Supplementary Information for:

Tunable genetic devices through simultaneous control of transcription and translation

Vittorio Bartoli, Grace A. Meaker, Mario di Bernardo and Thomas E. Gorochowski

Supplementary Note 1: Modelling the tunable expression system

To better understand the behavior of the tunable expression system (TES), we derived the following set of pseudo-chemical reaction equations to capture changes in the concentrations of the four main species in our system: toehold switch transcript *S*, tuner sRNA *T*, switchsRNA complex *C*, and output protein *P* (see **Figure 1a** for a schematic of the device)

$$
\frac{\alpha_{\tau}+\alpha_{\tau}^0}{\alpha_{\tau}+\alpha_{\tau}^0}T, \qquad (1)
$$
\n
$$
\frac{\alpha_{\tau}+\alpha_{\tau}^0}{\alpha_{\tau}+\alpha_{\tau}^0}S, \qquad (2)
$$
\n
$$
T+S\stackrel{k_{C+}}{\rightarrow}T+S, \qquad (3)
$$
\n
$$
C\stackrel{\alpha_{P}}{\rightarrow}C+P, \qquad (5)
$$
\n
$$
P\stackrel{\gamma_{P}}{\rightarrow}\emptyset, \qquad (6)
$$
\n
$$
T\stackrel{\gamma_{T}}{\rightarrow}\emptyset, \qquad (7)
$$
\n
$$
S\stackrel{\gamma_{S}}{\rightarrow}\emptyset, \qquad (8)
$$

$$
C \stackrel{\gamma_C}{\to} \emptyset \tag{9}
$$

Here, α_S and α_T are the induced promoter activities (rate of transcript production) of the main and tuner inputs, respectively, and α_P is the production rate of the output protein. α_S^0 and α_T^0 are the basal rates of transcription of the main and the tuner input promoters, respectively. k_{c+} and k_{c-} are the binding and unbinding rates of the toehold switch transcript with tuner sRNA. γ_s , γ_r , γ_c , γ_p are first-order degradation rates of the main toehold switch transcript, tuner sRNA, switch-sRNA complex, and output protein, respectively. Stochastic simulations of these chemical reaction equations were performed using COPASI (**Methods**) and a Systems Biology Markup Language (SBML) file capturing the model is available in **Supplementary Data 1**.

From these chemical reactions, we derived a set of ordinary differential equations (ODEs) to model the dynamics of the TES and assess how its function is affected by key design parameters. The model tracks changes in the concentrations of the main toehold switch transcript *S*, tuner sRNA *T*, switch-sRNA complex *C*, and output protein *P*. It consists of the following four coupled ODEs that capture production and loss of each species due to complex formation or degradation

$$
\frac{dS}{dt} = \alpha_S^0 + \alpha_S - k_{C+}ST + k_{C-}C - \gamma_S S \,, \tag{10}
$$

$$
\frac{dT}{dt} = \alpha_T^0 + \alpha_T - k_{C+}ST + k_{C-}C - \gamma_T T \,, \tag{11}
$$

$$
\frac{dC}{dt} = k_{c+}ST - k_{c-}C - \gamma_c C,\tag{12}
$$

All species and rates match those in the chemical reactions captured by **Equations (1)**–**(9)**. To analyze the behavior of this system, numerical simulations of these ODEs were performed using Python (**Methods**) across a range of biologically realistic parameter values (**Figure 3**; **Supplementary Table 1**).

Supplementary Note 2: Modelling possible retroactivity effects

 \overline{z}

 $\overline{1}$

One way that retroactivity [1, 2] could arise in the TES is due to the demand that synthesis of an output protein (e.g. YFP) places on the shared ribosome pool of a cell [3]. If expression of the output protein is sufficiently burdensome, there would be significant drop in the free ribosomes available to synthesize core machinery such as RNA polymerase (RNAP) and thereby a reduction in the transcription rates of the TES. To better understand this effect, we modified our existing model (**Supplementary Note 1**) to incorporate a factor *re* that captures the relative availability of host resources (i.e. RNAP)

$$
\frac{dS}{dt} = (\alpha_S^0 + \alpha_S)r_e - k_{c+}ST + k_{c-}C - \gamma_S S , \qquad (14)
$$

$$
\frac{dT}{dt} = (\alpha_T^0 + \alpha_T) r_e - k_{c+} ST + k_{c-} C - \gamma_T T \,, \tag{15}
$$

$$
\frac{ac}{dt} = k_{C+}ST - k_C_C - \gamma_C C. \tag{16}
$$

To calculate *re* we coupled these equations to a delay differential equation (DDE) model of ribosome allocation dynamics [3]. This model considers a fixed total pool of ribosomes *Rt* and their allocation between endogenous transcripts *E* (modelled as a single "average" type, see [3] for details), translationally active TES complexes *C*, and a free ribosome pool *Rf*. Conservation of the total number of ribosomes per cell is ensured by the additional algebraic constraint

$$
R_f = R_t - [R_E + R_C(t)],
$$
\n(17)

where R_E and R_C represent allocation of ribosomes to endogenous transcripts and active TES complexes, respectively, whose changes are captured by the following equations

$$
\frac{a_{E}}{dt} = \alpha_{E} R_{f}(t) E - \alpha_{E} R_{f}(t - \tau_{E}) E, \qquad (18)
$$

$$
\frac{dR_C}{dt} = \alpha_C R_f(t)C(t) - \alpha_C R_f(t - \tau_C)C(t - \tau_C).
$$
\n(19)

Here, α_E and α_C are the translation initiation rate of the endogenous transcripts and active TES complexes, respectively. Each equation captures the initiation of free ribosomes onto a respective transcript and the removal of ribosome from a transcript back into the free ribosome pool after the specific translation time of the transcript, either τ_E for endogenous transcripts or τ_c for active TES complexes. The relative availability of core cellular machinery like RNAP is then given by

$$
r_e = \frac{R_E}{R_0},\tag{20}
$$

where R_0 is the steady state value of R_E in the model when only endogenous genes are present (i.e. no TES complexes are considered, *C* = 0).

Analysis of the model was performed by numerical simulation of the DDEs above using Julia (**Methods**) across a range of biologically realistic parameter values (**Supplementary Table 1**). We then considered two cases: 1. when TES transcript synthesis rate was coupled

to ribosome availability (as above), and 2. when there was no coupling between the output protein production rate and the synthesis rate of the TES transcripts (i.e. r_e = 1). The simulations showed that retroactivity (case 1) could impact the response function of the TES device, but only when the output caused a significant burden on the cell and only for the most highly expressed outputs, i.e., when both the input and tuner promoter activities were high (**Supplementary Figure 2**).

Supplementary Figure 1: Characterization of sensor systems. (**a**) Response function of the P*tet* promoter to varying aTc concentrations: 0, 0.01, 0.04, 0.14, 0.5, 1, 2, 4, 8, 50 ng/μL. (**b**) Response function of the P*tac* promoter to varying IPTG concentrations: 0, 0.5, 5, 50, 100, 200, 450, 1000, 4000, 16000 mM. Points denote average output in relative promoter units (RPU) (**Methods**) from three biological replicates with error bars showing ±1 standard deviation. Solid grey lines show fitted Hill functions (P_{tet} fit parameters: $y_{min} = 0.05$ RPU, y_{max} = 6.77 RPU, *K* = 4.0 ng/μL aTc, *n* = 2.3; P*tac* fit parameters: *y*min = 0.001 RPU, *y*max = 2.73 RPU, *K* = 134 mM IPTG, *n* = 1.9).

Supplementary Figure 2: Simulated response functions for models. (**a**) Response functions when TES transcription is not coupled to varying availability of core cellular resources (r_e = 1). (b) Response functions when TES transcription is influenced by varying availability of core cellular resources due to synthesis of the output protein. Red arrow shows the decrease in the TES protein output due to retroactivity. All plots show individual curves for tuner promoter activities: 0.0001, 0.3, 1.5, 10, 190 transcripts/min (light grey to black). Model described in **Supplementary Note 2**.

Supplementary Figure 3: Plasmid maps. (**a**) pVB001: Tunable expression system. (**b**) pVB002: Tunable genetic NOT gate. (**c**) pVB003: Tunable expression system without RiboJ insulating element. (**d**) pVB004: Tunable genetic NOT gate without RiboJ insulating element. (**e**) pVB005: Tuner sRNA booster. (**f**) pVB006: Tunable NOR gate.

Supplementary Figure 4: Flow cytometry automated gating approach. (**a**) Raw event data from the flow cytometer. (**b**) Automatically gated event data with unwanted debris removed (**Methods**). Gating is performed using FlowCal which implements density-based filtering to discard events that deviate in their appearance from the main population. Top pseudo color density scatter plots show the forward and side scatter area (FSC-A and SSC-A) and bottom histogram shows the associated raw YFP fluorescence from these events. Data is collected for the initial TES construct with RiboJ present and an input of 6.6 RPU and tuner of 2.6 RPU.

Supplementary Table 1: Modelling parameters.

a. We assume that the induced translation initiation rate of P*tet* is similar to the transcription rate of an average constitutively expressed endogenous gene (20 transcripts min–1) [2] and that a plasmid with a p15A origin of replication has 15 copies per cell [3].

b. The induced transcription initiation rate of P*tet* is 1.57 times higher than P*tac* [1].

- c. We assume the pColE1 origin of replication leads to 50 copies per cell [7] and thus the transcription initiation rate from all P*tac* promoters is ~4 times higher when the sRNA booster is present.
- d. Uninduced basal expression calculated based on fold-change activation measured in previous work [1] and the estimated induced transcription initiation rates (α_S, α_T) .
- e. We assume the rates of DNA hybridization are similar to RNA hybridization [5].
- f. We assume the half-life of the switch transcript and tuner sRNA are similar to the average mRNA half-life in a cell, measured to be 3 min in exponentially growing cells [6].
- g. Dilution by growth is the dominant mode of degradation. The doubling time of cells during these experiments was determined to be ~20 min.

Supplementary Table 2: Sequences used for thermodynamic modeling.

	Dynamic range ^a (a.u.)			Fold-change ^a			Intersection ^a		
Tuner	Ara $/$ aTc			Ara / aTc			Ara $/$ aTc		
(RPU)		$-1+$	$+$ / +	$+$ / $\overline{}$	$-1+$	$+$ / $+$	$+$ $\overline{ }$ $\overline{}$	$-1+$	$+$ / $+$
0.002	$3210 +$	$2623 \pm$	$3671 \pm$	$5.6 \pm$	$4.3 \pm$	16.7 \pm	$0.21 +$	$0.13 \pm$	$0.04 \pm$
	573	322	233	3.9	2.6	3.3	0.12	0.07	0.003
2.6	$3808 \pm$	$4013 \pm$	$4012 \pm$	$8.4 \pm$	$15.3 \pm$	$13.9 +$	$0.07 \pm$	$0.08 \pm$	$0.05 \pm$
	604	718	729	0.1	4.5	4.5	0.04	0.02	0.02

Supplementary Table 3: Performance summary of tunable NOR gate.

a. Measurements are a comparison to experiments when both input inducers (Ara and aTc) are absent. Average values are shown ± 1 standard deviation calculated from flow cytometry data for three biological replicates.

Supplementary Table 4: Synthesized DNA fragments used in this study.

Supplementary Table 5: Primers used in this study.

a. Plasmid names prefixed with pAN are from Neilsen et al. [11].

b. Underlined lowercase sequences denote homology regions added for Gibson assembly.

Supplementary References

- [1] Del Vecchio D., Ninfa A.J. & Sontag E.D. (2008) Modular cell biology: retroactivity and indulation. *Molecular Systems Biology* **4**, 161.
- [2] Del Vecchio D. & Sontag E.D. (2009) Engineering Principles in Bio-molecular Systems: From Retroactivity to Modularity. *European Journal of Control* **3**, 389–397.
- [3] Gorochowski T.E., Avcilar-Kucukgoze I., Bovenberg R.A.L., Roubos J.A. & Ignatova Z. (2016) A minimal model of ribosome allocation dynamics captures trade-offs in expression between endogenous and synthetic genes. *ACS Synthetic Biology* **5**, 710– 720.
- [4] Stanton B.C., Nielsen A.A.K., Tamsir A., Clancy K., Peterson T. & Voigt C.A. (2014) Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nature Chemical Biology* **10**, 99–105.
- [5] Kennell D. & Riezman H. (1977) Transcription and translation initiation frequencies of the *Escherichia coli lac* operon. *Journal of Molecular Biology* **114**, 1–21.
- [6] Hiszczyńska-Sawicka E. & Kur J. (1997) Effect of *Escherichia coli* IHF Mutations on Plasmid p15A Copy Number. *Plasmid* **38**, 174–179.
- [7] Pai A. & You L. (2009) Optimal tuning of bacterial sensing potential. *Molecular Systems Biology* **5**, 286.
- [8] Xu S., Zhan J., Man, B., Jiang, S. Yue W., Gao S., Guo C., Liu H., Li Z., Wang J. & Zhou Y. (2017) Real-time reliable determination of binding kinetics of DNA hybridization using a multi-channel graphene biosensor. *Nature Communications* **8**, 14902.
- [9] Bernstein J.A., Khodursky A.B., Lin P.-H., Lin-Chao S. & Cohen S.N. (2002) Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proceedings of the National Academy of Sciences USA* **99**, 9697–9702.
- [10] Liang S.-T., Ehrenberg M., Dennis P. & Bremer H. (1999) Decay of *rplN* and *lacZ* mRNA in *Escherichia coli*. *Journal of Molecular Biology* **288**, 521–538.
- [11] Nielsen, A. A. K. *et al.* (2016) Genetic circuit design automation. *Science* **352**, aac7341.