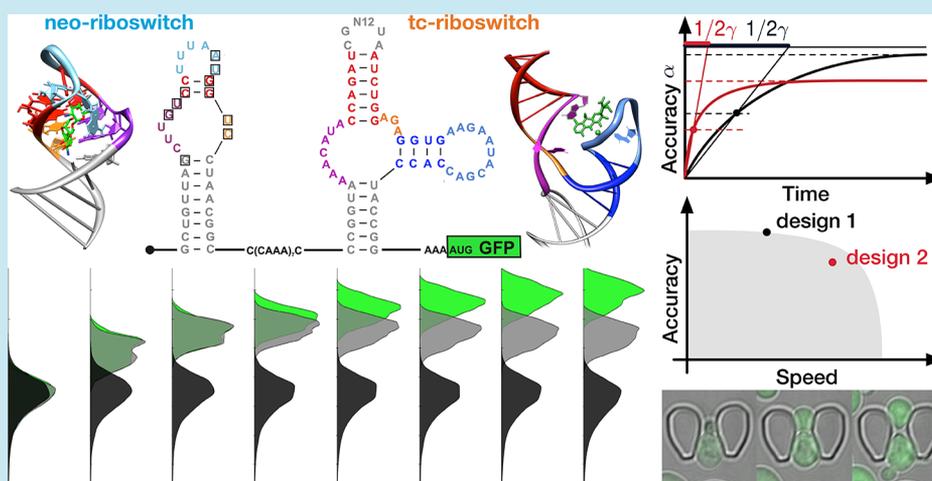


ROC'n'Ribo: Characterizing a Riboswitching Expression System by Modeling Single-Cell Data

Christopher Schneider,^{†,§} Leo Bronstein,^{‡,§} Jascha Diemer,[‡] Heinz Koepl,^{*,‡} and Beatrix Suess^{*,†} [†]Department of Biology, Synthetic Genetic Circuits, TU Darmstadt, Schnittspahnstrasse 10, 64287 Darmstadt, Germany[‡]Department of Electrical Engineering and Information Technology, Bioinspired Communications, TU Darmstadt, Rundeturmstrasse 12, 64283 Darmstadt, Germany Supporting Information

ABSTRACT: RNA-engineered systems offer simple and versatile control over gene expression in many organisms. In particular, the design and implementation of riboswitches presents a unique opportunity to manipulate any reporter device *in cis*, executing tight temporal and spatial control at low metabolic costs. Assembled to higher order genetic circuits, such riboswitch-regulated devices may efficiently process logical operations. Here, we propose a hierarchical stochastic modeling approach to characterize an *in silico* repressor gate based on neomycin- and tetracycline-sensitive riboswitches. The model was calibrated on rich, transient *in vivo* single-cell data to account for cell-to-cell variability. To capture the effect of this variability on gate performance we employed the well-known ROC-analysis and derived a novel performance indicator for logic gates. Introduction of such a performance measure is necessary, since we aimed to assess the correct functionality of the gate at the single-cell level—a prerequisite for its further adaption to a genetic circuitry. Our results may be applied to other genetic devices to analyze their efficiency and ensure their correct performance in the light of cell-to-cell variability.

For years special emphasis has been put on the creation and collection of various means to artificially control gene expression in both pro- and eukaryotic cells.^{1–4} Approaches range from simple but effective manipulations of promoter and terminator sequences,^{5,6} the organism-specific redesign of reporter genes⁷ and the introduction of protein- and RNA-based regulation to the reprogramming of fundamental cellular processes that operate orthogonally across several species.^{8–10} These advances have promoted an increased understanding of cellular functions and facilitated experiments where a tight regulation of gene expression was sought. Historically, the engineering focus has been put on proteins and as such, mainly on enzyme- and transcription factor-governed designs.¹¹ The importance of natural RNA regulation and the feasibility of RNA mimicry has been mostly neglected until the late discovery of such regulatory elements in prokaryotes in 2002.^{12–14} Since then, naturally occurring riboswitches have

been the subject of numerous research studies aiming at (i) structure determination,^{15,16} (ii) physiological relevance¹⁴ and (iii) their applicability and portability as artificial regulators of gene expression.^{17,18} The term riboswitches unites all RNA-based regulatory elements that are sensitive to an external input, typically a cellular metabolite, and convert that input signal by a change of their global folding into an altered expression of the controlled gene.¹⁴ Since the sequences of sensor and actuator are subdomains of a mutual architecture, riboswitches constitute an ideal scaffold to design synthetic derivatives with different ligand specificities and actuation profiles. Today, a growing number of RNA regulators including

Special Issue: IWBD 2016

Received: October 28, 2016

Published: June 7, 2017

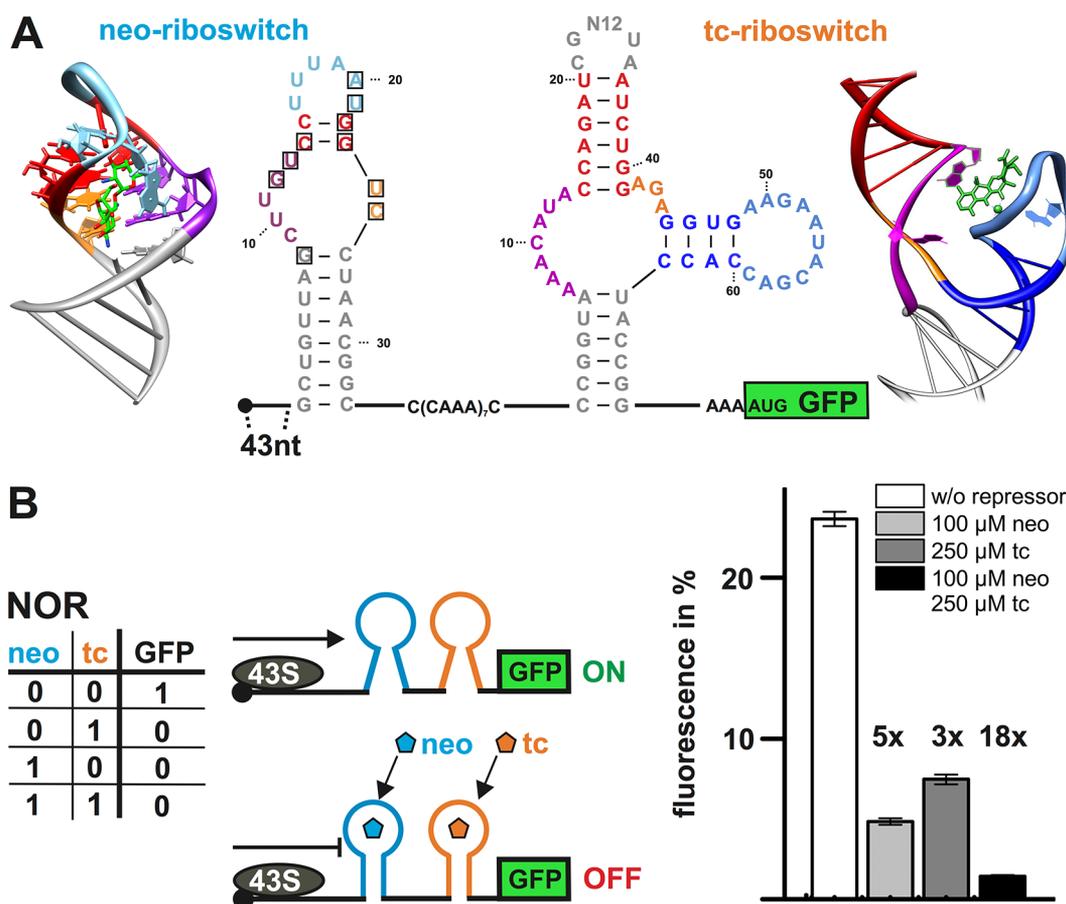


Figure 1. Design and analysis of the constructed NOR gate. (A) Constructed NOR gate at nucleotide resolution. The neomycin (neo) and the tetracycline (tc) riboswitch are placed cap-distal in the 5'UTR of a GFP+ reporter gene and spaced by a repetitive CAAA motif. P_{ADHI} and T_{ADHI} are used for transcription and a proposed yeast Kozak sequence (A at -3 and G at $+4$ relative to the translation start site) marks translation initiation.^{39,40} Motifs involved in ligand binding are highlighted in color and nucleotides involved in ligand binding are boxed in the 2D structure or specified in the 3D structure. Adopted from ref 41 and further modified. (B) Addition of either ligand leads to a decrease of gene expression due to an interference with 43S ribosomal scanning. Logic operations are illustrated by a scheme of the constructed gate. A single-input corresponds to a NOT logic that can be expanded to a NOR logic upon addition of the orthogonal ligand. Cells transformed with the gate were cultured in the absence and presence of neomycin and/or tetracycline and analyzed by fluorescence spectrometry in a bulk measurement after 24 h. Output fluorescence was normalized to cell growth, background corrected and is displayed relative to a positive control that is not riboswitch-controlled but otherwise treated equally. Experiments were performed in three independent triplicates and repeated twice.

synthetic riboswitches may be readily used for conditional gene expression in *Escherichia coli*, *Saccharomyces cerevisiae* and several human cell lines.^{19–21}

Typically, a dedicated expression system is set up as a proof-of-principle study for a newly engineered riboswitch and some of its kinetic and thermodynamic properties are explored.^{22–24} While this strategy has greatly contributed to create and collect novel parts, it turned out that especially riboswitches exhibit high sensitivity to the neighboring sequences and may thus forfeit their original performance when transferred to a different genetic context,^{17,18} although several design principles have been proposed.^{21,25–27} Suffering from unresolved structural interferences imposed by the genetic context and a lack of data accounting for the pronounced cell-to-cell variability, these parts are deemed to vanish from the synbio toolbox.

As a consequence, only a few examples of logic devices and larger genetic circuits composed of RNA-controlled devices, in particular riboswitches, have been set up to date.^{18,28–31} Mostly operating in *E. coli* and partly in *S. cerevisiae*, the developed logic gates employ distinctly different RNA-based approaches highlighting the versatility of this general RNA-based design

strategy. The RNA toehold switches generated by Collins and co-workers²⁸ feature high orthogonality of two independent RNA parts connected by *in trans* interactions and have been yielded by a combination of *in silico* computation and rational design. Hartig and co-workers have engineered the twister ribozyme,²⁹ an *in cis* acting RNA part, to sense different ligands mainly by permutation experiments guided by rational design principles from previous studies.^{21,24,32} The performance of any logic gate presented in these two studies appears to be generally lower compared to the single-input counterparts, underlining the need for connectable RNA parts without compromised functionality, *i.e.*, performing better than the single parts of which they are composed.

To further bridge the gap between the creation and collection of modular and portable RNA parts toward the assembly into devices and higher order circuits, an accurate characterization of all parts is mandatory. On one side, this involves the computation of mRNA structures to model folding constraints and aid in the design of the actuating domains and their embedment in a neutral (noninterfering) genetic context. On the other side, defining the scope of our research, it involves

the characterization of the input-output behavior and associated cell-to-cell variability of devices in terms of intrinsic and extrinsic noise. Extrinsic noise, defined as variations in device behavior through fluctuations of molecular components that are considered extrinsic to the device, adversely affects the development and maintenance of a homogeneous cell population.³³ Differences in cell cycle stage, ribosome copy number, ribosome utilization, and other resource competitions lead to performance deficiencies across a cell population. The dependency of device performance on such contexts that are considered extrinsic to the device was identified to be the major stepping stone toward a rational design.^{6,34} For robust functionality and reproducibility it is essential to account for all those features when characterizing a device. Single-cell studies allow to characterize phenotypic variability and to identify the different sources contributing to the overall variability within the population. Moreover, time-lapse recordings allow to recapitulate the expression dynamics at the single-cell level, revealing changes in cell-to-cell heterogeneity and yielding nonaveraged kinetics of a device's performance. This is in strong contrast to the widespread bulk measurements that are recorded during the steady state phase of the cell, losing all aforementioned information. As a consequence, plausible computational models need to account for the kinetics of a device and for the accompanied cell-to-cell variability in the device's behavior.

In this study we present an approach that facilitates the comprehensive transient single-cell characterization of a novel RNA-controlled device constructed from the neomycin- and tetracycline-sensitive riboswitches and implemented in *S. cerevisiae*. By coupling the two riboswitches that both block translation initiation by the same mechanism, representing a NOR operation, we sought to improve the dynamic range between ON and OFF state of reporter gene expression. The device can also realize single-input NOT operations, if only one ligand is added. We studied the transient dynamics of this logic gate at the single-cell level through time-lapse flow cytometry and live cell imaging, and along with that, derived a novel gate performance indicator based on the well-known concept of receiver-operating-characteristics (ROC). Importantly, this indicator accounts for cell-to-cell variability and can be applied to any logic gate. Introduction of such a performance measure is necessary, since we wanted to assess the correct functionality of the gate at the single-cell level—a prerequisite for its further adaptation to a genetic circuitry. We also determined the NOR and NOT functionality for different expression levels of the gate by adopting the GEV system developed for inducible transcriptional regulation.³⁵ From the acquired data we built and calibrated a stochastic single-cell device model that accordingly accounts for intrinsic but also for extrinsic noise. We then used the model to complete the characterization of the repressor gate *in silico*, e.g., determined the gate's full dose response profile at steady state. Using the model we investigated a redesign option to improve gate performance. Although initial considerations of noise propagation hinted toward reduction of extrinsic noise stemming from GEV variability, the model-based analysis predicted little effect by a tighter control of that parameter. To test this prediction we conducted a genomic integration of the GEV device and found reduced GEV variability, but no increase in gate performance, validating this *in silico* prediction.

RESULTS AND DISCUSSION

Construction of a Repressor Gate Operating by NOT and NOR Logic. Starting with the neomycin- (aptamer M4²²) and tetracycline- (aptamer cb32-ANsh³⁶) binding riboswitches, we designed a repressor gate capable of transmitting NOT and NOR logic behavior (Figure 1A). The two riboswitches are separated by multiple CAAA-spacer units to minimize global folding constraints between the two RNA architectures. This strategy has also been used in other studies, since CAA(A) repeats do not form any secondary structures and may thus be seen solely as an insulation module. As both riboswitches operate by conformational capture mechanisms, it is crucial for their functionality to facilitate free fluctuations between their ground-state conformations.³⁷ The initial expression platform was set up on a plasmid with 2 μ ori, using the *ADHI* promoter and terminator to drive the expression of the GFP+ reporter gene.³⁸ To assess gate functionality we recorded fluorescence of the *S. cerevisiae* strain RS453 α transformed with the reporter device in a standard steady-state measurement after 24 h.

Figure 1B shows the fluorescence output of repressor gate-regulated cells relative to a positive control that is not controlled by a riboswitch. Inhibition of gene expression is achieved by blocking ribosomal scanning once the 43S complex encounters the stabilized RNA-folding of the ligand-bound riboswitch. Basal gene expression is reduced to about 24%, owed to the rather extensive global RNA-architecture imposed by the two riboswitches in the 5'-UTR of the reporter gene, but leaving enough room for the testing of an ON/OFF inverter gate. Logic operations are governed by a sequential repression profile. The gate shows a stronger repression for the neomycin single-input (5-fold) as compared to the tetracycline single-input (3-fold) and processes an increased repression through the concerted action of both ligands (18-fold). On a first level, the device can thus be used to only moderately inhibit gene expression, whereas its application as a NOR gate leads to a multiplied decrease of residual reporter expression to about 2%, yielding an increased level of regulation. Preliminary data of the implementation of the inverse TcNeo gate showed unexpected change in gate behavior, hinting toward unresolved structural interferences (Supporting Information S2.1). Hence, we decided to exclude the gate from our single-cell analysis and continued with the population analysis of the constitutively expressed NeoTc gate that underlined the broad phenotype arising in the fluorescence profile of cells harboring 2 μ -replicated plasmids; the major drawback of systems with a multicopy replication origin and addressed in^{5,42} (see Supporting Information S2.2). Besides the high cell-to-cell variability conferred by 2 μ plasmids, they impose an increased metabolic burden on the host cell, especially when equipped with strong promoters that drive the overexpression of reporter genes. Genetic circuits of any complexity constructed from several 2 μ plasmids will therefore exhibit severe noise propagation and ultimately broad distributions of phenotypes, rendering performance on the single-cell level rather poor.⁴³ Based on these findings we chose to continue with the chromosomally integrated version of the NeoTc gate, as it presented promising switching rates and an orthogonal dual-input repression behavior, serving as a prototype for further analysis and computational modeling purposes.

Transient Characterization of the NeoTc Gate. Time course measurements of device behavior during induction provide insight into the kinetics of the device, such as the time

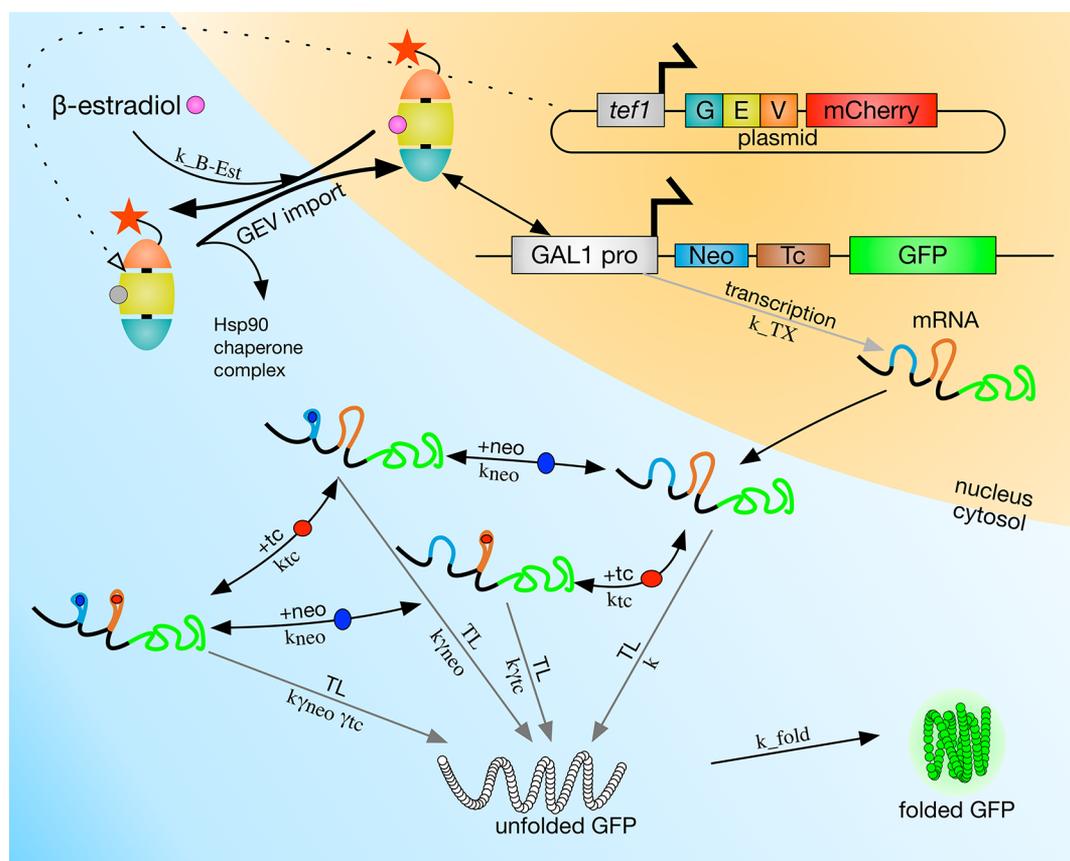


Figure 2. Illustration of the repressor gate and its computational model. The transcription factor GEV is expressed constitutively from an ARS/CEN plasmid. The NeoTc gate is chromosomally integrated with GFP+ as reporter gene. All arrows indicate reactions covered by the stochastic kinetic model with corresponding reaction rates. Upon induction with β -estradiol GEV localizes to the nucleus and activates transcription of the gate. The mRNA binds a repressor and forms its stable stem loop that interferes with 43S ribosomal scanning. The binding of neomycin and tetracycline to the mRNA are assumed to be independent from each other. Each mRNA species has its own translation rate.

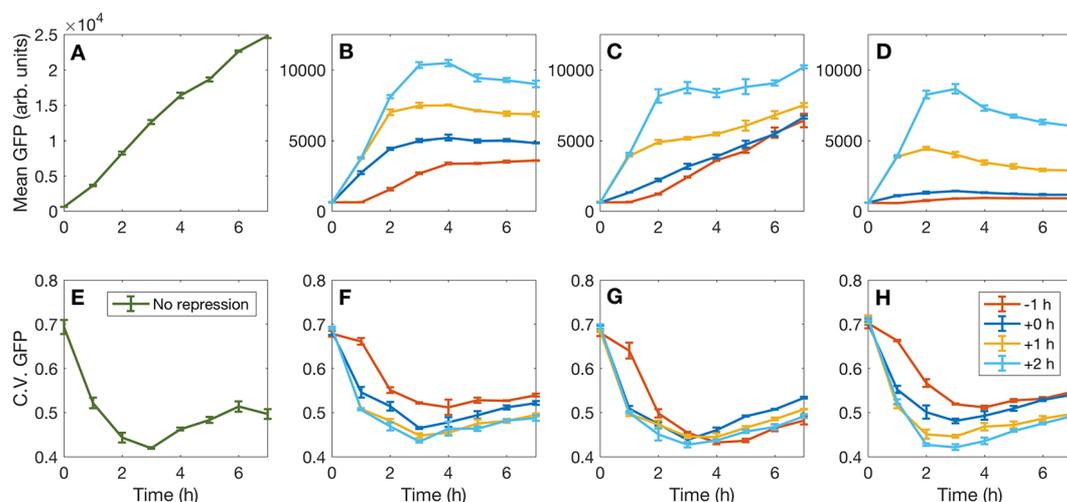


Figure 3. Gate characterization on the single cell level. (A–D) Fluorescence means over time for unrepresed cells (A) and cells repressed with 100 μ M neomycin (B), 250 μ M tetracycline (C) and both (D). (E–H) show the corresponding coefficients of variation (CV) summarizing the width of the cell population. Trajectories are color-coded to show different times of repression; 1 h of repression before induction (red), simultaneous repression and induction (dark blue), repression after 1 h of induction (yellow) and repression after 2 h (light blue). Induction for all experiments with 10 nM β -estradiol. Samples were drawn every hour for 7 h from cultures grown in 12-well plates. Data is averaged over three technical replicates as indicated by the error bars and is representative for three independent biological replicates. The complete data set of the biological replicates is presented in S2.5. Fluorescence background is not subtracted for the plotted analysis.

interval it requires until it equilibrates to a new expression profile. Characterizing the device kinetics on the single-cell level

moreover, allows to capture the heterogeneity in the onset of a response to the inducer directly visible in live cell imaging and

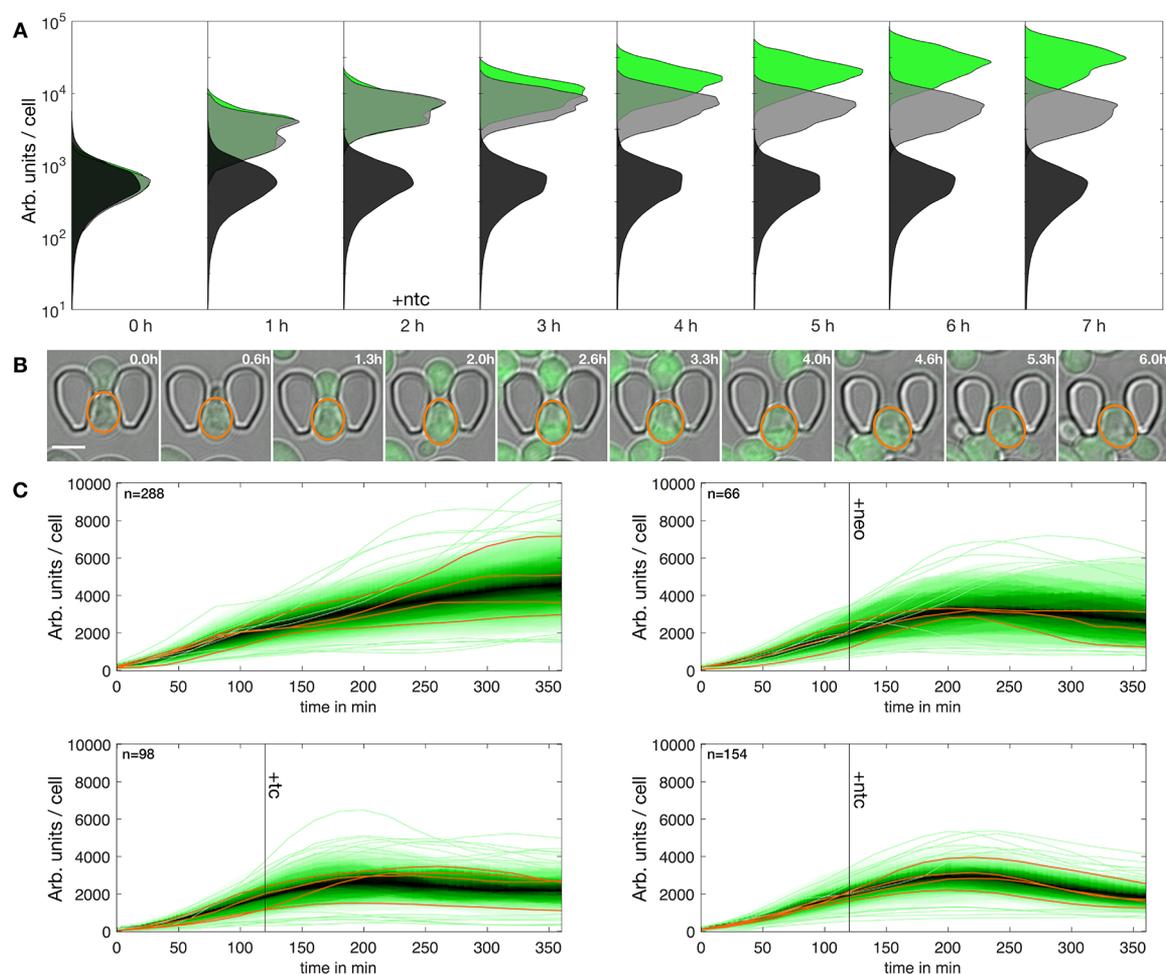


Figure 4. Qualitative population dynamics of the NeoTc gate. (A) Histogram of time-lapse cytometry data from a +2 h experiment. Uninduced cells (black) show a broader distribution than cells induced at 0 h (green) and cells induced at 0 h and repressed at +2 h (gray). (B) The picture series shows a representative cell trapped on the microfluidic chip and recorded over time by fluorescence microscopy. Scale bar 5 μm . (C) Density plots of segmented single-cell traces. Random examples for individual traces are highlighted in orange. Induction with 10 nM β -estradiol, repression with 100 μM neomycin and/or 250 μM tetracycline. Color becomes darker the more cell traces approach the median. For cytometry samples were drawn every hour for 7 h from cultures grown in 12-well plates. Plotted data is representative for at least three independent replicates. The same inoculums were used to conduct microfluidic measurements and analyzed data is representative for at least three independent replicates. Fluorescence background is not subtracted in both experimental setups.

at least discernible in flow cytometry data through transient fluctuations of the measured distributions. Resolving the transient dynamics of a device is also important for building and calibrating computational models for it. Characterization at steady state provide only poor data for the calibration of such models, often leading to severe identifiability issues. Recent studies^{44,45} have shown that the distributional information (e.g., variance additionally to the mean) available from single-cell data can partly alleviate this shortcoming, but transient data is generally required to circumvent nonidentifiability. Here we present a characterization protocol for the considered riboswitch gate that generates sufficient data for the kinetic characterization of the device and for its subsequent modeling.

To prepare our platform for kinetic analysis and to circumvent the variable copy number of the 2 μ replication origin we adopted the GEV system³⁵ and conducted a CRISPR/Cas9-guided⁴⁶ genomic integration of the gate (Figure 2). GEV is a chimeric transcription factor composed of a VP16 transcription activation domain, a GAL4 DNA binding site and, most importantly, a β -estradiol-receptive domain. The transcription factor itself is constitutively

expressed from an ARS/CEN plasmid using the *TEF1* promoter and *CYC1* terminator. Additionally, the RFP derivative mCherry is C-terminally coupled. Transcription of P_{GALI} -controlled genes is switched on by supplementation of β -estradiol. Only a few other off-target genes have been identified by Isaacs *et al.* and were found to be negligible. GEV must be transcribed by a strong promoter and activated by a prior titrated concentration of β -estradiol to achieve fast ON-rates and simultaneously avoid diminished cell growth.

In a first experiment, GEV drove the transcription of the chromosomally integrated NeoTc gate. Output data was recorded by flow cytometry during exponential growth to avoid errant expression patterns by a metabolic adaptation (see Supporting Information S2.3). Additionally, cultivation of yeast cells transformed with the positive control plasmid in absence and presence of both repressors revealed no influence of both repressors on GFP and GEV expression (see Supporting Information S2.4). To gain information about the time scale and transient strength of translation inhibition by both repressors, neomycin and tetracycline were added at different time points (−1 h, 0 h, +1 h, +2 h) relative to the induction

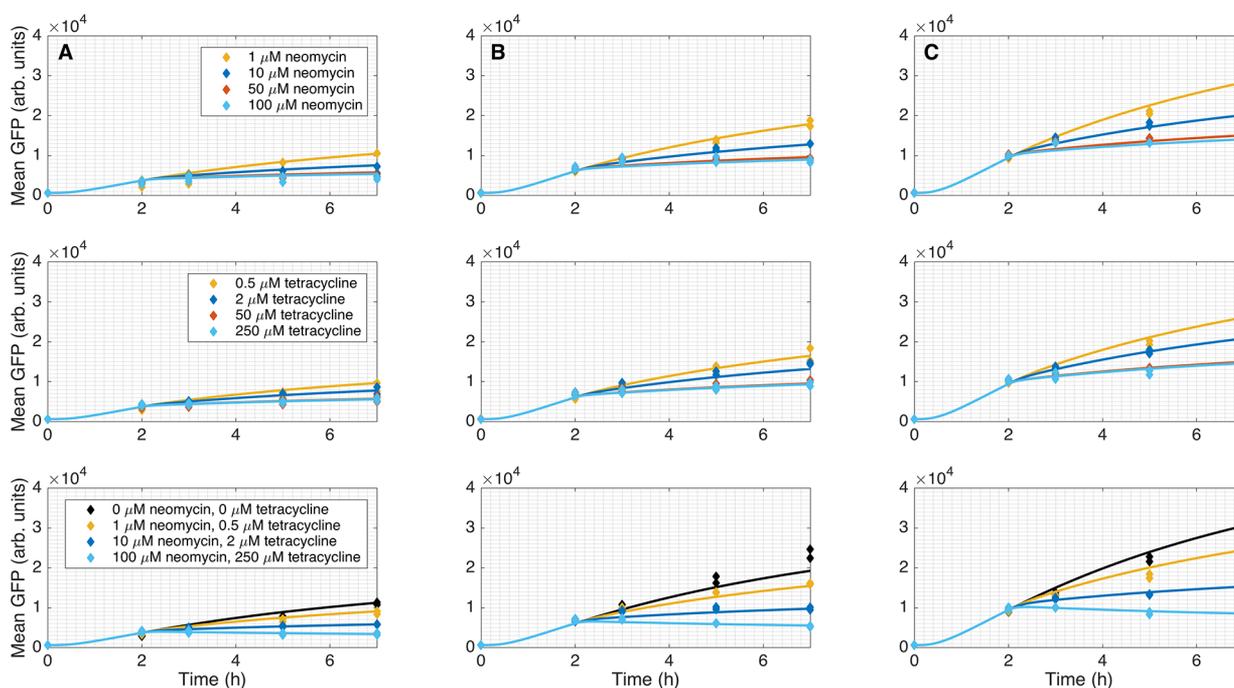


Figure 5. Matrix of dose–response profiles at varying induction levels. NeoTc gate induced with (A) 2.5 nM β -estradiol, (B) 5 nM β -estradiol and (C) 10 nM β -estradiol and repressed with indicated ligand concentrations after 2 h induction. Fluorescence output was recorded by flow cytometry. Mean fluorescence values (dots) are plotted over the time course of the experiment from samples drawn at the indicated time points. The calibrated computational model reproduces the full time-resolved response of the device (lines) for all gate induction levels and for all combinations of repressor concentrations. Cultures were grown in 12-well plates. Plotted data is representative for at least five independent replicates. Fluorescence background is not subtracted.

with β -estradiol at 0 h (Figure 3A–D). Onset of GFP+ expression is fast enough to clearly discern induced cells from uninduced cells after 1 h by an 8-fold increase of the mean fluorescence value (Figure 3A, 4A). By 2 h, the induced populations have shifted to a near complete baseline separation from background fluorescence (13-fold induction) and show an almost linear increase of fluorescence over the time course of the experiment (Figure 3A). The device features a high repression efficiency for the dual-input, as indicated by the –1 and 0 h experiments, where only a 2-fold increase of the mean fluorescence over background level is observed for 7 h. Importantly, onset kinetics of the repression cascade triggered by ligand addition show an immediate effect on gene expression within the first hour after ligand addition, irrespective of the time point after induction (Figure 3D). Thus, the device processes a NOR operation within a short time scale, enabling low-delay repressive operations for genetic circuit design.

The single-input data suggest different repression kinetics for the neomycin- and tetracycline-inputs, respectively (Figures 3B,C). The device exhibits a progressive decrease of repression efficiency for the tetracycline-input, reflecting a fast onset followed by a gradual reduction of repression with advancing induction (Figure 3C). The differences in the onset of repression for both ligands are best seen by comparing the slopes of the graph between 1 and 2 h for the +1 h experiments and between 2 and 3 h for the +2 h experiments. For the single-input these slopes are parallel with respect to the different time points of repression. This is also observed for the dual-input experiments and hints toward a timing-independent application of the device as stated above (Figures 3B–D). The comparison of these slopes for the two single-inputs reveals the 4-fold faster onset that is achieved with tetracycline. Interestingly, the onsets

of the +1 h and +2 h experiments are very similar for the tetracycline-input (750 arb units/h) and the dual-input (800 arb units/h). This points toward a predominant role of the tetracycline-sensitive riboswitch in the fast repression kinetics of the NOR gate that are complemented by the more efficient repression by neomycin in the long run (see Figure 12 for comparison). The observed time course of the repression profile could be reproduced in three biological replicates. The data is presented in Supporting Information S2.5. Figures 3E–H show the coefficients of variation (CV) for all experimental setups to account for cell-to-cell variability. Although the CVs change overall less than 2-fold, the CVs decrease upon induction indicates a stabilizing effect on phenotypic heterogeneity. This effect correlates mainly with the disappearance of all background fluorescence that can be traced back to autofluorescence of the cells and β -estradiol independent GEV activity broadening the phenotypic distribution of uninduced cells, but vanishing upon induction among all other induced cells, ultimately reducing the CV.

In Figure 4A we illustrate the population dynamics exemplified by the dual-input repression for the +2 h experiment. In accordance with the CV analysis, cell-to-cell variability is reduced by induction, turning cells toward an ON-state at varying time scales reflected in a smooth population shift. Repression at +2 h causes the population shift to halt as further translation is blocked efficiently by the NeoTc gate. Although flow cytometry is a powerful technique to analyze a bulk culture at the single-cell level, it does not allow us to track a single cell over time to determine the characteristics of a transient response on the single-cell level. To link the snapshot data of a population recorded by flow cytometry to the single-cell transient dynamics we analyzed the NeoTc gate for a +2 h

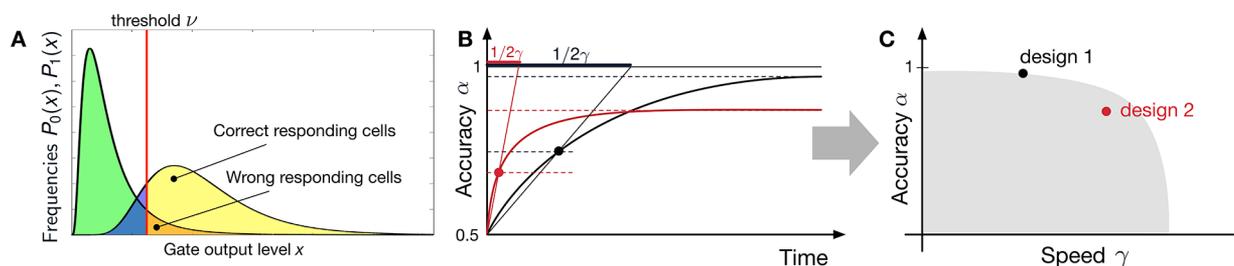


Figure 6. Illustrating performance characterization using ROC analysis. (A) Empirical distributions for the ON state (right) and the OFF state (left) of a gate across a cell population. Generally, both distributions are divided into two parts by a choice of decision threshold. Corresponding to each, there are cells in the OFF state producing output levels above (false positive rate, orange) and cells in the ON state correctly producing output levels above that threshold (true positive rate, yellow). Note that the true negative rate and the false negative rate are functions of the false positive rate and true positive rate, respectively. Thus, they do not enter explicitly into the calculation of the ROC curve. (B) The accuracy of the gate is computed as the achieved asymptotic area under the ROC curve (AUC), whereas the speed of the gate is determined by the time it takes until its AUC reaches half-maximum. (C) Different gate designs will be located in the two-dimensional performance space, where trade-offs for a given gate architecture define fundamental performance constraints that cannot be overcome for that architecture (gray-shaded area).

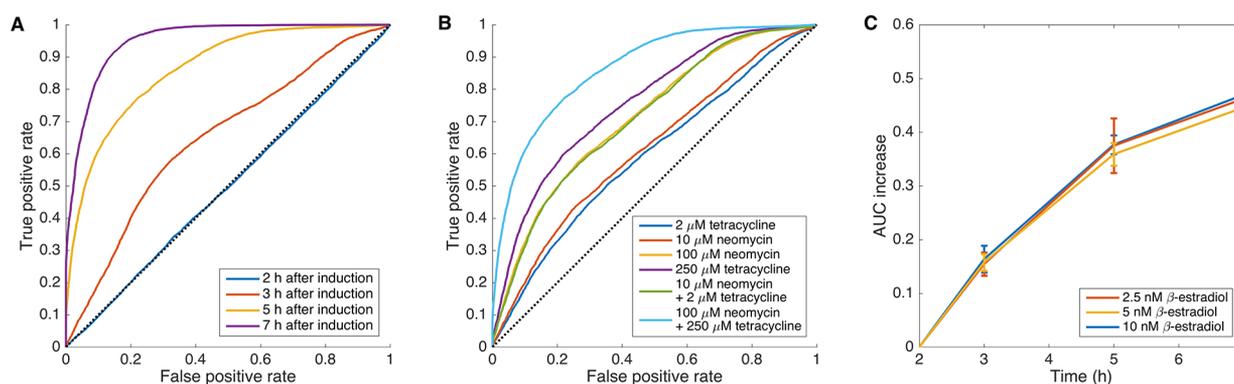


Figure 7. ROC analysis of flow cytometry data. (A) ROC curve evolution over time for full induction (10 nM β -estradiol) and maximum repression (100 μ M neomycin and 250 μ M tetracycline) initiated after 2 h of induction. To understand how to read these curves, consider, for example, that we have chosen the decision threshold ν to achieve a false positive rate of 0.2 at 5 h after induction. Then a true positive rate of approximately 0.75 is achieved. (B) ROC curves at different concentrations of repressors (added after 2 h), induced with 10 nM β -estradiol and measured after 5 h. In (A) and (B), the dotted diagonal line shows the performance of a hypothetical, completely random (*i.e.*, dysfunctional) gate. (C) Increase of AUC relative to its value at the time of repressor addition (100 μ M neomycin and 250 μ M tetracycline, 2 h after induction) over time for different inducer concentrations.

experiment on a microfluidic chip by time-lapse fluorescence microscopy (Figure 4B). In addition to the density profile, we randomly picked single-cell traces to demonstrate their coherently developing phenotypes over time. That is, a highly expressing cell emits a stronger fluorescence signal after several hours of repression than a lowly expressing cell, thus their dynamic fluorescent phenotypes develop coherently. In total, our live-cell imaging data provides further backing for the considered time-lapse flow cytometry data set.

Having identified that the repression kinetics and efficiencies are independent of the preceding induction period, we next assessed gate performance at different induction levels and determined its dose dependence on different repressor concentrations (Figure 5). Yeast cultures were sampled at indicated time points (dots), whereas the calibrated model reproduces the full time-resolved response of the device (lines). Induction with progressively increasing levels of β -estradiol leads to a proportional increase of the GFP+ signal. The gate features an equal repression onset and efficiency for all three induction levels indicated by the slopes within the hour of repression and the expression levels after 7 h relative to the time point of repressor addition. Consistent with our precedent results, these experiment series highlight the high repression efficiency after addition of both ligands. For 100 μ M neomycin

and 250 μ M tetracycline translation is fully blocked and the observed signal decay is congruent with the half-life of the reporter GFP+. By supplementing different repressor concentrations we can map out the full dose–response profile of the gate and compute its performance with respect to the three inducer concentrations as presented in the following section and summarized in Figure 7 and 10.

Performance Characterization. For logic-based devices, the fold-change at the bulk level is a poor descriptor of the gate performance. Importantly, the cell-to-cell variability in the device behavior needs to be accounted for. Hence, apart from the fold-change between the Boolean states, the overlap between the single-cell distributions associated with these two states is crucial. The natural design goal is to require correct functionality of the gate at the single-cell level and not just at the bulk level; one can well imagine designs where the Boolean states seem well separated in terms of means, but for which their standard deviations are as large as the separation of their means and thus device functionality at the single-cell level is very poor (see Figure 6A). Here, we provide a new measure for gate performance that incorporates cell-to-cell variability and relies on well-established facts from decision theory.^{47,48} In particular, we interpret single-output logic gates as detectors or classifiers, for which the natural performance characterization is

the ROC (receiver-operator-characteristics) curve (see Figure 7A,B).

The ROC and its cumulative measure, the AUC (area under the ROC curve), summarize the classifier performance independent of a particular threshold. To define the ROC of a general n -input single-output gate, denote the collection of measured gate outputs across a cell population for all input combinations corresponding to output ON by p_1, \dots, p_N , and by q_1, \dots, q_M the respective gate outputs for all input combinations with output OFF (see Figure 6A). For any $\nu \in \{p_1, \dots, p_N, q_1, \dots, q_M\}$, we obtain one point on the ROC curve, with coordinates $\beta(\nu)$ and $\delta(\nu)$ in the false positive rate and true positive rate plane, respectively,

$$\beta(\nu) = \frac{1}{M} \sum_{i=1}^M \mathbb{I}(q_i > \nu)$$

$$\delta(\nu) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(p_i > \nu)$$

and

$$\alpha = \frac{1}{NM} \sum_{i=1}^M \sum_{j=1}^N \mathbb{I}(p_j > q_i)$$

where we introduced AUC as α . Since the number of cells measured in flow cytometry is very large, the precise method of interpolation between the discrete ν values is not important. For a chosen loss function (e.g., probability of false response $1 - \delta(\nu) + \beta(\nu)$), which will generally depend on the desired application, the optimal threshold can then be determined, yielding one point along the ROC curve. The incurred loss for that threshold is then the lower bound, achievable only if the gate is cascaded with an ideal threshold device (detector). With that, the loss characterizes the gate independent of the particular implementation of the threshold device. In practice the threshold device needs to be realized through additional reactions, i.e., through high cooperativity and fast time scales. For the case of normally distributed data, which is practically often encountered if the flow cytometry measurements are log-transformed, the AUC admits an explicit and intuitive expression that just involves the empirical means μ_0, μ_1 and standard deviations σ_0, σ_1 of the measured distributions, computed in the log-domain

$$\alpha = \Phi(\rho)$$

with

$$\rho = \frac{\mu_1 - \mu_0}{\sqrt{\sigma_1^2 + \sigma_0^2}}$$

where Φ is the cumulative distribution function of the standard normal distribution.⁴⁷ Similarly, the full ROC curve can then also be expressed solely in terms of these quantities. As Φ is a monotone increasing function any gate superior under α will remain superior under ρ (in detection theory, ρ is called the sensitivity index⁴⁹). Hence, ρ can be used as an easy to compute, yet principled, performance indicator that accounts for the cell-to-cell variability in gate output. More specifically, it puts the gate's fold-change, given by the numerator, in relation to the width of the ON and OFF distributions, given by its denominator. Figures 7A shows the resulting performance of the gate for a specific repressor concentration over the course

of a 7 h experiment, while Figure 7B shows the performance for a combination of the repressors at the same induction level after 5 h. Correspondingly, Figure 7C depicts the time course of AUC during the course of three different induction levels. We emphasize that this ROC and AUC characterization is done solely on the experimental data. Figure 7C indicates that gate performance within the observation time interval is independent of the gate induction levels that were considered, suggesting that AUC α characterizes the gate as such and is not affected by its expression level.

Note that if the two distributions corresponding to the Boolean states show negligible overlap, then α plateaus at one, indicating that a perfect classifier is feasible in every single cell.⁵⁹ If the separation of the distributions beyond that nonoverlapping point should enter the performance characterization, then ρ provides a better measure as it further increases with increasing separation of the two distributions (i.e., increasing the fold change).

The slope of the AUC α (and also of ρ) accordingly provides a measure of how fast the gate becomes functional; in particular we define the slope or speed γ in terms of the inverse time it takes to reach half-maximum performance (see Figure 6B). On the basis of this observation we extract two orthogonal features of the gate that describe (i) accuracy of the gate and (ii) speed or responsiveness of the gate. Every gate design can then be mapped into a two-dimensional performance space as indicated in Figure 6C. For a given device architecture only a certain region of the performance space will be accessible through parameter variations and some parametrization thereof will lead to designs that are at the Pareto-front and hence optimal. The characterization of the NeoTc gate according to this performance space will be done in [In Silico Completion of Device Characterization](#) section.

Hierarchical Stochastic Models. In order to build a computational model that can quantitatively account for the single-cell device characterization data it is mandatory to include intrinsic noise but also extrinsic noise, i.e., the dependency on the molecular makeup of the host cell. Following recent work,^{50,51} we formulate a hierarchical probabilistic graphical model of the device that captures both sources of variability. The graphical model comprising latent and observable variables is shown in Figure 8. The model includes the forward scattering channel (FSC) as an observable, because we found positive correlation between it and the expression levels of the reporter proteins. The set of reactions that the model accounts for are displayed in Figure 2, whereas the complete list of reactions is given in the [Supporting Information S1.1](#).

To have a sound biophysical basis at the stochastic level, all propensities are confined to mass-action kinetics of at most arity two. The reaction system involves ligands, transcription factor translocation and gene expression and therefore exhibits a multiscale nature that usually results in high computational burden for stochastic simulations. Hence, for calibrating the model one can resort to more efficient approximations in terms of moment-closure⁵⁰ or through linear noise approximation (LNA).⁵² For both approximations extrinsic noise can be incorporated with ease.^{50,53} Here we utilize LNA. The inference follows a Bayesian approach, where Markov chain Monte Carlo (MCMC) together with parallel tempering is used to efficiently sample the posterior distribution over model parameters. The model inference used mean and variance of the experimental data. Taking the posterior mean for each parameter (see

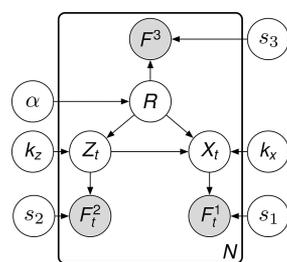


Figure 8. Probabilistic graphical model of the complete device across a cell population. Nodes indicate random variables or random processes (indicated through time dependency subscript) and edges indicate statistical dependency; box refers to so-called plate-notation, indicating that the graphical model within the box is repeated N times, where N is the number of considered cells. Latent molecule states corresponding to the GEV device (Z_t) and the observable fluorescent readout (F_t^2) of GEV abundance; latent states of the NOR gate (X_t) and observable fluorescence readout of the GFP+ reporter abundance (F_t^1); sets of latent kinetic parameters (k_z, k_x) for both subsystems, heterogeneous but constant translation efficiency (R) and the forward scattering channel (F^3) as observable covariate of R ; distribution or hyperparameters α, s_1, s_2, s_3 .

Supporting Information S1.2 for details) we generated model outputs for all measured inducer and repressor concentrations that are stated in Figure 5. The trajectories predicted by the calibrated model in terms of means are shown as lines in Figure 5. Since our inference only used measured means and variances, we investigated if the stochastic model is able to predict the full distributions over GFP+ for the cell population. In order to do so, we performed 10 000 stochastic simulations of the model for each condition and recorded the marginal distributions. The results are shown in Figure 9, suggesting that the model is able to recover the approximate shape and trend of the entire distribution across conditions.

In Silico Completion of Device Characterization. The short and transient experimental characterization of the gate does not provide a full account of the gate's behavior, but is

directly geared toward providing minimal, yet informative data for translation of the system to the *in silico* domain. Having a computational model that accurately predicts the repression dynamics and its cell-to-cell variability, one can now use the model to complete the device characterization *in silico*. In particular, one can fully map out the dose–response curve of the repressor and the dose–response surface of the NOR gate at arbitrary time points and especially at steady state phase with respect to all activating and inhibiting rate constants. In such a way, we can emulate the device's output over the time course of simulated prolonged exponential growth; data that are beyond the currently experimentally feasible time window. Figure 10 shows the dose–response of the NOR gate at 7 h and at steady state for different induction levels of the gate. The experimental data at 7 h aligns remarkably well with model predictions. As has been previously observed, different induction levels do not influence the performance of the gate and thus computed dose–response surfaces superimpose each other, only distanced by the respective induction strength. The model predicts a residual fluorescence output for maximum repressed cells at the higher induction levels of 5 nM (3-fold) and 10 nM β -estradiol (6-fold) relative to the starting level of uninduced cells at ~ 500 arb units, suggesting the continuous translation of a small number of mRNAs escaping any repressor interaction that is also observed with the constitutively expressed gate (compare Figure 1B).

From the steady state predictions we then locate different NOT and NOR gate configurations in the proposed two-dimensional performance space of accuracy and speed. Figure 11 shows the coordinates of the NOT/NOR gate at various repressor concentrations in the two-dimensional performance space. It is evident that addition of higher repressor concentrations lead to a faster response of the gate and a better separation of the two Boolean states with respect to cell-to-cell variations and the fold-change between them. The performance space mapped here can then be used to characterize the gate design as shown in Figure 6C by simply setting a core coordinate for the gate design itself and defining

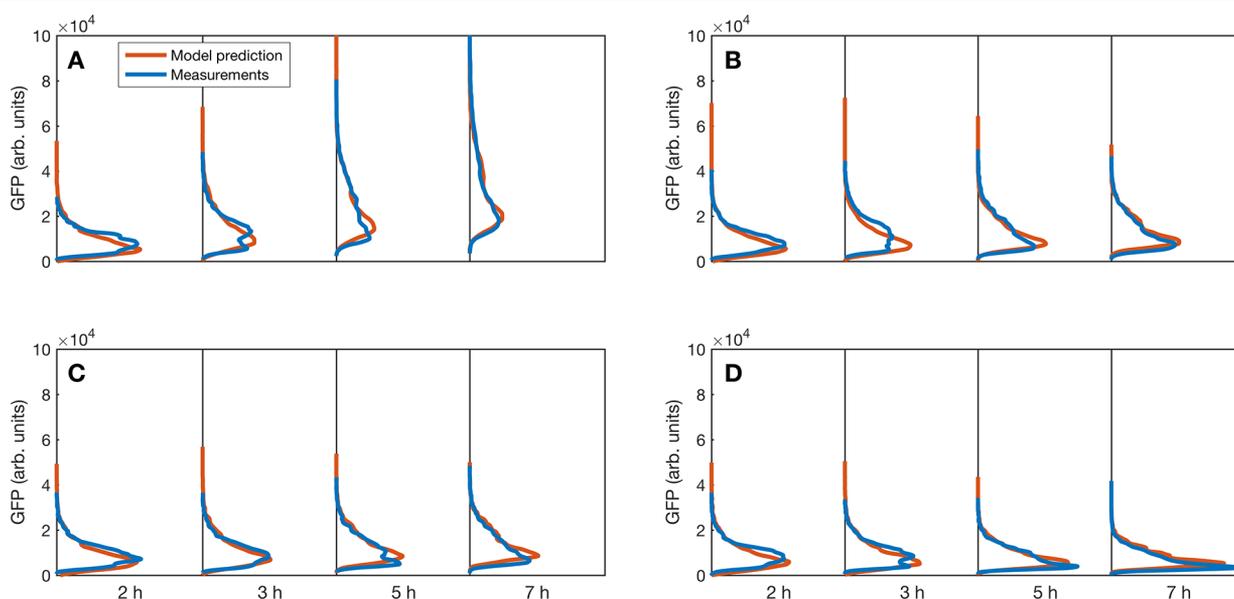


Figure 9. Predicting the full shape of cell-to-cell variability from a model. Time evolution of the empirical distribution function from flow cytometry (blue) and from stochastic simulations (red) starting at the time of repression at 2 h; no repression (A), 100 μ M neomycin (B), 250 μ M tetracycline (C) and both neomycin and tetracycline at respective concentrations (D).

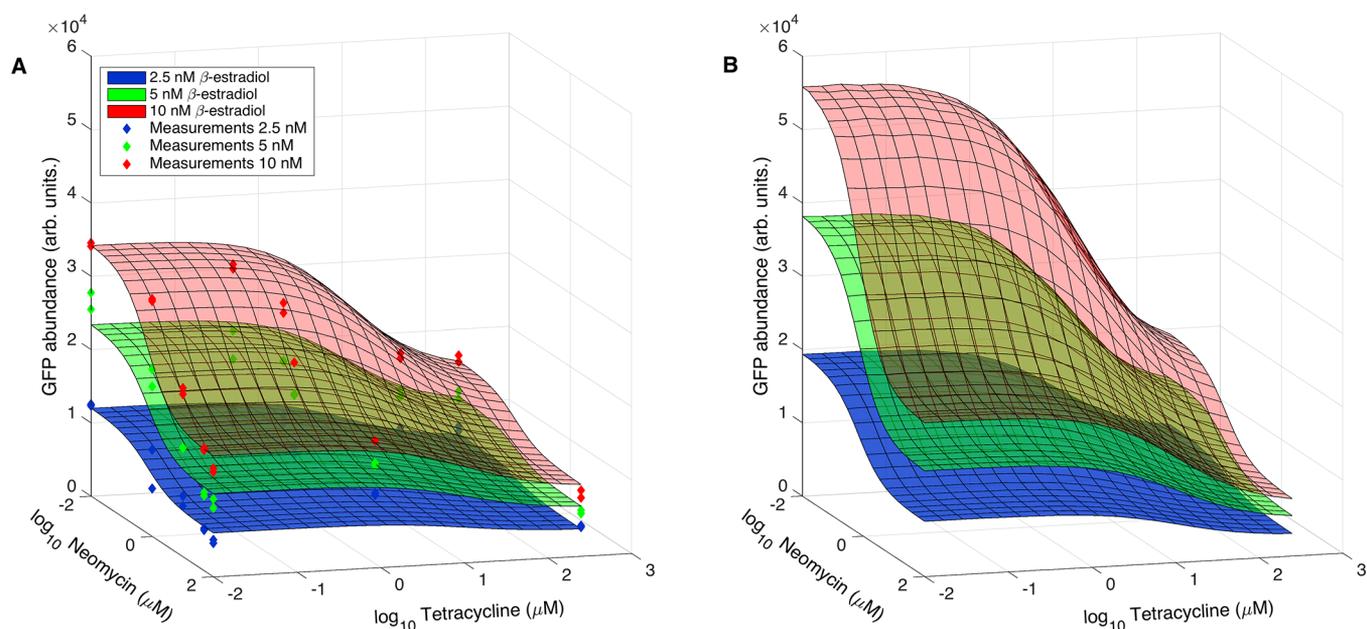


Figure 10. Dose–response surfaces at various induction levels of the gate, showing the dependence of gate output on the concentrations of repressors. The system was measured until 7 h after induction. The model, calibrated from the measured data, interpolates the missing parts of the dose–response surface. (A) measured data at 7 h and interpolated dose–response surface. (B) the unobserved dose–response after 24 h is predicted using the model since this time point is beyond the currently possible sampling interval that requires exponential growth. Note that the measurements with highest GFP+ abundance correspond to the unrepressed cells.

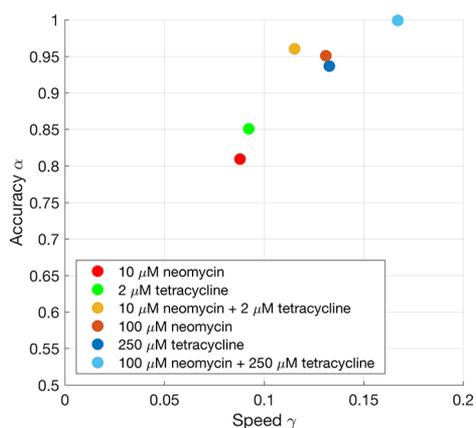


Figure 11. NOT/NOR gate performance characterized through position in the performance space for various combinations of repressor concentrations. Inducer concentration was 10 nM β -estradiol.

an area around it determined by the tested repressor concentrations to visualize the performance constraints that cannot be overcome by the design of the 5'-NeoTc gate.

The reporter measurements presented in Figures 3 and 5 imply differences in the repression kinetics of tetracycline and neomycin and a closer look at the concurrent events for the case of a double-repression would be desirable. However, due to the stability and hence low-pass filtering of the reporter, quantifications based solely on the experimental data are unfeasible. The model-based reconstruction provides mRNA estimates at their comparably faster time-scales, allowing us to better dissect the temporal sequence of events at the onset of repression after 2 h induction. By this only newly transcribed mRNAs in the presence of repressors are captured and used to calibrate the model. Figure 12 shows the evolution of occupancies for the different mRNA subpopulations (*i.e.*, free, single-bound, double-bound) at the onset of repression. In accordance to protein levels (Figures 3 and 5), free mRNA abundances decrease faster in the presence of tetracycline as compared to neomycin (Figure 12A,B). Increase of double-bound mRNAs is slowed by transient occupancy of the mRNA

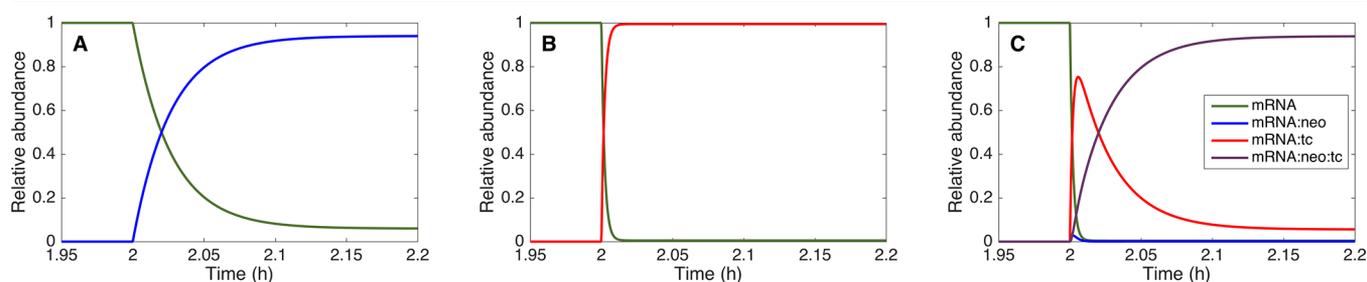


Figure 12. Relative abundance of mRNA subpopulations after addition of repressors at +2 h. (A) 100 nM neomycin. (B) 250 nM tetracycline. (C) 100 nM neomycin and 250 nM tetracycline. Due to the faster binding kinetics of tetracycline, the majority of mRNA molecules binds tetracycline immediately after addition. Subsequently, the balance progressively shifts to double-bound mRNA due to the slower binding kinetics of neomycin.

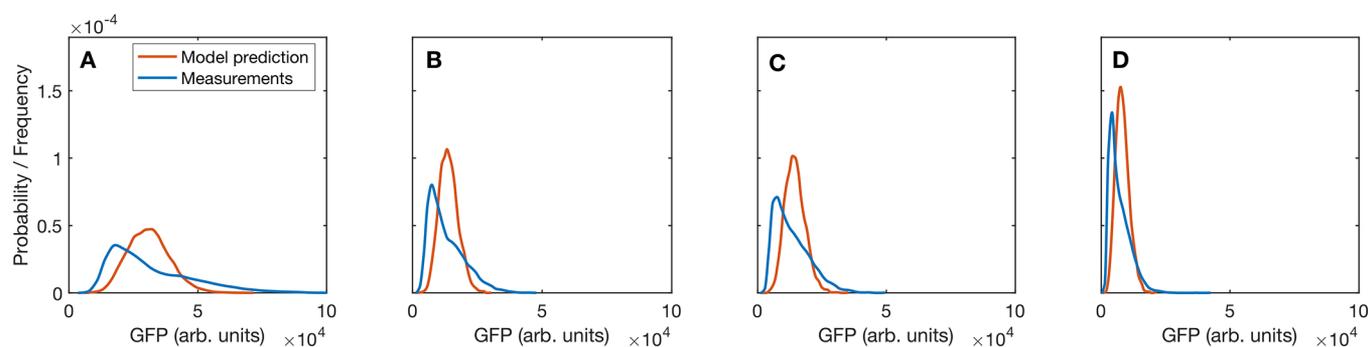


Figure 13. Predicted GFP+ abundance for fixed cell size after 7 h of induction. (A) No repressors. (B) 100 μM neomycin. (C) 250 μM tetracycline. (D) 100 μM neomycin + 250 μM tetracycline. Repressors were added after 2 h induction with 10 nM β -estradiol.

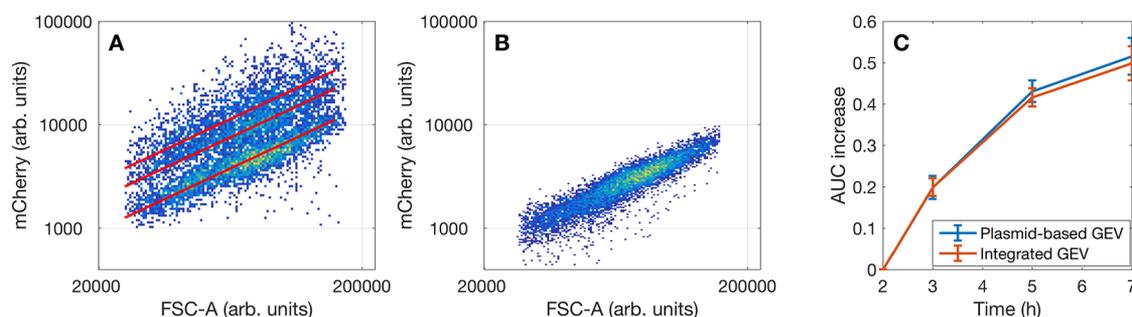


Figure 14. Influence of GEV variability on gate performance. GEV vs forward scatter for plasmid-based GEV (A) and for chromosomally integrated GEV (B). For the plasmid-based case, we see at least two subpopulations. The distance on the mCherry-axis between the red lines corresponds to doubling or tripling of GEV abundance, which is the simplest assumption for the behavior when 2 or 3 plasmid copies are present in a cell. For chromosomally integrated GEV, only the subpopulation corresponding to a single copy of the plasmid remains. (C) AUC increase over time for plasmid-based and for integrated GEV for induction at 2.5 nM β -estradiol and repression with 100 μM neomycin and 250 μM tetracycline. The differences between the two are within the margin of error.

with tetracycline alone as neomycin exhibits relatively slow association kinetics (Figure 12C). For the same reason, no single-bound mRNAs by neomycin are found in the presence of tetracycline. The overall abundance of mRNAs encoding the NeoTc gate is assumed to be unaffected by the integration of the two riboswitches in absence and presence of their cognate ligands as has been shown in earlier studies.^{36,54}

Our model also allowed us to investigate the noise in GFP+ expression which is left after removing the effect of cell size. Hence, we performed an *in silico* homogenization or cell-cycle synchronization. The resulting distributions are shown in Figure 13 for several combinations of inducer and repressor concentrations. It indicates that a considerable variability due to stochastic gene expression remains after such homogenization that is inherent to the system and cannot be decreased by a further reduction of gene copy number for instance.

Computer-Aided Redesign. Having general performance measures in place, we then asked whether performance can be further improved by a redesign of the inducible expression system, for instance by changing specific reaction rate constants that are amenable to experimental modification. As the width of the distribution and in particular the distributional overlap between the Boolean states are critical to achieve high AUC we specifically investigated, whether reduction of extrinsic noise results in improved gate performance. Although we used ARS/CEN plasmids that should be ideally maintained as a single copy within the cell, our data and previous publications suggest that cells do not always carry just a single plasmid, but multiple.⁵⁵ In particular, Figure 14A provides evidence for a considerable population of cells carrying two and three

plasmids. This plasmid heterogeneity leads to a broadened distribution of the GEV transcription factor. Generally, through noise propagation a broadened GEV distribution will give rise to a broadened distribution of the gate output, and hence, deteriorating the gate's performance. Using our model, we queried the sensitivity of the gate's AUC with respect to changes in GEV heterogeneity. Practically, we computed α for different standard deviations of GEV heterogeneity in the model and found that variability in GEV abundance had little effect on α and hence did not contribute much to the output variability. Thus, the model suggests that redesigns targeting a reduction of GEV variability through, for instance, a chromosomal integration of GEV, would not yield a performance increase. To test this prediction, we inserted the plasmid-derived GEV expression cassette into the *his3* locus of yeast cells already bearing the NeoTc gate integration and repeated the above single-cell device characterization. Figure 14B shows the reduction in GEV heterogeneity after integration. In terms of performance the device behaves the same under the same conditions (Figure 14C). Thus, the integration of GEV and thereby the reduction of its noisy expression has no influence on the AUC increase of the gate.

The results of the device characterization are summarized in Supporting Information S2.6. On top of the AUC analysis the model also provides an explanation for this counterintuitive insensitivity of gate performance with respect to GEV variability. A plain saturating effect through GEV at the promoter would be incompatible with the observed gradual dose dependency of gate output with respect to different induction levels (e.g., Figure 10) and the model indeed

determined that the insensitivity arises because β -estradiol is the limiting factor, while GEV is in excess for a very broad range of values, exceeding the threshold needed for maximum GFP+ expression even with a single gene copy.

In conclusion, we demonstrate a novel approach to characterize a genetic device by capturing its transient dynamics. The collected time-lapse data provide a solid source for the design and calibration of stochastic kinetic models at the single-cell level. Specification of gate performance accounts for the fold-change of the Boolean states and cell-to-cell variability. To this end, we developed a new general performance measure for genetic devices based on ROC analysis. Further, using the transient data we built a minimal stochastic model capable to correctly reproduce various induction and repression settings and to complete experimental inaccessible data *in silico*. The proposed model is able to capture surprisingly fine distributional features on a short temporal scale. Albeit it is not yet suited to identify context-dependencies of adjacent elements, it may be used to predict redesign strategies of the greater system by incorporating extrinsic noise parameters. Our approach intends to initiate the in-depth investigations of genetic circuits and especially their building blocks to enhance their performance at the single-cell level.

MATERIALS AND METHODS

Algorithms for Calibration and *In Silico* Characterization. All computational aspects of the work are detailed in the [Supporting Information S1](#).

Cloning. DNA-vector maps, integration parts and PCR primer lists are provided upon request. A brief description of all steps involved in the construction of all DNA parts and devices is summarized in the following text.

Custom oligonucleotides were synthesized by Sigma-Aldrich. Plasmids were constructed by standard cloning techniques using PCR with Q5 Polymerase (NEB) and restriction and ligation reactions with HF restriction enzymes (NEB) and T4-DNA ligase (NEB). For Gibson assembly, instructions provided at [neb.com](#) were followed and the homemade mix from [openwetware.org/wiki/Gibson_Assembly](#) was prepared.

P_{ADH1} -Driven Device. A modified version of the plasmid pWHE601 was used as the backbone vector to clone the riboswitch parts downstream of the *ADH1* promoter, generating the NOR and NOT devices shown in [Figure 1](#). Plasmids were transformed into the electrocompetent *Escherichia coli* strain TOP10 grown in LB medium supplemented with 100 $\mu\text{g}/\text{mL}$ ampicillin.

P_{GALT} -Driven Device. The plasmid pSP212-pTEF1-GLC-tCYC1 expresses the GEV transcription factor that is C-terminally coupled to mCherry and was transformed into competent yeast containing the NeoTc device integrated into the *ade2* locus.

Genome Manipulations. Genotype RS453 α : *mat α ade2-1 trp1-1 can1-100 leu2-3 his3-1 ura3-52*.

Genotype CEN.PK111-27B: *mat α leu2-3 trp1-1*.

Genomic integrations were performed by the CRISPR/Cas9-based method CASEMBL.⁴⁶ The yeast shuttle plasmids p414-pTEF1-Cas9-tCYC1 (plasmid #83946)⁵⁶ and p426-pSNR52-gRNA.CAN1.Y-tSUP4 (plasmid #43803)⁵⁷ were ordered from [addgene.org](#). p414 (TRP marker) expresses the Cas9 endonuclease and p426 (URA marker) harbors the gRNA expression cassette that was transferred to p425 (LEU marker) as the URA3 locus presented a putative integration site and the CEN.PK111-27B strain is Ura3⁺. Gibson assembly was used to

simultaneously ligate the gRNA expression cassette amplified in one part from p426, and broken in two by a ClaI restriction site, with one of the three gRNAs targeting the ADE2, HIS3 and URA3 loci and the linearized p425 plasmid, excluding the original gRNA. Initially, as described in the publication from Jakoucinas *et al.*, yeast strain CEN.PK 111-27B was used for integrations into the ADE2 locus, resulting in null mutants accumulating a red pigment that absorbs green fluorescence light. Experiments with GFP+ as a reporter protein were thus found to be impracticable. We therefore switched the strain to RS453 α that is a null mutant with respect to *ade2*, but passaged long enough to have a white phenotype due to mutations of enzymes in the upstream de novo biosynthesis of purine nucleotides, preventing the occurrence of a red pigment. For integrations into the *his3* locus strain RS453 α was taken as well. DNA parts for integration were assembled by PCR, purified, and EtOH precipitated to increase their concentration. 2–4 pmol/part were mixed with 1 μg p414-Cas9 and 1 μg p425-gRNA. Transformations were conducted according to the instructions provided with the Frozen-EZ Yeast Transformation II Kit from Zymo Research and spread on appropriate SCD drop-out agar plates (see below for details on growth medium and protocol). As may be deduced from the genotype of this strain, neither integration site could be exploited for a phenotypic screening to detect locus destruction. However, we found, in accordance with the developers of the method, that integration efficiency based on yeast colony PCR is between 80 and 100% for both loci, using primer pairs within and upstream or downstream of the integration sites.

Cultivation of Yeast. The Frozen-EZ Yeast Transformation II Kit was used for the preparation of competent yeast cells from strain RS453 α and their subsequent transformation on appropriate SCD drop-out plates (0.2% YNB w/o AA (Difco), 0.15% Yeast Synthetic Drop-out Medium w/o uracil, leucine and tryptophan (Sigma-Aldrich), 2% dextrose (Carl Roth) and 1.8% agar (Carl Roth). Uracil, leucine and tryptophan were added as indicated on the complete formula of the Yeast Synthetic Drop-out Medium.

After 3 d growth at 30 °C in a humidified incubator, single colonies were cultured in 5 mL SCD drop-out medium for 24 h, then diluted to an OD₆₀₀ of approximately 0.05 (flow cytometry) or 0.01 (microfluidic measurements) and grown for additional 10 h (flow cytometry) or 5 h (microfluidic measurements) in 5 mL SCD drop-out medium to obtain actively growing cells in early to mid log phase for the subsequent experiments.

Cells for bulk measurements ([Figure 1](#)) were grown in 3 mL SCD drop-out medium without repressors or supplemented with 100 μM neomycin and/or 250 μM tetracycline in 12-well plates for 24 h and analyzed in two independent triplicates. Fluorescence of all cultures was OD₆₀₀- and background-corrected and is displayed relative to a positive control that is not riboswitch-controlled, but treated equally.

Cytometry. All cytometry measurements were performed on a CytoFlex S instrument from Beckman Coulter equipped with a 488 nm laser and a 561 nm laser for excitation of GFP+ and mCherry, respectively. Emission light was bandpass-filtered at 510/20 nm or 610/20 nm.

Cells taken from the 10 h preculture were split to 12-well plates at an OD₆₀₀ of 0.05 and cultured for 10 h at 30 °C in 3 mL SCD drop-out medium. This protocol generates yeast cultures that are in the mid log phase at the time point of

induction and grow until late log phase over the time course of the experiment (Supporting Information S2.3). Four separate experiments with regard to the timing of repression at -1 h, 0 h, $+1$ h and $+2$ h relative to the induction with 10 nM β -estradiol at 0 h were performed. 10 – 30 μ L samples were drawn hourly for 7 h (Figure 3). Experiments to investigate dose-dependencies of the gate on repressor and inducer concentrations were conducted with 2.5 , 5 , or 10 nM β -estradiol and 2 , 10 , 50 , and 100 μ M neomycin and 0.5 , 2 , 10 , 250 μ M tetracycline. Here, repressors were added $+2$ h relative to induction and 10 – 30 μ L samples were drawn at 0 , 2 , 3 , 5 and 7 h (Figure 5).

Time-Lapse Microscopy. For *in vivo* time-lapse experiments, cells from the 5 h preculture were trapped on a PDMS-based microfluidic chip, inspired by the Alcatraz⁵⁸ chip. Single yeast cells were kept in place in a continuous flow for several hours and images were recorded every 10 min at 3 different focal planes. Imaging was performed on a Nikon Eclipse TI with a $100\times$ NA 1.45 objective, ORCA Flash4.0 camera (Hamamatsu) with SpectraX light engine (Lumencor) at $470/24$ nm for GFP+ and $575/25$ nm for mCherry, kept at 30 °C with a incubator box (life imaging services). Concentrations of inducer and repressors were used according to the cytometry measurements: 10 nM β -estradiol, 100 μ M neomycin and 250 μ M tetracycline. Cells were induced for 2 h, then exposed to single or both repressors. Single cell traces were extracted with a FIJI/Matlab script.

■ ASSOCIATED CONTENT

■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acssynbio.6b00322](https://doi.org/10.1021/acssynbio.6b00322).

Description of hierarchical stochastic model; description of inference procedure; posterior distribution plots; reporter gene assay of the alternative NOR gate TcNeo; population analysis of the constitutively expressed NeoTc gate by flow cytometry; growth curve of the NeoTc gate; ligand influence on GFP and GEV expression; variability of flow cytometry results over three biological replicates; influence of GEV variability on GFP+ variability (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: heinz.koeppl@bcs.tu-darmstadt.de.

*E-mail: bsuess@bio.tu-darmstadt.de.

ORCID

Beatrix Suess: [0000-0001-8666-6716](https://orcid.org/0000-0001-8666-6716)

Author Contributions

§C.S. and L.B. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Landesoffensive für wissenschaftliche Exzellenz (LOEWE; initiative to increase research excellence in the state of Hessen, Germany) as part of the LOEWE Schwerpunkt CompuGene, the European Union (FP7 Promys) and DFG SFB902 A2.

■ REFERENCES

- (1) Purnick, P. E. M., and Weiss, R. (2009) The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* **10**, 410–422.
- (2) Nandagopal, N., and Elowitz, M. B. (2011) Synthetic biology: integrated gene circuits. *Science* **333**, 1244–1248.
- (3) Groher, F., and Suess, B. (2014) Synthetic riboswitches—a tool comes of age. *Biochim. Biophys. Acta, Gene Regul. Mech.* **1839**, 964–973.
- (4) Nielsen, A. A. K., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., Ross, D., Densmore, D., and Voigt, C. A. (2016) Genetic circuit design automation. *Science* **352**, aac7341.
- (5) Redden, H., Morse, N., and Alper, H. S. (2014) The synthetic biology toolbox for tuning gene expression in yeast. *FEMS Yeast Res.* DOI: [10.1111/1567-1364.12188](https://doi.org/10.1111/1567-1364.12188).
- (6) Brophy, J. A. N., and Voigt, C. A. (2014) Principles of genetic circuit design. *Nat. Methods* **11**, 508–520.
- (7) Lee, S., Lim, W. A., and Thorn, K. S. (2013) Improved blue, green, and red fluorescent protein tagging vectors for *S. cerevisiae*. *PLoS One* **8**, e67902.
- (8) Karginov, F. V., and Hannon, G. J. (2010) The CRISPR system: small RNA-guided defense in bacteria and archaea. *Mol. Cell* **37**, 7–19.
- (9) Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., and Lim, W. A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183.
- (10) Chin, J. W. (2014) Expanding and reprogramming the genetic code of cells and animals. *Annu. Rev. Biochem.* **83**, 379–408.
- (11) Cameron, D. E., Bashor, C. J., and Collins, J. J. (2014) A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12**, 381–390.
- (12) Winkler, W. C., Cohen-Chalamish, S., and Breaker, R. R. (2002) An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 15908–15913.
- (13) Winkler, W., Nahvi, A., and Breaker, R. R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952–956.
- (14) Breaker, R. R. (2012) Riboswitches and the RNA world. *Cold Spring Harbor Perspect. Biol.* **4**, a003566.
- (15) Serganov, A., and Patel, D. J. (2012) Molecular recognition and function of riboswitches. *Curr. Opin. Struct. Biol.* **22**, 279–286.
- (16) Peselis, A., and Serganov, A. (2014) Themes and variations in riboswitch structure and function. *Biochim. Biophys. Acta, Gene Regul. Mech.* **1839**, 908–918.
- (17) Eitzel, M., and Mörl, M. (2017) Synthetic Riboswitches - from Plug and Pray towards Plug and Play. *Biochemistry* **56**, 1181.
- (18) Berens, C., Groher, F., and Suess, B. (2015) RNA aptamers as genetic control devices: the potential of riboswitches as synthetic elements for regulating gene expression. *Biotechnol. J.* **10**, 246–257.
- (19) McKeague, M., Wong, R. S., and Smolke, C. D. (2016) Opportunities in the design and application of RNA for gene expression control. *Nucleic Acids Res.* **44**, 2987–2999.
- (20) Chappell, J., Watters, K. E., Takahashi, M. K., and Lucks, J. B. (2015) A renaissance in RNA synthetic biology: new mechanisms, applications and tools for the future. *Curr. Opin. Chem. Biol.* **28**, 47–56.
- (21) Ausländer, S., Stücheli, P., Rehm, C., Ausländer, D., Hartig, J. S., and Fussenegger, M. (2014) A general design strategy for protein-responsive riboswitches in mammalian cells. *Nat. Methods* **11**, 1154–1160.
- (22) Weigand, J. E., Sanchez, M., Gunnesch, E.-B., Zeiher, S., Schroeder, R., and Suess, B. (2008) Screening for engineered neomycin riboswitches that control translation initiation. *RNA* **14**, 89–97.
- (23) Liang, J. C., Chang, A. L., Kennedy, A. B., and Smolke, C. D. (2012) A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity. *Nucleic Acids Res.* **40**, e154.
- (24) Klausner, B., Atanasov, J., Siewert, L. K., and Hartig, J. S. (2015) Ribozyme-based aminoglycoside switches of gene expression engi-

neered by genetic selection in *S. cerevisiae*. *ACS Synth. Biol.* 4, 516–525.

(25) Beisel, C. L., and Smolke, C. D. (2009) Design principles for riboswitch function. *PLoS Comput. Biol.* 5, e1000363.

(26) Espah Borujeni, A., Mishler, D. M., Wang, J., Huso, W., and Salis, H. M. (2016) Automated physics-based design of synthetic riboswitches from diverse RNA aptamers. *Nucleic Acids Res.* 44, 1–13.

(27) Wachsmuth, M., Findeiß, S., Weissheimer, N., Stadler, P. F., and Mörl, M. (2013) De novo design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Res.* 41, 2541–2551.

(28) Green, A. A., Silver, P. A., Collins, J. J., and Yin, P. (2014) Toehold switches: de-novo-designed regulators of gene expression. *Cell* 159, 925–939.

(29) Felletti, M., Stifel, J., Wurmthaler, L. A., Geiger, S., and Hartig, J. S. (2016) Twister ribozymes as highly versatile expression platforms for artificial riboswitches. *Nat. Commun.* 7, 12834.

(30) Callura, J. M., Cantor, C. R., and Collins, J. J. (2012) Genetic switchboard for synthetic biology applications. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5850–5855.

(31) Win, M. N., and Smolke, C. D. (2008) Higher-order cellular information processing with synthetic RNA devices. *Science* 322, 456–460.

(32) Wieland, M., and Hartig, J. S. (2008) Improved aptazyme design and in vivo screening enable riboswitching in bacteria. *Angew. Chem., Int. Ed.* 47, 2604–2607.

(33) Bowsher, C. G., and Swain, P. S. (2012) Identifying sources of variation and the flow of information in biochemical networks. *Proc. Natl. Acad. Sci. U. S. A.* 109, E1320–8.

(34) Cardinale, S., and Arkin, A. P. (2012) Contextualizing context for synthetic biology—identifying causes of failure of synthetic biological systems. *Biotechnol. J.* 7, 856–866.

(35) McIsaac, R. S., Silverman, S. J., McClean, M. N., Gibney, P. A., Macinskas, J., Hickman, M. J., Petti, A. A., and Botstein, D. (2011) Fast-acting and nearly gratuitous induction of gene expression and protein depletion in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 22, 4447–4459.

(36) Hanson, S., Berthelot, K., Fink, B., McCarthy, J. E. G., and Suess, B. (2003) Tetracycline-aptamer-mediated translational regulation in yeast. *Mol. Microbiol.* 49, 1627–1637.

(37) Schneider, C., and Suess, B. (2016) Identification of RNA aptamers with riboswitching properties. *Methods* 97, 44–50.

(38) Scholz, O., Thiel, A., Hillen, W., and Niederweis, M. (2000) Quantitative analysis of gene expression with an improved green fluorescent protein. *Eur. J. Biochem.* 267, 1565–1570.

(39) Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene* 234, 187–208.

(40) Dvir, S., Velten, L., Sharon, E., Zeevi, D., Carey, L. B., Weinberger, A., and Segal, E. (2013) Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 110, E2792–E2801.

(41) Wittmann, A., and Suess, B. (2012) Engineered riboswitches: Expanding researchers' toolbox with synthetic RNA regulators. *FEBS Lett.* 586, 2076–2083.

(42) Lee, M. E., DeLoache, W. C., Cervantes, B., and Dueber, J. E. (2015) A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol.* 4, 975–986.

(43) Lee, J. W., Gyorgy, A., Cameron, D. E., Pyenson, N., Choi, K. R., Way, J. C., Silver, P. A., Del Vecchio, D., and Collins, J. J. (2016) Creating Single-Copy Genetic Circuits. *Mol. Cell* 63, 329–336.

(44) Munsy, B., Trinh, B., and Khammash, M. (2009) Listening to the noise: random fluctuations reveal gene network parameters. *Mol. Syst. Biol.* 5, 318.

(45) Fröhlich, F., Thomas, P., Kazerooni, A., Theis, F. J., Grima, R., and Hasenauer, J. (2016) Inference for Stochastic Chemical Kinetics Using Moment Equations and System Size Expansion. *PLoS Comput. Biol.* 12, e1005030.

(46) Jakociūnas, T., Rajkumar, A. S., Zhang, J., Arsovska, D., Rodriguez, A., Jendresen, C. B., Skjoldt, M. L., Nielsen, A. T., Borodina, I., Jensen, M. K., and Keasling, J. D. (2015) CasEMBLR: Cas9-

Facilitated Multiloci Genomic Integration of in Vivo Assembled DNA Parts in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 4, 1226–1234.

(47) Kay, S. M. (1998) *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice-Hall PTR.

(48) Lasko, T. A., Bhagwat, J. G., Zou, K. H., and Ohno-Machado, L. (2005) The use of receiver operating characteristic curves in biomedical informatics. *J. Biomed. Inf.* 38, 404–415.

(49) Macmillan, N. A., and Creelman, C. D. (2004) *Detection Theory: A User's Guide*, Taylor & Francis.

(50) Zechner, C., Ruess, J., Krenn, P., Pelet, S., Peter, M., Lygeros, J., and Koepl, H. (2012) Moment-based inference predicts bimodality in transient gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 109, 8340–8345.

(51) Zechner, C., Unger, M., Pelet, S., Peter, M., and Koepl, H. (2014) Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings. *Nat. Methods* 11, 197–202.

(52) Van Kampen, N. G. (1992) *Stochastic Processes in Physics and Chemistry*, Elsevier Science.

(53) Finkenstädt, B., Woodcock, D. J., Komorowski, M., Harper, C. V., Davis, J. R. E., White, M. R. H., and Rand, D. A. (2013) Quantifying intrinsic and extrinsic noise in gene transcription using the linear noise approximation: An application to single cell data. *Ann. Appl. Stat.* 7, 1960–1982.

(54) Grate, D., and Wilson, C. (2001) Inducible regulation of the *S. cerevisiae* cell cycle mediated by an RNA aptamer–ligand complex. *Bioorg. Med. Chem.* 9, 2565–2570.

(55) Clarke, L., and Carbon, J. (1980) Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature* 287, 504–509.

(56) Stovicek, V., Borodina, I., and Forster, J. (2015) CRISPR–Cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains. *Metabolic Engineering Communications* 2, 13–22.

(57) DiCarlo, J. E., Norville, J. E., Mali, P., Rios, X., Aach, J., and Church, G. M. (2013) Genome engineering in *Saccharomyces cerevisiae* using CRISPR–Cas systems. *Nucleic Acids Res.* 41, 4336–4343.

(58) Crane, M. M., Clark, I. B. N., Bakker, E., Smith, S., and Swain, P. S. (2014) A microfluidic system for studying ageing and dynamic single-cell responses in budding yeast. *PLoS One* 9, e100042.

(59) Note that requiring nonoverlapping distributions is a restrictive (sufficient) condition for proper gate function on the single-cell level but is the only testable condition for snapshot data such as flow-cytometry data. In contrast, using live-cell imaging one could declare proper functioning of a gate within a single cell if repeated ON/OFF switching yields an acceptable AUC, *i.e.*, if fold-changes between the Boolean states are large and reproducible within the same cell. This does not necessarily imply nonoverlapping distributions on the population level.