$k_i^{\text{eff}} = \frac{\epsilon}{n_i} \cdot \frac{1/k_i}{1 + \frac{1}{1-\phi_q} \cdot \sum_{j \in \{a,r\}} m_j/k_j} \quad (22)$$

where

$$k_i = \frac{k_i^- + \epsilon/n_i}{k_i^+} \quad (23)$$

is the apparent mRNA-ribosome dissociation constant, determined by the binding and unbinding rates between the RBS and the ribosome ( $k_i^+$  and  $k_i^-$ , respectively) and the rate at which the ribosome completes translation to slide off the mRNA. The latter can be obtained by dividing  $\epsilon$ , the translation elongation rate in amino acids per hour, by  $n_i$ , the number of amino acid residues in the protein encoded by the gene. Notably, owing to our assumption that the mass fraction of housekeeping proteins in the cell  $\bar{\phi}_q \approx 0.59$  always remains constant<sup>20,26</sup>, we avoid modeling the expression of housekeeping genes explicitly. Hence, Eq. (22) captures the housekeeping mRNAs' ribosome demand via the factor  $\frac{1}{1-\phi_q}$  (Supplementary Note S1.5).

**Dilution and degradation of species.** As for the removal of molecules, all species are diluted at a rate  $\lambda$  due to the cell growing and dividing. The overwhelming majority of proteins are not actively degraded in the bacterial cell<sup>69</sup>. Hence, save for extremely unfavorable culture conditions that our model does not aim to describe, the total rate of protein breakdown in the cell is at most  $0.02\text{--}0.025 \text{ h}^{-1}$ , at least an order of magnitude smaller than the typical growth rates of  $0.3\text{--}2 \text{ h}^{-1}$  considered in our study<sup>70</sup>. Protein degradation is therefore negligible compared to dilution. Conversely, mRNA degradation is widespread and happens at a high rate<sup>49</sup>. Hence, we include additional constant degradation terms  $\beta_a$  and  $\beta_r$  in Eqs. (15) and (16), which describe mRNA concentration dynamics. In contrast to mRNAs, the highly stable secondary and tertiary structure of tRNA molecules protect them from degradation both in and out of steady-state cell growth regimes, which makes the degradation rate of tRNA vanishingly small on the timescale of cell division<sup>71,72</sup>. Therefore, the species  $t^c$  and  $t^u$  in Eqs. (19) and (20) are removed by dilution only.

**Cell growth rate regulation.** The growth rate  $\lambda$  is related to the rate of protein synthesis in the cell by the finite proteome cellular trade-off, identified by Weisse et al.<sup>28</sup>. While the shares of different protein classes in the cell's protein mass can vary, the total mass of proteins per unit of cell volume has been observed to be constant in the exponential growth phase<sup>73</sup>. Consequently, the overall rate of production of



all proteins must equal the total rate of protein mass dilution due to cell division. If  $M$  is defined as the mass of proteins (in amino acid residues) in the average volume of the cell over the cell cycle and

$$B = \frac{\sum_{j \in (a,r)} m_j/k_j}{1 + \frac{1}{1-\phi_q} \sum_{j \in (a,r)} m_j/k_j} R \quad (24)$$

is the total concentration of actively translating ribosomes (Supplementary Note S1.4), Supplementary Eq. (8) in Supplementary Note S1 shows that  $\lambda$  can be found as the quotient of the overall protein synthesis rate (i.e., abundance of translating ribosomes times the rate of translation elongation) and the cell's protein mass, which yields Eq. (25):

$$\lambda(\epsilon, B) = \frac{\epsilon B}{M} \quad (25)$$

**Translation elongation.** We now proceed to define the translation elongation rate  $\epsilon$ , which depends on  $t^c$ , the size of the pool of protein synthesis precursors (i.e., charged tRNAs)<sup>22</sup>. This relationship can be described by Michaelis-Menten kinetics<sup>23</sup> as shown in Eq. (26), where  $\epsilon_{\max}$  is the maximum possible translation elongation rate<sup>74</sup> and  $K_\epsilon$  is the half-saturation constant:

$$\epsilon(t^c) = \epsilon_{\max} \frac{t^c}{t^c + K_\epsilon} \quad (26)$$

**Protein precursor synthesis.** To replenish the protein precursors consumed during translation, the uncharged tRNAs ( $t^u$ ) are aminoacylated by metabolic proteins  $p_a$ . In line with our coarse-grained approach, we represent this process as a single reaction, whose rate per molecule of metabolic protein, which acts as the enzyme, is given by a Michaelis-Menten relation in Eq. (27)<sup>26</sup>. In this relation,  $t^u$  is the substrate,  $K_v$  is the corresponding half-saturation constant, and  $\nu_{\max}$  is the maximum tRNA aminoacylation rate per enzyme. The nutrients are assumed to be present in excess—hence their concentration having no effect on the reaction rate—and the aminoacyl-tRNA yield per nutrient molecule (alias the medium's nutrient quality) is captured by a constant factor  $0 \leq \sigma \leq 1$ . Notably, some steps of converting the nutrients into protein precursors may in fact be catalyzed by housekeeping proteins. However, these enzymes' concentrations are unchanging across all culture conditions, so the rates of the reactions catalyzed by them can be factored into the constant parameters  $\nu_{\max}$  and  $\sigma$ .

$$\nu(t^u, \sigma) = \nu_{\max} \sigma \frac{t^u}{t^u + K_v} \quad (27)$$

**Flux-parity resource allocation.** Finally, we consider how the cell allocates resources between the expression of different genes, which is captured by the ribosomal gene transcription regulation function  $F_r$ . We implement the near-optimal regulation of bacterial gene expression by the ppGpp signaling pathway according to the recently proposed Flux-Parity Regulation Theory. Yielding reliable predictions of bacterial behaviors in both steady-state and dynamic scenarios, it postulates that the cell equates and maximizes the protein synthesis and tRNA aminoacylation fluxes, which allows it to achieve maximum growth rate<sup>26</sup>. The cellular concentration of ppGpp ( $[ppGpp]$ ) therefore reflects the ratio of the aminoacylated and uncharged tRNA concentrations, although the exact mechanism enabling this is not yet entirely clear<sup>31,75</sup>. Consequently, instead of explicitly modeling the ppGpp concentration's dynamics, we describe it with a variable  $T$  that follows the biochemically motivated phenomenological relation<sup>26</sup>:

$$T = \frac{t^c}{t^u} \propto \frac{1}{[ppGpp]} \quad (28)$$

ppGpp represses ribosome synthesis<sup>30,76</sup>, which can be captured by the Hill kinetics outlined in Eq. (29), where  $\tau$  is the half-saturation constant:

$$F_r(T) = \frac{T}{T + \tau} \quad (29)$$

Since tRNA genes are co-regulated with the ribosomal genes<sup>77</sup>,  $F_r(T)$  is likewise included in Eq. (30) for the rate of tRNA transcription  $\psi$ . This value is calculated per unit of growth rate because tRNA transcription is enabled by the same  $\sigma^0$  factor as for most mRNAs<sup>78</sup>, whose availability we assume to be growth-dependent<sup>44</sup>:

$$\psi(T) = \psi_{\max} F_r(T) = \psi_{\max} \frac{T}{T + \tau} \quad (30)$$

## Numerical simulations

Our Matlab model implementation, along with all other scripts and data used to obtain the results described here, can be found at [https://github.com/KSechkar/rc\\_e\\_coli](https://github.com/KSechkar/rc_e_coli)<sup>63</sup>. The deterministic simulations were run using Matlab R2022a's `ode15s` solver on a Dell OptiPlex 7000 PC with a 2.10 GHz 12th Gen Intel(R) Core(TM) i7-12700 processor and 16 GB RAM, running on Windows 11.

To investigate the integral controller's stochastic performance, we used a hybrid tau-leaping simulation algorithm<sup>79</sup>. Our host cell variables are coarse-grained, representing the average dynamics of multiple variables whose fluctuations can be expected to average out, so their dynamics were captured by the hybrid model's deterministic component. Meanwhile, heterologous gene expression dynamics were described by the stochastic model component<sup>38</sup>. The Matlab implementation of this simulation algorithm—the code for which can be found together with the rest of our scripts in this manuscript's GitHub repository—was run on a PC with a 3.20 GHz Intel(R) Xeon(R) w5-2455X processor and 125 GB RAM, running on Ubuntu 20.04.1. Matlab R2023b's `ode15s` solver was used to simulate the model's deterministic component, while the stochastic component was simulated using a tau-leap algorithm with a time step of  $10^{-6}$  h. The details of our hybrid simulation method are provided in Supplementary Note S3.3.

All ODEs used in synthetic circuit simulations are provided in Supplementary Notes S3.1, S4.2.1, S4.3.1 and S4.4.1. Synthetic circuit parameters, displayed in Supplementary Tables 1, 4, 6, 7 and 9, were picked from the feasible ranges based on published literature<sup>80–83</sup> as outlined in Supplementary Note S4.1. Whenever a system's steady state was found by numerical simulation, this meant simulating the ODEs for 72 h by default, 48 h when calculating the likelihood function during parameter fitting (as a way of shortening the runtime of our code that needed to evaluate it many consecutive times), and 480 h when sweeping through different synthetic gene concentrations in Fig. 5 and Supplementary Fig. 4 due to the cell taking a long time to reach its steady state in presence of very high burden. The simulation was terminated prematurely if the changes in all variables over 12 h yielded less than  $10^{-6}$  when squared and summed together.

## Parameter fitting

Parameter fitting was used to estimate the maximum tRNA aminoacylation rate  $\nu_{\max}$  and the ribosome-chloramphenicol binding rate  $k_{cm}$  (required to model chloramphenicol's action on the cell as outlined in Supplementary Note S2.1), as well as  $K_\epsilon$  and  $K_v$ , the Michaelis constants determining the translation elongation and tRNA aminoacylation rates. While the metabolic and ribosomal genes' promoter strengths  $\alpha_a$  and  $\alpha_r$  also had to be determined this way, we found that the optimality of the fits remained approximately unchanged over a wide range of values as long as the  $\alpha_r$ :  $\alpha_a$  ratio was the same (Supplementary Fig. 2b). Therefore, we defined  $\alpha_a$  using an order-of-magnitude estimate and fitted this ratio's value to experimental data. As we outline in

Supplementary Note S2.3, this then allowed us to obtain more accurate promoter strengths by matching our model's predictions to experimentally measured RNA production rates<sup>84</sup>. Moreover, since the mutual maximization of tRNA charging and protein synthesis fluxes is known to be achieved when  $K_e \approx K_v^{26}$ , we assumed them to be equal, reducing the number of parameter values to be fitted down to only four.

Experimental datapoints for fitting were taken from the study by Scott et al.<sup>19</sup>, which measured steady-state growth rates and RNA:protein mass ratios of *E. coli* subjected to different concentrations of the translation-inhibiting drug chloramphenicol in various culture media. In order to convert RNA:protein mass ratios into ribosomal mass fractions, they were multiplied by a conversion factor of 0.4558<sup>26</sup>. Due to our model's lack of consideration of the cell's metabolic regulation mechanisms at very low growth rates, only the datapoints with  $\lambda > 0.3$  were used for parameter fitting.

Similarly to Weisse et al.<sup>28</sup>, who fitted their model's parameters to the same measurements and also used a constant scaling factor between 0 and 1 to quantify the yield of translation rate-defining precursor molecules per nutrient molecule, we estimated different media's nutrient qualities as six points equally spaced on the logarithmic scale between  $\sigma = 0.08$  and  $\sigma = 0.5$ . The effect of chloramphenicol on the cell was modeled by Supplementary Eqs. (59)–(64), derived and explained in Supplementary Note S2.1. Briefly, the ODEs were obtained similarly to the original cell model, first defining a model that explicitly considers all reactions involved in competitive ribosome binding then simplifying the translation dynamics. However, as chloramphenicol binds and disables translating ribosomes, extra terms were introduced to the definitions of apparent mRNA-ribosome dissociation constants  $\{k_i\}$ , as well as to the ODE for  $R$ , redefined as the concentration of the cell's operational ribosomes. As the cell's measured ribosome content also includes inactivated ribosomes, their levels had to be considered and were denoted as a new variable  $B_{cm}$ , whose ODE was added to the model (Supplementary Note S2.1).

Fitting was performed using the Differential Evolution Adaptive Metropolis (DREAM) algorithm, which is a variation of the Markov Chain Monte Carlo scheme for inferring parameters' probability distribution given a set of experimental observations. This involves following the trajectory of a Markov chain, whose states are sets of possible parameter values and whose stationary distribution is equal to the probability distribution of interest. Generally, such a chain is simulated by choosing a potential next set of parameter values according to some proposal distribution and then either accepting or rejecting it based on the probability of observing a known outcome if they are true. In the DREAM modification, the simulation's efficiency is increased by continuously adapting the proposal distribution based on the states considered in the past, as well as tracking several (in our case, 10) trajectories in parallel<sup>85</sup>.

Priors for all parameters were defined as normal distributions. The mean was assumed to be 1 for  $\alpha_r$ ;  $\alpha_a$  and taken from literature for the other parameters<sup>26,28,37</sup>. The variances were assumed to comprise a quarter of the mean. The admissible intervals for all parameters were set between  $1/50$  and 50 times the prior's mean (Supplementary Table 3). The DREAM simulation was run for 20,000 steps, using Matlab's Parallel Computing Toolbox v7.6 and the DREAM v2.4 Matlab package, part of the HydroSight v1.3.1 toolbox<sup>86</sup>. The mode of the fitted probability distributions provided the parameter values for our model. The parameters' starting values and other simulation details, as well as the details of parameter sensitivity analysis based on the fitting's outcome<sup>87,88</sup>, can be found in Supplementary Note S2.2. The parameter fitting script is provided in this manuscript's GitHub repository<sup>63</sup>.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

The data generated in this study are provided in the GitHub repository [https://github.com/KSechkar/rc\\_e\\_coli](https://github.com/KSechkar/rc_e_coli)<sup>63</sup> under <https://doi.org/10.5281/zenodo.10700011>. The data used for parameter fitting and in Fig. 3a and Supplementary Fig. 2a are from the study of ref. 19 and are available as Supplementary Material at <https://doi.org/10.1126/science.1192588>. The data used in Figs. 3b–d and 5c and Supplementary Fig. 1 are from the study of ref. 26 and are available in the GitHub repository [https://github.com/cremerlab/flux\\_parity](https://github.com/cremerlab/flux_parity)<sup>89</sup> under 10.5281/zenodo.5893799. Source data are provided with this paper.

## Code availability

All code used to implement the model and obtain the figures is provided in the GitHub repository [https://github.com/KSechkar/rc\\_e\\_coli](https://github.com/KSechkar/rc_e_coli)<sup>63</sup> under <https://doi.org/10.5281/zenodo.10700011>.

## References

- El Karoui, M., Hoyos-Flight, M. & Fletcher, L. Future trends in synthetic biology—a report. *Front. Bioeng. Biotechnol.* **7**, 175 (2019).
- Immethun, C. M., Kathol, M., Changa, T. & Saha, R. Synthetic biology tool development advances predictable gene expression in the metabolically versatile soil bacterium *Rhodospseudomonas palustris*. *Front. Bioeng. Biotechnol.* **10**, 3018 (2022).
- Perrino, G., Hadjimitsis, A., Ledesma-Amaro, R. & Stan, G.-B. Control engineering and synthetic biology: working in synergy for the analysis and control of microbial systems. *Curr. Opin. Microbiol.* **62**, 68–75 (2021).
- Ceroni, F. et al. Burden-driven feedback control of gene expression. *Nat. Methods* **15**, 387–393 (2018).
- Zhang, R. et al. Winner-takes-all resource competition redirects cascading cell fate transitions. *Nat. Commun.* **12**, 853 (2021).
- McBride, C. D. & Del Vecchio, D. Predicting composition of genetic circuits with resource competition: demand and sensitivity. *ACS Synth. Biol.* **10**, 3330–3342 (2021).
- McBride, C. D., Grunberg, T. W. & Del Vecchio, D. Design of genetic circuits that are robust to resource competition. *Curr. Opin. Syst. Biol.* **28**, 100357 (2021).
- Eguchi, Y. et al. Estimating the protein burden limit of yeast cells by measuring the expression limits of glycolytic proteins. *eLife* **7**, e34595 (2018).
- Frei, T. et al. Characterization and mitigation of gene expression burden in mammalian cells. *Nat. Commun.* **11**, 4641 (2020).
- Boo, A., Ellis, T. & Stan, G.-B. Host-aware synthetic biology. *Curr. Opin. Syst. Biol.* **14**, 66–72 (2019).
- Darlington, A. P. S., Kim, J., Jimenez, J. I. & Bates, D. G. Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes. *Nat. Commun.* **9**, 695 (2018).
- Qian, Y., Huang, H.-H., Jimenez, J. I. & Del Vecchio, D. Resource competition shapes the response of genetic circuits. *ACS Synth. Biol.* **6**, 1263–1272 (2017).
- Cella, F. et al. MIRELLA: a mathematical model explains the effect of microRNA-mediated synthetic genes regulation on intracellular resource allocation. *Nucleic Acids Res.* **51**, 3452–3464 (2023).
- Gyorgy, A. et al. Isocost lines describe the cellular economy of genetic circuits. *Biophys. J.* **109**, 639–646 (2015).
- Perrino, G. & Stan, G.-B. Robust set-point regulation of gene expression using resource competition couplings in mammalian cells. in *American Control Conference (ACC)*, 1373–1378. <https://doi.org/10.23919/ACC53348.2022.9867225> (2022).
- Melendez-Alvarez, J., He, C., Zhang, R., Kuang, Y. & Tian, X.-J. Emergent damped oscillation induced by nutrient-modulating growth feedback. *ACS Synth. Biol.* **10**, 1227–1236 (2021).
- Blanchard, A. E., Liao, C. & Lu, T. Circuit-host coupling induces multifaceted behavioral modulations of a gene switch. *Biophys. J.* **114**, 737–746 (2018).

18. Melendez-Alvarez, J. R., Zhang, R. & Tian, X.-J. Growth feedback confers cooperativity in resource-competing synthetic gene circuits. *Chaos Solit. Fractals* **173**, 113713 (2023).
19. Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).
20. Hui, S. et al. Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Mol. Syst. Biol.* **11**, 784 (2015).
21. Scott, M. & Hwa, T. Bacterial growth laws and their applications. *Curr. Opin. Biotechnol.* **22**, 559–565 (2011).
22. Scott, M., Klumpp, S., Mateescu, E. M. & Hwa, T. Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol. Syst. Biol.* **10**, 747 (2014).
23. Giordano, N., Mairet, F., Gouze, J.-L., Geiselmann, J. & De Jong, H. Dynamical allocation of cellular resources as an optimal control problem: novel insights into microbial growth strategies. *PLoS Comput. Biol.* **12**, e1004802 (2016).
24. Edwards, J. S., Ibarra, R. U. & Palsson, B. O. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* **19**, 125–130 (2001).
25. Molenaar, D., Van Berlo, R., De Ridder, D. & Teusink, B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5**, 323 (2009).
26. Chure, G. & Cremer, J. An optimal regulation of fluxes dictates microbial growth in and out of steady-state. *eLife* **12**, e84878 (2023).
27. Calabrese, L., Ciandrini, L. & Lagomarsino, M. C. How total mRNA influences cell growth. *bioRxiv*. <https://doi.org/10.1101/2023.03.17.533181> (2023).
28. Weiße, A. Y., Oyarzun, D. A., Danos, V. & Swain, P. S. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc. Natl Acad. Sci. USA* **112**, E1038–E1047 (2015).
29. Santos-Navarro, F. N., Vignoni, A., Boada, Y. & Pico, J. RBS and promoter strengths determine the cell-growth-dependent protein mass fractions and their optimal synthesis rates. *ACS Synth. Biol.* **10**, 3290–3303 (2021).
30. Potrykus, K., Murphy, H., Philippe, N. & Cashel, M. ppGpp is the major source of growth rate control in *E. coli*. *Environ. Microbiol.* **13**, 563–575 (2011).
31. Wu, C. et al. Cellular perception of growth rate and the mechanistic origin of bacterial growth law. *Proc. Natl Acad. Sci. USA* **119**, e2201585119 (2022).
32. Kremling, A., Geiselmann, J., Ropers, D. & de Jong, H. Understanding carbon catabolite repression in *Escherichia coli* using quantitative models. *Trends Microbiol.* **23**, 99–109 (2015).
33. Macklin, D. N. et al. Simultaneous cross-evaluation of heterogeneous *E. coli* datasets via mechanistic simulation. *Science* **369**, eaav3751 (2020).
34. Erickson, D. W. et al. A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. *Nature* **551**, 119–123 (2017).
35. Calabrese, L. et al. Protein degradation sets the fraction of active ribosomes at vanishing growth. *PLoS Comput. Biol.* **18**, 1–23 (2022).
36. Korem Kohanim, Y. et al. A bacterial growth law out of steady state. *Cell Rep.* **23**, 2891–2900 (2018).
37. Klumpp, S., Scott, M., Pedersen, S. & Hwa, T. Molecular crowding limits translation and cell growth. *Proc. Natl Acad. Sci. USA* **110**, 16754–16759 (2013).
38. Liao, C., Blanchard, A. E. & Lu, T. An integrative circuit-host modelling framework for predicting synthetic gene network behaviours. *Nat. Microbiol.* **2**, 1658–1666 (2017).
39. Lin, J. & Amir, A. Homeostasis of protein and mRNA concentrations in growing cells. *Nat. Commun.* **9** <https://doi.org/10.1038/s41467-018-06714-z> (2018).
40. Roy, A., Goberman, D. & Pugatch, R. A unifying autocatalytic network-based framework for bacterial growth laws. *Proc. Natl Acad. Sci. USA* **118**, e2107829118 (2021).
41. Doan, D. T., Hoang, M. D., Heins, A.-L. & Kremling, A. Applications of coarse-grained models in metabolic engineering. *Front. Mol. Biosci.* **9** <https://doi.org/10.3389/fmolb.2022.806213> (2022).
42. Dai, X. et al. Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. *Nat. Microbiol.* **2** <https://doi.org/10.1038/nmicrobiol.2016.231> (2016).
43. Moore, S. J. et al. EcoFlex: a multifunctional MoClo kit for *E. coli* synthetic biology. *ACS Synth. Biol.* **5**, 1059–1069 (2016).
44. Balakrishnan, R. et al. Principles of gene regulation quantitatively connect DNA to RNA and proteins in bacteria. *Science* **378**, eabk2066 (2022).
45. Hermesen, R., Erickson, D. W. & Hwa, T. Speed, sensitivity, and bistability in auto-activating signaling circuits. *PLoS Comput. Biol.* **7**, 1–9 (2011).
46. Tan, C., Marguet, P. & You, L. Emergent bistability by a growth-modulating positive feedback circuit. *Nat. Chem. Biol.* **5**, 842–848 (2009).
47. Tan, S.-I. & Ng, I.-S. New insight into plasmid-driven T7 RNA polymerase in *Escherichia coli* and use as a genetic amplifier for a biosensor. *ACS Synth. Biol.* **9**, 613–622 (2020).
48. De Cesare, I., Salzano, D., Di Bernardo, M., Renson, L. & Marucci, L. Control-based continuation: a new approach to prototype synthetic gene networks. *ACS Synth. Biol.* **11**, 2300–2313 (2022).
49. Jain, C. Degradation of mRNA in *Escherichia coli*. *IUBMB Life* **54**, 315–321 (2002).
50. Tsoi, R. et al. Metabolic division of labor in microbial systems. *Proc. Natl Acad. Sci. USA* **115**, 2526–2531 (2018).
51. Briat, C., Gupta, A. & Khammash, M. Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Syst.* **2**, 15–26 (2016).
52. Aoki, S. K. et al. A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature* **570**, 533–537 (2019).
53. Qian, Y. & Del Vecchio, D. Realizing ‘integral control’ in living cells: how to overcome leaky integration due to dilution? *J. R. Soc. Interface* **15**, 20170902 (2018).
54. Yoo, S. M., Na, D. & Lee, S. Y. Design and use of synthetic regulatory small RNAs to control gene expression in *Escherichia coli*. *Nat. Protoc.* **8**, 1694–1707 (2013).
55. Belliveau, N. M. et al. Fundamental limits on the rate of bacterial growth and their influence on proteomic composition. *Cell Syst.* **12**, 924–944.e2 (2021).
56. Ceroni, F., Algar, R., Stan, G.-B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12**, 415–418 (2015).
57. Thorsen, K., Drenth, T. & Ruoff, P. The effect of integral control in oscillatory and chaotic reaction kinetic networks. *Phys. D Nonlinear Phenom.* **393**, 38–46 (2019).
58. Ingram, D. & Stan, G.-B. Modelling genetic stability in engineered cell populations. *Nat. Commun.* **14** <https://doi.org/10.1038/s41467-023-39654-4> (2023).
59. Gutierrez Mena, J., Kumar, S. & Khammash, M. Dynamic cybergenetic control of bacterial co-culture composition via optogenetic feedback. *Nat. Commun.* **13** <https://doi.org/10.1038/s41467-022-32392-z> (2022).
60. Dinh, C. V., Chen, X. & Prather, K. L. J. Development of a quorum-sensing based circuit for control of coculture population composition in a Naringenin production system. *ACS Synth. Biol.* **9**, 590–597 (2020).
61. Huang, H.-H. et al. dCas9 regulator to neutralize competition in CRISPRi circuits. *Nat. Commun.* **12** <https://doi.org/10.1038/s41467-021-21772-6> (2021).



62. Barajas, C., Huang, H.-H., Gibson, J., Sandoval, L. & Del Vecchio, D. Feedforward growth rate control mitigates gene activation burden. *Nat. Commun.* **13** <https://doi.org/10.1038/s41467-022-34647-1> (2022).
63. Sechkar, K., Steel, H., Perrino, G. & Stan, G.-B. A coarse-grained bacterial cell model for resource-aware analysis and design of synthetic gene circuits GitHub. <https://doi.org/10.5281/zenodo.10700011> (2024).
64. Chandran, D., Copeland, W., Sleight, S. & Sauro, H. Mathematical modeling and synthetic biology. *Drug Discov. Today Dis. Models* **5**, 299–309 (2008).
65. Karpinets, T. V., Greenwood, D. J., Sams, C. E. & Ammons, J. T. RNA:protein ratio of the unicellular organism as a characteristic of phosphorous and nitrogen stoichiometry and of the cellular requirement of ribosomes for protein synthesis. *BMC Biol.* **4** <https://doi.org/10.1186/1741-7007-4-30> (2006).
66. Blumenthal, R. M., Lemaux, P. G., Neidhardt, F. C. & Dennis, P. P. The effects of the *relA* gene on the synthesis of aminoacyl-tRNA synthetases and other transcription and translation proteins in *Escherichia coli* B. *Mol. Gen. Genet.* **149**, 291–296 (1976).
67. Paget, M. S. & Helmann, J. D. The  $\sigma 70$  family of sigma factors. *Genome Biol.* **4**, 203 (2003).
68. Kennell, D. & Riezman, H. Transcription and translation initiation frequencies of the *Escherichia coli* lac operon. *J. Mol. Biol.* **114**, 1–21 (1977).
69. Nath, K. & Koch, A. L. Protein degradation in *Escherichia coli*. *J. Biol. Chem.* **245**, 2889–2900 (1970).
70. Pine, M. J. Steady-state measurement of the turnover of amino acid in the cellular proteins of growing *Escherichia coli*: existence of two kinetically distinct reactions. *J. Bacteriol.* **103**, 207–215 (1970).
71. Li, Z. & Deutscher, M. P. *Analyzing the Decay of Stable RNAs in E. coli* Ch. 2, 31–45 (Academic Press, 2008).
72. Prossliner, T., Agrawal, S., Heidemann, D. F., Sørensen, M. A. & Svenningsen, S. L. tRNAs are stable after all: pitfalls in quantification of tRNA from starved *Escherichia coli* cultures exposed by validation of RNA purification methods. *mBio* **14**, e02805-22 (2023).
73. Basan, M. et al. Inflating bacterial cells by increased protein synthesis. *Mol. Syst. Biol.* **11**, 836 (2015).
74. Milo, R. & Phillips, R. *Cell Biology by the Numbers* (Garland Science, 2015).
75. Potrykus, K. & Cashel, M. (p)ppGpp: still magical? *Annu. Rev. Microbiol.* **62**, 35–51 (2008).
76. Paul, B. J., Ross, W., Gaal, T. & Gourse, R. L. rRNA transcription in *Escherichia coli*. *Annu. Rev. Genet.* **38**, 749–770 (2004).
77. Jinks-Robertson, S., Gourse, R. L. & Nomura, M. Expression of rRNA and tRNA genes in *Escherichia coli*: evidence for feedback regulation by products of rRNA operons. *Cell* **33**, 865–876 (1983).
78. Miura, C. et al. Functional characterization of the principal sigma factor RpoD of phytoplasmias via an in vitro transcription assay. *Sci. Rep.* **5**, 11893 (2015).
79. Hepp, B., Gupta, A. & Khammash, M. Adaptive hybrid simulations for multiscale stochastic reaction networks. *J. Chem. Phys.* **142**, 034118 (2015).
80. Kelly, J. R. et al. Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.* **3**, 4 (2009).
81. Hirst, C. D. *Automated BioPart Characterisation for Synthetic Biology*. PhD Thesis, Imperial College London (2014).
82. Zucca, S. et al. Multi-faceted characterization of a novel LuxR-repressible promoter library for *Escherichia coli*. *PLoS ONE* **10**, 1–26 (2015).
83. Liu, X. et al. De novo design of programmable inducible promoters. *Nucleic Acids Res.* **47**, 10452–10463 (2019).
84. Bremer, H. & Dennis, P. P. Modulation of chemical composition and other parameters of the cell at different exponential growth rates. *EcoSal Plus* **3** <https://doi.org/10.1128/ecosal.5.2.3> (2008).
85. Vrugt, J. A. Markov chain Monte Carlo simulation using the DREAM software package: theory, concepts, and MATLAB implementation. *Environ. Model Softw.* **75**, 273–316 (2016).
86. Peterson, T. *HydroSight: Open-Source Data-driven Hydrogeological Insights* GitHub. <https://github.com/peterson-tim-j/HydroSight> (2022).
87. Gutenkunst, R. N. et al. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.* **3**, e189 (2007).
88. Komorowski, M., Costa, M. J., Rand, D. A. & Stumpf, M. P. H. Sensitivity, robustness, and identifiability in stochastic chemical kinetics models. *Proc. Natl Acad. Sci. USA* **108**, 8645–8650 (2011).
89. Chure, G. & Cremer, J. *An Optimal Regulation of Fluxes Dictates Microbial Growth In and Out of Steady-State* GitHub. <https://doi.org/10.5281/zenodo.5893799> (2023).

## Acknowledgements

This work was funded by the RAE Chair in Emerging Technologies program by the Royal Academy of Engineering [RAEng CiET 1819\5 to G.B.S.].

## Author contributions

G.B.S., G.P., and K.S. conceived the study. K.S. developed the model and performed mathematical derivations. K.S. and H.S. developed the model's Matlab implementation. K.S. wrote the manuscript. G.P., G.B.S., and H.S. reviewed and edited the paper. G.P. and G.B.S. co-supervised the project. G.B.S. secured all funding.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-46410-9>.

**Correspondence** and requests for materials should be addressed to Giansimone Perrino or Guy-Bart Stan.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024