

Supplementary Information for:

Harnessing the central dogma for stringent multi-level control of gene expression

F. Veronica Greco, Amir Pandi, Tobias J. Erb, Claire S. Grierson and Thomas E. Gorochoowski

Supplementary Notes	Page
Supplementary Note 1: Derivation of mathematical models	2
Supplementary Note 2: MLC toolkit user guide	3
 Supplementary Figures	
Supplementary Figure 1: Key features of a controller's response function	5
Supplementary Figure 2: Growth curves	6
Supplementary Figure 3: STAR ₂ performance comparison	7
Supplementary Figure 4: Plasmid maps for the MLC toolkit	8
Supplementary Figure 5: Plasmid maps for the gene expression controllers used	9
Supplementary Figure 6: Flow cytometry automated gating approach	10
 Supplementary Tables	
Supplementary Table 1: Model parameters	11
Supplementary Table 2: Genetic parts used in this study	12
Supplementary Table 3: Performance of controllers in a cell-free expression system	14
 Supplementary References	15

Supplementary Note 1: Derivation of mathematical models

To explore how single- and multi-level regulation affected the output protein production rate as a function of an input inducer chemical concentration, we derived mathematical models for each type of system. For the single-level controller we defined equations to track the concentration of GOI transcripts R and protein P as:

$$\frac{dR}{dt} = \alpha_I - \gamma_R R, \quad (1)$$

$$\frac{dP}{dt} = \alpha_P R - \gamma_P P. \quad (2)$$

Here, α_I is the production rate of GOI transcripts, α_P is the production rate proteins per transcript, and γ_R and γ_P are first-order degradation rates of the transcripts and output proteins, respectively. In this system, α_I is directly related to the activity of the P_{L1} promoter (**Figure 1C**). To capture a realistic response of a promoter for a small molecule sensor (e.g. the P_{tac} system) that we might use as input to the controller, we made use of empirically measured steady-state response functions that fit to a Hill equation such that

$$\alpha_I = f(x) = y_{\min} + (y_{\max} - y_{\min}) \frac{x^n}{K^n + x^n}. \quad (3)$$

Here, x is the concentration of the inducer (e.g. a small molecule), $y_{\min}$ and $y_{\max}$ are the minimum and maximum activities of the P_{L1} promoter, K is the inducer concentration where the P_{L1} promoter activity is half its maximum, and n is the cooperativity.

For the MLC, we extended this model to include additional states for the concentrations of the $L2$ regulator S , and the complex C of the GOI transcript with the $L2$ regulator that forms through non-cooperative RNA-RNA hybridisation. By assuming that only these complexes can be translated into output protein, we end up with the following set of equations:

$$\frac{dR}{dt} = \alpha_I - k_{C+}RS + k_{C-}C - \gamma_R R, \quad (4)$$

$$\frac{dS}{dt} = \alpha_I - k_{C+}RS + k_{C-}C - \gamma_S S, \quad (5)$$

$$\frac{dC}{dt} = k_{C+}RS - k_{C-}C - \gamma_C C, \quad (6)$$

$$\frac{dP}{dt} = \alpha_P C - \gamma_P P, \quad (7)$$

where k_{C+} and k_{C-} are the binding and unbinding rates of the GOI transcript with the $L2$ RNA-based regulator.

To analyse the behaviour of these systems, numerical simulation of these ODEs was performed using the `odeint` function of the `SciPy.integrate` Python package and biologically realistic parameters assuming the input sensor behaved similarly to the P_{tac} system (**Supplementary Table 1**).

Supplementary Note 2: MLC toolkit user guide

The MLC toolkit is composed of a set of plasmids and assembly rules to create controllers in a combinatorial manner. It includes a backbone plasmid called pMLC-BB1 that is used to hold newly built controllers and a set of 8 types of part plasmid (pA–pH) corresponding to the individual blocks that make up a controller (**Figure 2A**). pMLC-BB1 is based on the pET28a(+) plasmid and contains a pBR322 origin of replication, kanamycin resistance marker and expression cassette for an orange fluorescent protein (*ofp*) gene that is highly visible in transformed cells. The *ofp* expression cassette is flanked by BsaI sites such that it is removed before insertion of a fully assembled MLC. To minimise controller malfunctions due to individual part failure, the orientation of each transcriptional unit has been chosen such that potential transcriptional read-through at any terminator does not cause unwanted expression of other controller elements. The toolkit has also been designed to ensure the accurate and rapid assembly of new designs by minimising the chance of incorrect assembly through the selection of 4 bp overhangs with low cross reactivity and the use of a fluorescent protein drop-out upon successful insertion of a controller construct into the backbone to reduce the subsequent screening/sequencing of transformants needed to verify a perfect assembly.

Creating new parts

Parts for each block of a controller can be created in the form of pre-cloned plasmids, amplicons or even annealed complementary oligos if the length is short. Pre-cloned plasmids are the preferred method as this reduces the chance of part mutations. Amplicons are easy to generate by PCR and allow for the easy addition of 5' and 3' sequences needed for assembly, however, also tend to see higher numbers of mutations. Complementary oligos annealed to create a double stranded DNA part are ideal for smaller sequences up to 60 nt, although again using this method will lead to the highest error rates in the final parts produced. Every new part must be flanked by a standardised sequence that is used to create the necessary single stranded overhangs that are used for correct assembly. These sequences contain a BsaI restriction site and a specific 4 bp overhang. The following table provides details of the necessary 5' and 3' sequences that are needed.

Part block	Sequence to add at 5' end	Sequence to add at 3' end
A	NNNNNNGGTCTCAgtga	ggagTGAGACCNNNNNN
B	NNNNNNGGTCTCaggag	aggtTGAGACCNNNNNN
C	NNNNNNGGTCTCAaggt	gcttTGAGACCNNNNNN
D	NNNNNNGGTCTCagctt	cgctTGAGACCNNNNNN
E	NNNNNNGGTCTCacgct	ccgcCGAGACCNNNNNN
F	NNNNNNGGTCTCaccgc	cataCGAGACCNNNNNN

G	NNNNNNGGTCTC <u>ac</u> atc	tactTGAGACCNNNNNN
H	NNNNNNGGTCTC <u>a</u> tact	tgccTGAGACCNNNNNN

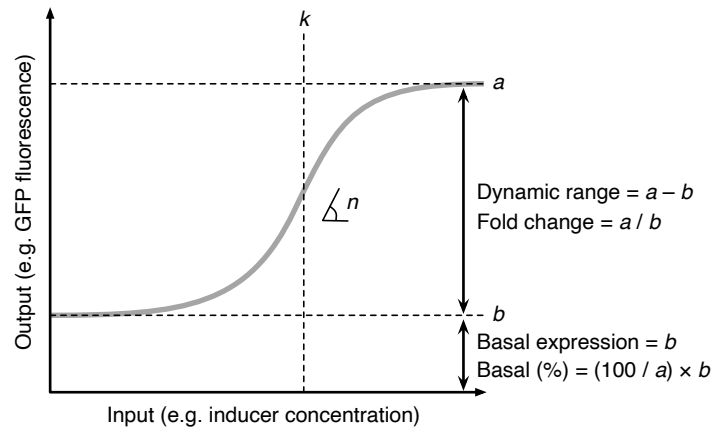
Underlined sequences denote BsaI recognition sites used for Golden Gate assembly. Bold lowercase sequences denote 4 bp overhangs generated after digestion by BsaI which are used for assembly. The six undefined bases at the 5' and 3' ends are optional but help improve digestion efficiency.

Assembling new controllers

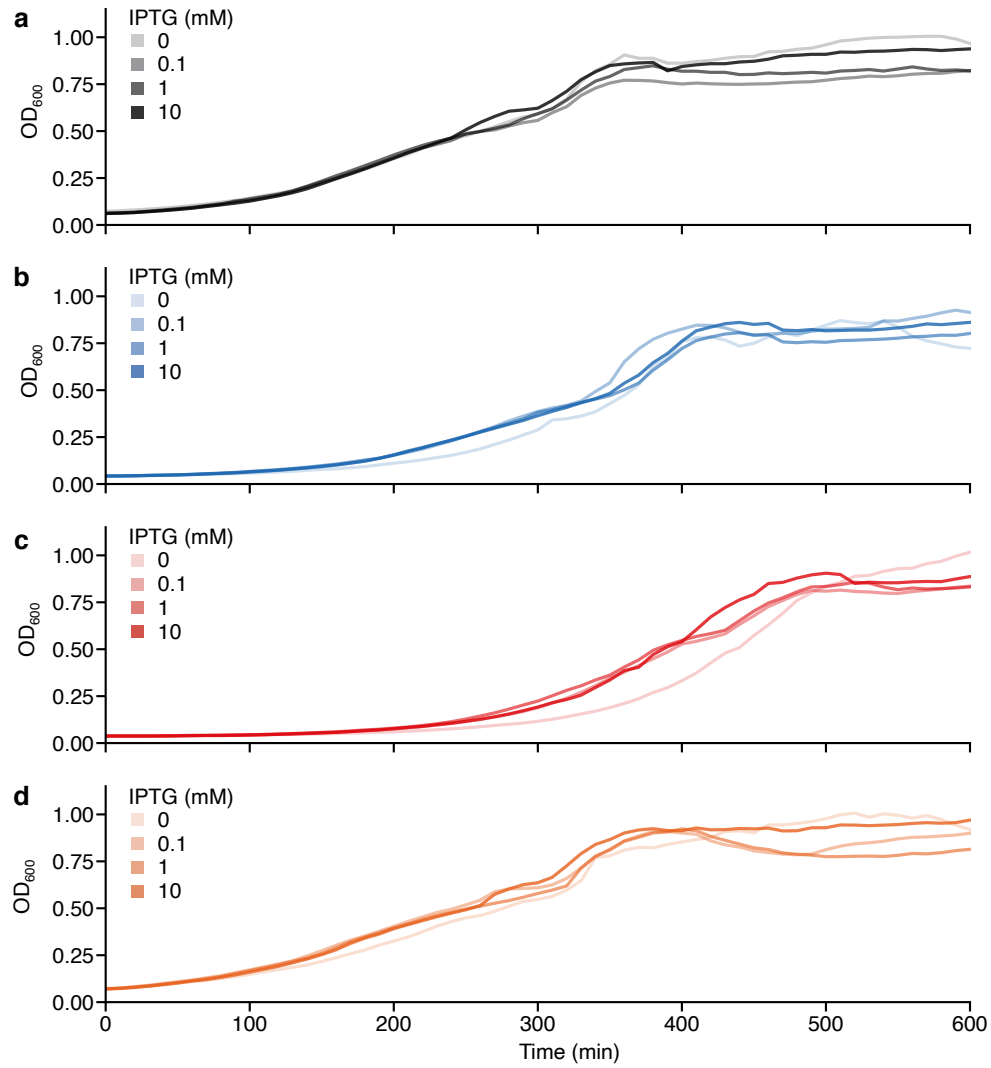
Assembly of a controller is performed using a standard Golden Gate reaction. We recommend the use of the NEB Golden Gate Assembly Kit (BsaI-HFv2, E1601) and have found that assemblies can be scaled down using this kit to a quarter of their normal volume (5 μ L) with still large numbers of successful transformants produced for chemically competent *E. coli* DH10- β cells using 2 μ l of the reaction mix. Each reaction should include the backbone plasmid and a single plasmid type for each of the blocks in the controller (**Figure 2A**). Although we rarely see mutations in the part sequences, there is a tendency for mutations to occur within the 4 bp overhangs used for assembly. Therefore, it is recommended that all controllers are fully sequence verified before use.

Assembling single-level controllers using the MLC toolkit

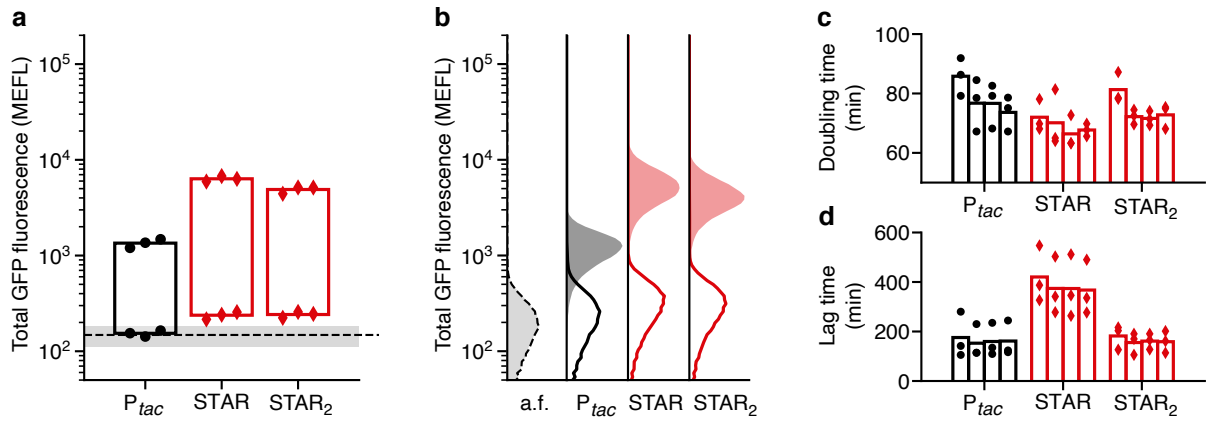
Although the toolkit is tailored to produce MLCs, it is possible to create SLCs by substituting several blocks with a specially designed spacer. Specifically, blocks B and C can be replaced by a single spacer element ('BC-spacer', **Supplementary Table 2**) that is able to assemble with block A at its 5'-end and block D at its 3'-end. As mutations in this part are less problematic, we have used annealed oligos to produce this part for creating the P_{tac} SLC. In addition, the removal of the L2 regulator also requires the use of a standard RBS part for block G. We include a strong RBS part in the toolkit for this use ('RBS', **Supplementary Table 2**).



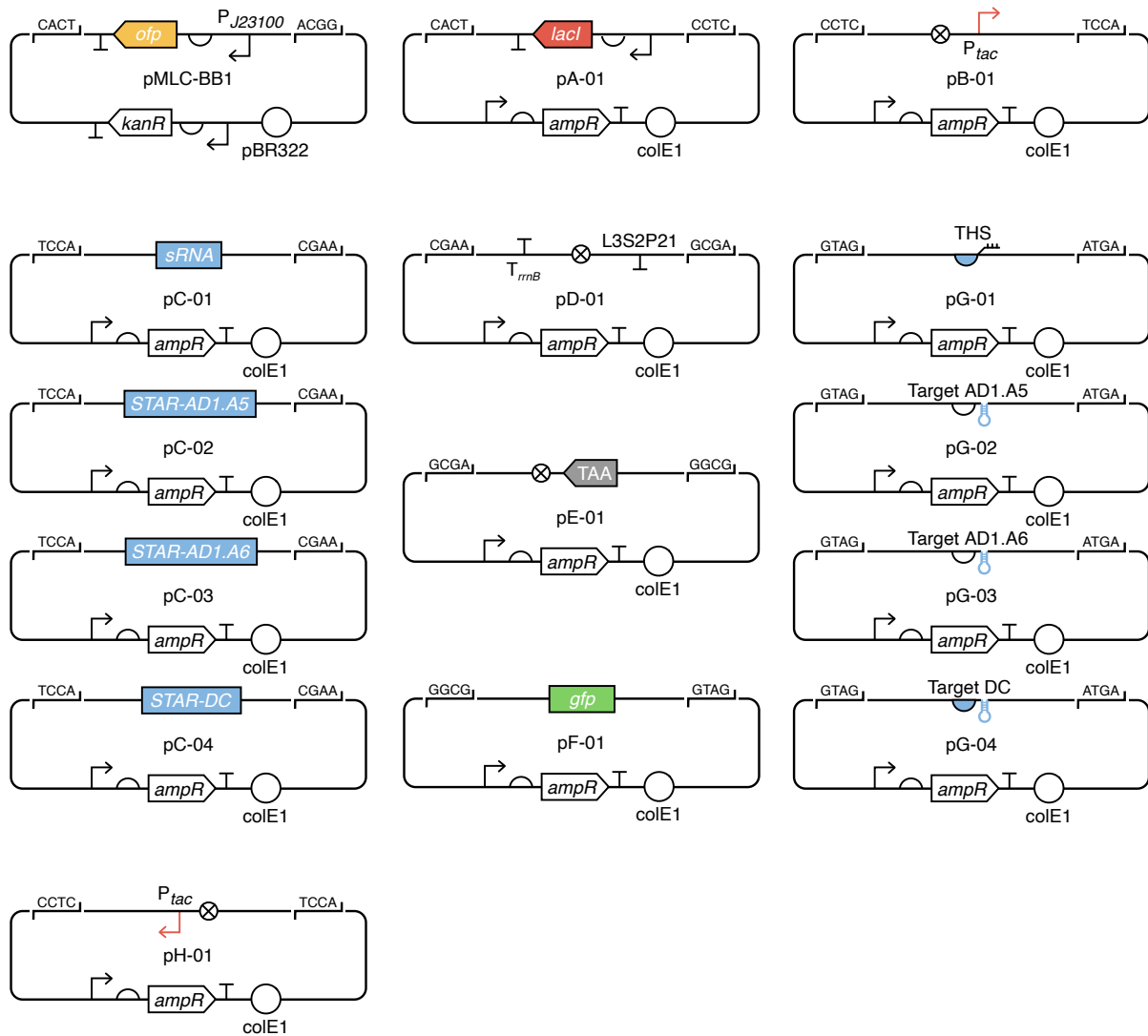
Supplementary Figure 1: Key features of a controller's response function. Typical response function for an inducible system shown by thick grey line. The parameters a , b , k , and n can be found by fitting to experimental data that measures the controller output across a range of inputs x to a Hill function of the form $f(x) = b + (a - b) \left[\frac{x^n}{k^n + x^n} \right]$. Parameter a corresponds to the maximum output, b to the basal output when no input is present, k is the input where half the maximal output minus basal output is achieved, and n is the cooperativity (that is manifested as the steepness of the response function).



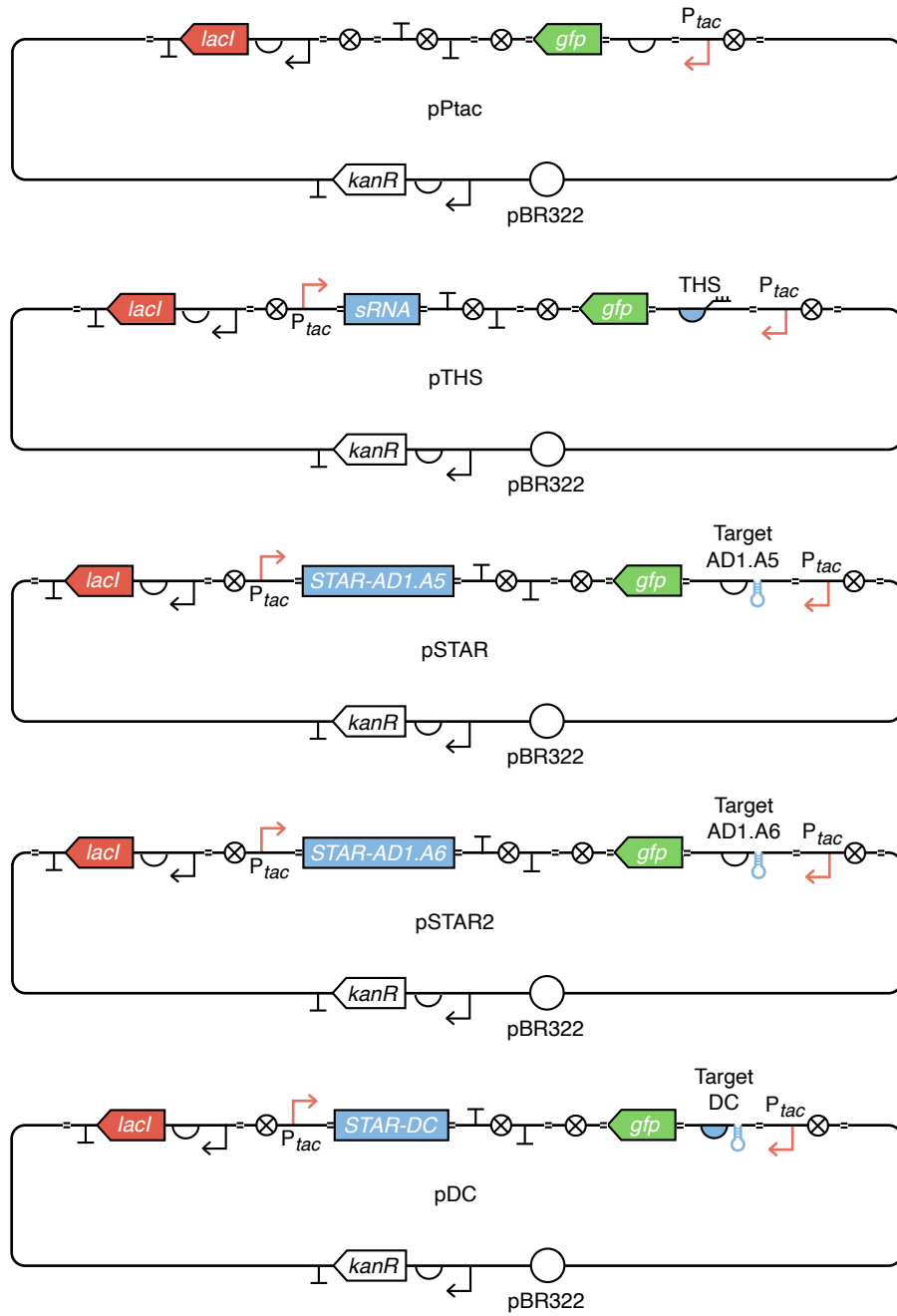
Supplementary Figure 2: Growth curves. Time course measurements of the optical density at 600 nm (OD₆₀₀) of *E. coli* DH10-β cells harboring (a) pPtac, (b) pTHS, (c) pSTAR and (d) pDC plasmids in the presence of different concentrations of IPTG. Source data are provided as a Source Data file.



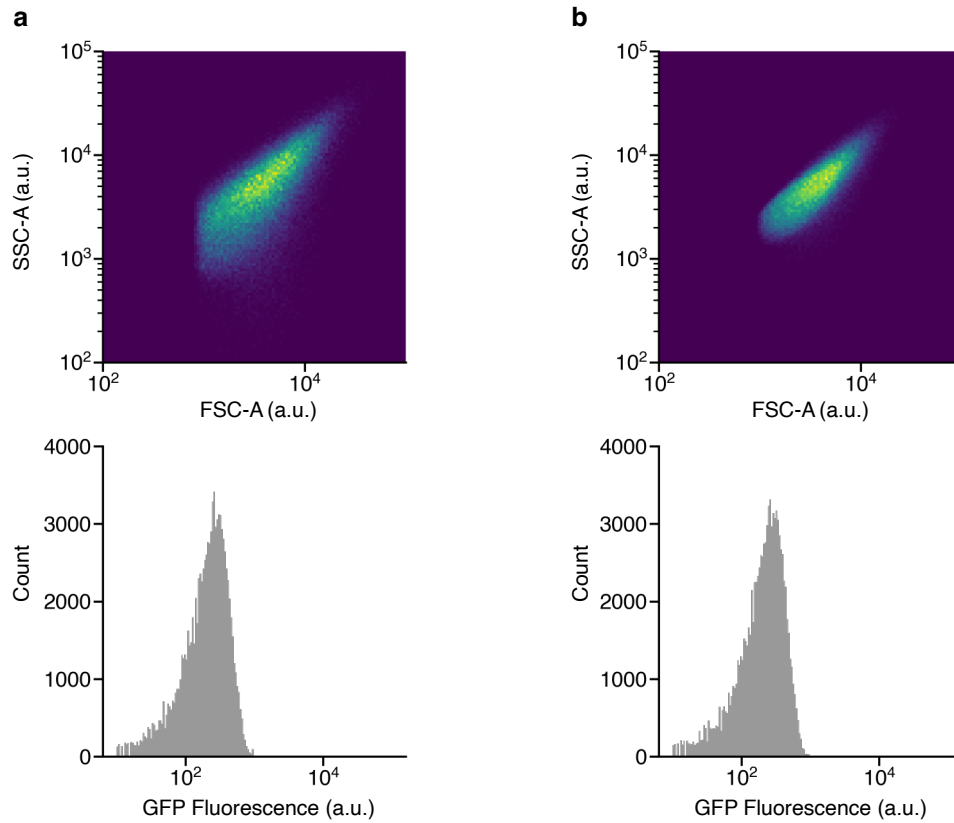
Supplementary Figure 3: STAR₂ performance comparison. Data for the P_{tac} and STAR designs shown for comparison. **(a)** Total GFP fluorescence for 'off' and 'on' input states (0 and 1 mM IPTG, respectively). Points show the three biological replicates for each controller and condition (black circles, P_{tac} ; red diamonds, STAR, STAR₂). Black dashed line denotes the mean fluorescence of cell autofluorescence (a.f.) controls containing no plasmid with grey shared region showing ± 1 standard deviation of 11 biological replicates. Fluorescence given in calibrated molecules of equivalent fluorescein (MEFL) units. **(b)** Flow cytometry distributions of total GFP fluorescence for 'off' (line) and 'on' (shaded) input states. Cell autofluorescence (a.f.) controls containing no controller are shown by black dashed line and light grey filled distributions. **(c)** Doubling time of cells harbouring single- and multi-level controllers for varying concentrations of IPTG (bars left to right for each design: 0, 0.1, 1, 10 mM IPTG). **(d)** Lag time calculated as the time to reach an $OD_{600} = 0.15$ after inoculation of cells harbouring controllers for varying concentrations of IPTG (bars left to right for each design: 0, 0.1, 1, 10 mM IPTG). Source data are provided as a Source Data file.



Supplementary Figure 4: Plasmid maps for the MLC toolkit. Genetic diagrams are shown using SBOL Visual symbols. BsaI restriction sites flanking each part are shown with their 4 bp overhang sequence after digestion that is used for assembly. The *gfp* part in plasmid pF-01 is in a reverse orientation and lacks a stop codon. The *gfp* open reading frame is completed after assembly with plasmid pE-01 which includes a TAA stop codon.



Supplementary Figure 5: Plasmid maps for the gene expression controllers used. Genetic diagrams are shown using SBOL Visual symbols and include the scars introduced between core elements of the genetic template.



Supplementary Figure 6: 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 GFP fluorescence from these events. Data is shown for *E. coli* cells harboring the pTHS construct with no inducer present.

Supplementary Table 1: Model parameters

Parameter	Description	Value(s)	Units	Refs.
α_I	Transcription rate of input promoter	0.04–100	transcript min ⁻¹	[1,2]
α_P	Translation rate of sRNA:transcript complex	5	protein complex ⁻¹ min ⁻¹	[3]
k_{C+}	sRNA:transcript complex association rate	0.0257	complex transcript ⁻¹ min ⁻¹	[4]
k_{C-}	sRNA:transcript complex disassociation rate	0.0067	transcript complex ⁻¹ min ⁻¹	[4]
γ_R	Transcript degradation rate	0.231	min ⁻¹	[5]
γ_S	sRNA degradation rate	0.231	min ⁻¹	[5]
γ_C	sRNA:transcript complex degradation rate	0.231	min ⁻¹	[5]
γ_P	Output protein degradation rate	0.035	min ⁻¹	[6]
$y_{\min}$	Minimum input promoter activity	0.04	transcript min ⁻¹	[1,2]
$y_{\max}$	Maximum input promoter activity	100	transcript min ⁻¹	[1,2]
K	Concentration of input inducer at which sensor output is half-maximal	134	mM	[1]
n	Cooperativity of input sensor system	1.9	–	[1]

Supplementary Table 2: Genetic parts used in this study.

Plasmid	Part	Sequence ^a	Ref.
pA1	<i>lacI</i> -TU	<u>GGTCTCAgtgaAGTGAAGAAAAAGGCGCGAGCGGCCTTTTAGTTAGATCGGAC</u> <u>CCCACATCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGTCGATTAATGA</u> <u>ATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCTATTGGGCGCCAGGGTGGTTTTCTTT</u> <u>TCACCAGTGAGACTGGCAACAGCTGATTGCCCTTACCAGCTGGCCCTGAGAGAGTTGC</u> <u>AGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAAATCCTGTTTATGGTGGTTAA</u> <u>CGGCGGGATATAACATGAGCTATCTTCGGTATCGTCGTATCCCACTACCGAGATATCCG</u> <u>CACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCAGCGCCATCTGATCG</u> <u>TTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTAGCATTGTCATGGTTTGTG</u> <u>AAAACCGGACATGGCACTCCAGTCGCCTTCCCCTTCCGCTATCGGCTGAATTTGATTGC</u> <u>GAGTGAGATATTTATGCCAGCCAGCCAGACGCGAGACGCGCCGAGACAGAACTTAATGGG</u> <u>CCCGCTAACAGCGCGATTGCTGGTGACCCAAATGCGACCAGATGCTCCACGCCAGTCG</u> <u>CGTACCGTCCCTCATGGGAGTAAATAACTGTTGATGGGTGTCTGGTCAGAGACATCAA</u> <u>GAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCAGCTCTGTCATCC</u> <u>AGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTCACCGCCGC</u> <u>TTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCGCTGGCACCAGTT</u> <u>GATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGGTGTCAGGGCCAGACTG</u> <u>GAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGT</u> <u>GGGAATGTAATTCAGCTCCACCATCGCCGCTTCCACTTTTCCCGCGTTTTCGCAGAAA</u> <u>CGTGGCTGGCCTGGTTTACCACGCGGGAACGGTCATATAAGAGACACCGGCATACCT</u> <u>GCGACATCGTATAACGTTACTGGTTTCATATTACCTCCATGAATGACTCTCTTCCGG</u> <u>GCGCTATCATGCCATACCGCGAAAGGTTTTTACCATTGATGGCGCGCCGCGCTGGG</u> <u>AATCTGAGACATGAGTCAAGATATTTGCTCGGTAACGTATGCTggagTGAGACC</u>	[7]
pB1	<i>P_{tac}</i>	<u>GGTCTCAggagTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTTGTGAGCGC</u> <u>TCACAATTaggTGAGACC</u>	[7]
pC1	STAR- pT1810-DC	<u>GGTCTCAaggTAACAAAAATAAGCAATAAGGAATCGCTCACCCAAAGGATCTgcttTGA</u> <u>GACC</u>	[8]
pC2	STAR- AD1.A5	<u>GGTCTCAaggTGAAGTGTATACATTCCCCGCTGCTCCAACATTTATACAACATAATTAA</u> <u>AACAATTCAGTGTAAAACTgcttTGAGACC</u>	[9]
pC3	STAR- AD1.A6	<u>GGTCTCAaggTGAAGTGTATACATTCCCCGCTGCTCCAACATTTATACAACATAATTAA</u> <u>AACAATTCAGTGTAAAACTTTCTTAGACgcttTGAGACC</u>	[9]
pC4	sRNA-THS	<u>GGTCTCAaggTGGGACCTATTGGACCCGTTTCCAATAGGTGAACAAGACGATAGAACAA</u> <u>GCATTTGCACTTATAGAgctTGAGACC</u>	[10]
pD1	TrmB- L3S2P21	<u>GGTCTCAgcttATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCTGTTTTATCTGT</u> <u>TGTTTGTGCGTGAACGCTCTCTGAGTAGGACAAATCCGCGGGAGCGGATTTGAACGT</u> <u>TGCGAAGCAACGGCCCGAGGGTGGCGGGCAGGACGCCCGCCATAAATGCCAGGCATC</u> <u>AAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTGGCCCCGACCCTTAGACTCTG</u> <u>TACTCAGGGACCAAAACGAAAAAGGCCCCCTTTCGGGAGGCCTCTTTCTGGAATTT</u> <u>GGTACCGAGcgtTGAGACC</u>	[11]
pE1	Spacer-TAA	<u>GGTCTCAcgcTAgTTGTATCGTACGTCGGTCTAGGAATCGTTTAGccgcCGAGACC</u>	—
pF1	<i>gfp</i>	<u>GGTCTCAccgcCTCACTTATACAACATCATCCATGCCTAAAGTAATCCCCGCCGCGGTTA</u> <u>CGAACTCCAAAAGGACCATGTGATCACGTTTCTCATTGGGTCTTCGATAATTTGGAT</u> <u>TGTGTACTCAGGTAGTGATTGTCAGGAAGTAAACCGGGCCGTCGCCAATGGGTGTGTT</u> <u>TTGTTGATAATGGTCAGCCAATTGCACGCTGCCATCCTCAATGTTGTGACGAATCTTGA</u> <u>AATTTACCTTAATCCGTTCTTCTGCTTGTGAGCCATGATATACAGTTGTGCGAATTA</u> <u>TAGTTTACTCCAGTTTATGTCCCAAAATGTTGCCATCTCTTTAAAGTCAATGCCTTT</u> <u>CAATTCAATGCGATTCACTAACGTATCCCCCTCAAACCTTGACCTCGCGCGGTGTCTTAT</u> <u>AGTTCCCGTCATCCTTGAAGAAGATTGTGCGCTCTTGACATAGCCTTCTGGCATTGCA</u> <u>GATTTAAAGAAGTCGTGCTGCTTCATGTGATCTGGGTAACGCGAGAAACATTGGACCCC</u> <u>GTATGTAAGTGTGTAAACAGAGTTGGCCAGGGGACTGGAAGTTTTCCTGTAGTACAAA</u> <u>TGAACCTAAGCGTCAATTTGCCGTAAGTAGCGTCGCCCTTCGCCCTCGCCGCTAACCGAA</u> <u>AACTTATGACCATTAAAGTCCCATCCAACCTCAACCAAGATGGGTACAACGCTGTAAA</u> <u>CAGCTCCTCTCCTTTACTcactTGAGACC</u>	—
pG1	Target- pT181- DC+RBS	<u>GGTCTCAcactcAAAAAATCGACTCCTTAATCTCAATTTCTGTTTAAAGGAATCGCTCACCC</u> <u>AtactTGAGACC</u>	[8]
pG2	Target- AD1.S5+RBS	<u>GGTCTCAcactcAGATCCTTCTCCTAGATCCAAAAAAGCGGGGAATATATACATGA</u> <u>ACTGTATACATTCCCCGCTGCTCCAACATTTATACAACATAATTAACAATTCAGTGA</u> <u>AAAACCTactTGAGACC</u>	[9]

pG3	Target-AD1.S6+RBS	<u>GGTCTC</u> acatc AGATCCTTCCTCCTAGATCCAAAAAAGCGGGGAATATATACATGA ACTGTATACATTCCCCGCTGCTCCAACATTTATACAATAATTAAACAATTCACCTGTA AAAACTTTTCCTAGACT tact <u>TGAGACC</u>	[9]
pG4	Target-THS	<u>GGTCTC</u> acatc ACGCTTGTTCCTCATGTCTTGTCTCCTCTGTTTCAAGACGATAGAACAA GCATTTGCACTTAT tact TGAGACC	[10]
pH1	P _{tac}	GGTCTC tact AATTGTGAGCGCTCACAATTCACACATTATACGAGCCGATGATTAAT TGTCACAT gccc TGAGACC	[7]
— ^b	RBS	Fwd: catc AGATCCTTCCTCCTAGATCCGCATCCGGGC Rev: agta GCCCGGATGCGGATCTAGGAGGAAGGATCT	—
— ^b	BC-spacer ^c	Fwd: ggag CTCGTTCATCCCGTGGGACATCAAGCTTCGCCTTGATAAA Rev: aagc TTTATCAAGGCGAAGCTTGATGTCCACGGGATGAACGAG	—

- Underlined sequences denote BsaI recognition sites used for Golden Gate assembly. Bold lowercase sequences denote 4 bp overhangs generated after digestion by BsaI which are used for assembly.
- No plasmid was used for these parts. Instead, due to their short length complementary oligos were annealed (**Methods**) to generate double-stranded DNA parts with necessary single-stranded overhangs for assembly.
- This spacer is used to replace block B and C in an assembly when no L2 regulation is present.

Supplementary Table 3: Performance of controllers in a cell-free expression system^a

Controller	Type ^b	Basal ^c (%)	Dynamic range ^d (au/hr)	Fold change ^d	Cooperativity ^e , <i>n</i>
P _{tac}	SLC	10.0	192	10	3.4
THS	MLC	0.3	4481	349	4.0
STAR	MLC	1.8	855	55	6.3
DC	MLC	3.7	462	27	4.1

- a. All values are averages calculated from three biological replicates. Key performance features of the controllers are visually shown in **Supplementary Figure 1**.
- b. SLC refers to 'single-level controller' and MLC refers to 'multi-level controller'.
- c. Relative basal expression rate calculated when no IPTG is present and as a percentage of the GFP expression rate for the 'on' state (10 mM IPTG).
- d. Calculated between 'on' and 'off' states for cells grown in 0 and 10 mM IPTG, respectively, and given in GFP fluorescence production in arbitrary units per hour (au/hr).
- e. From the Hill function fitting of the response functions (**Figure 5A**).

Supplementary References

- [1] 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.
- [2] Kennell D. & Riezman H. (1977) Transcription and translation initiation frequencies of the *Escherichia coli lac* operon. *Journal of Molecular Biology* **114**, 1–21.
- [3] Pai A. & You L. (2009) Optimal tuning of bacterial sensing potential. *Molecular Systems Biology* **5**, 286.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] Meyer A.J., Segall-Shapiro T.H., Glassey E., Zhang J. & Voigt C.A. (2019) *Escherichia coli* “Marionette” strains with 12 highly optimized small-molecule sensors. *Nature Chemical Biology* **15**, 196–204.
- [8] Westbrook, A.M. & Lucks, J.B. (2017) Achieving large dynamic range control of gene expression with a compact RNA transcription–translation regulator. *Nucleic Acids Research* **45**, 5614–5624.
- [9] Chappell, J., Takahashi, M.K., Lucks, J.B. (2015) Creating small transcription activating RNAs. *Nature Chemical Biology* **11**, 214–220.
- [10] Green, A.A., Kim, J., Ma, D., Silver, P.A., Collins, J.J. & Yin, P. (2017) Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121.
- [11] Chen Y.-J., Liu P., Nielsen A.A.K., Brophy J.A.N., Clancy K., Peterson T. & Voigt C.A. (2013) Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nature Methods* **10**, 659–664.