nature biotechnology

Ribozyme-based insulator parts buffer synthetic circuits from genetic context

Chunbo Lou[1], Brynne Stanton[1], Ying-Ja Chen[1], Brian Munsky[2] & Christopher A Voigt[1]

Synthetic genetic programs are built from circuits that integrate sensors and implement temporal control of gene expression[1–4]. Transcriptional circuits are layered by using promoters to carry the signal between circuits. In other words, the output promoter of one circuit serves as the input promoter to the next. Thus, connecting circuits requires physically connecting a promoter to the next circuit. We show that the sequence at the junction between the input promoter and circuit can affect the input-output response (transfer function) of the circuit[5–9]. A library of putative sequences that might reduce (or buffer) such context effects, which we refer to as 'insulator parts', is screened in *Escherichia coli*. We find that ribozymes that cleave the 5′ untranslated region (5′-UTR) of the mRNA are effective insulators. They generate quantitatively identical transfer functions, irrespective of the identity of the input promoter. When these insulators are used to join synthetic gene circuits, the behavior of layered circuits can be predicted using a mathematical model. The inclusion of insulators will be critical in reliably permuting circuits to build different programs.

A fundamental principle of synthetic biology is that genetic parts can be independently characterized and used to predict their combined behavior[10,11]. This principle underpins the high-throughput fabrication of valuable parts[12], the storage of useful part information in a registry[13,14] and the use of these data for computer-aided design[15]. In practice, however, the functions of parts are influenced by their genetic, environmental and cellular context[16]. This limits the designability of genetic systems, because each part needs to be characterized in its final context.

The function of a part can be affected by the sequence of neighboring parts, and we refer to this effect as 'part-junction interference'. For example, a barcode part was found to contain a sequence similar to the –10 region of the constitutive promoter (http://2010.igem.org/Team:UC_Davis/notebook/c0051debug.html). Depending on the sequence of the upstream part, a promoter can be spontaneously formed at the junction between the barcode and neighboring part. Similarly, promoter parts are often defined by a relatively short (~50 bp) sequence, but regions 100 bp or more upstream can affect promoter strength[17], and the effect of remote sequences can be reduced by including an insulator region[18]. A strong hairpin sequence has also been shown to insulate the ribosome binding site (RBS) from the sequence of the 5′ UTR[19]. Recently, it was shown that the processing of mRNA using CRISPR RNA processing elements could be used to reliably maintain relative promoter strengths and reduce interference between genes within an operon[20].

Genetic circuits use biochemical interactions to implement a computational operation. Transcriptional circuits implement this operation on the level of promoters, where promoters serve as both the input and the output of the circuit. Thus, the connection of a new input to a circuit creates the possibility of junction interference. How the output of a logic gate changes as a function of the inputs is described by its transfer function. This should be an intrinsic property of the circuit, which is independent from the inducible system that is used to measure it. As part of building a program for an unrelated application, we characterized the transfer function of a circuit and serendipitously discovered the influence of part-junction interference. The circuit is a NOT gate where the signal of an input promoter is inverted by having it drive the expression of the CI repressor that turns off an output promoter (**Fig. 1a**)[21–24]. Because the transfer function is characterized using an inducible system, the raw data are in the form of concentrations (the input) and fluorescence (the output). So that the units of the inputs and outputs are the same, the activity of the inducible promoter is characterized independently (**Fig. 1a,b**) and used to report the transfer function[6].

Initially, the IPTG-inducible pTAC promoter was used to characterize the NOT gate. The output of the CI-repressible pOR1 promoter was measured as a function of [IPTG] using green fluorescent protein (gfp). The activity of pTAC was also measured as a function of [IPTG]. These two data sets were used to generate the transfer function of the NOT gate (**Fig. 1c**). Each point on the curve is the fluorescence measured using two genetic systems (pTAC-gfp and pTAC_cI_pOR1-gfp). This results in a cooperative function, where the output promoter activity declines as the input promoter turns on.

We characterized the transfer function again using the pBAD promoter to initiate transcription and drive the circuit. The activity of the pBAD promoter was first measured as a function of arabinose concentration. Then, the output of the pBAD-driven NOT gate was measured, and the two data sets were used to generate the transfer function (**Fig. 1c**). Although the qualitative trend was similar, the quantitative details differed greatly. The pBAD-driven gate showed a 20-fold higher basal state and a less cooperative transition. This result was surprising;

[1]Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [2]Center for Nonlinear Studies and the Computer, Computational, and Statistical Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA. Correspondence should be addressed to C.A.V. (cavoigt@gmail.com).

Received 25 May; accepted 19 September; published online 3 October 2012; doi:10.1038/nbt.2401

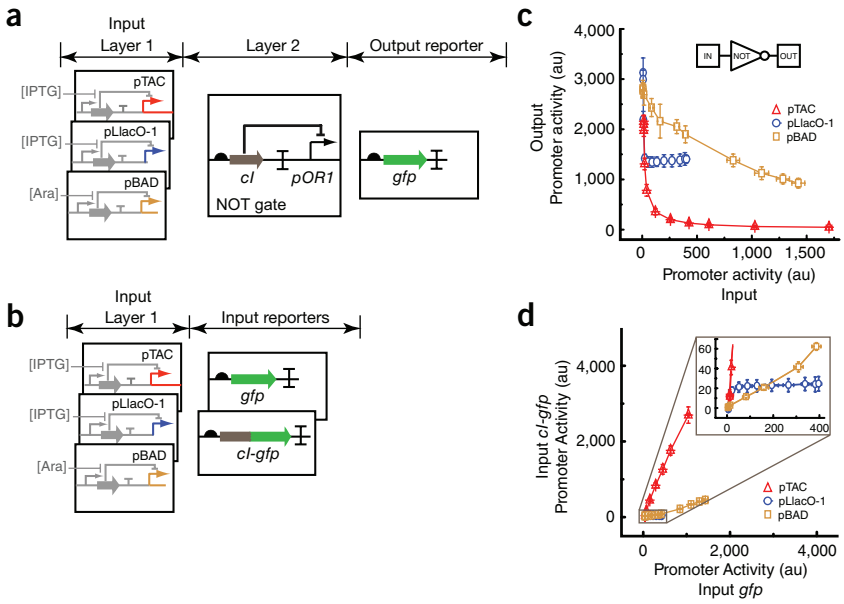

Figure 1 The transfer function of the NOT gate depends on the inducible system used to measure it. (**a**) Three inducible systems were connected to the NOT gate: pTAC, pLlacO-1 and pBAD. The NOT gate is based on the *cI* gene, which represses the pOR1 promoter. (**b**) The promoter activity of each inducible system is determined using gfp as a reporter alone and fused to the *cI* gene. (**c**) Data gathered using the constructs from parts **a** and **b** are used to determine the transfer function of the NOT gate. The transfer functions are shown as measured by the inducible systems of **a**. Each point represents one concentration of inducer that corresponds to expression from pOR1 (part (**a**)) and the output promoter of the inducible system (part (**b**)). The inducer concentrations are: pTAC, 0, 0.1, 1, 5, 10, 20, 30, 40, 50, 70, 100 μM IPTG; pLlacO-1, 0.1, 1, 10, 50, 100, 150, 200, 300, 500, 1,000, 2,000, 3,000 μM IPTG; and pBAD, 0, 0.1, 1, 2, 5, 7, 10, 12.5, 25, 37.5, 50, 62.5 mM Ara. (**d**) The fluorescence measured from each inducible promoter driving the expression of *gfp* is compared to *cI-gfp* (part (**b**)). The expression ratios are the slope of each line: 2.6 (pTAC) and 0.3 (pBAD). Expression of *cI-gfp* saturates at ~20 au (inset). As in part **c**, each point represents a single concentration of inducer. Error bars, mean ± s.d. obtained from at least three experiments performed on different days.

because the inputs are measured and reported as promoter activity, one would have expected the transfer functions to be identical.

The input promoter activity for both systems was measured using *gfp*, but the promoters controlled *cI* in the context of the circuit. Thus, we suspected that the expression ratios of CI to GFP were different under pBAD and pTAC control. To determine whether this was true, we constructed a fusion protein (*cI-gfp*) to monitor CI expression, and we compared it with GFP only (**Fig. 1b**). Indeed, these promoters yielded different ratios of CI-GFP to GFP expression: 0.3 for pBAD and 2.6 for pTAC (**Fig. 1d**). In other words, the identity of the upstream promoters caused the expression of CI-GFP to differ by an order of magnitude compared to GFP.

Whereas promoters technically end at the transcription start (+1) site, parts that are labeled as promoters often contain additional sequences downstream of position +1. This may be due either to the inclusion of additional regulatory elements or an unknown or misannotated start site. A potential explanation for the pBAD/pTAC discrepancy is that the two promoters generate transcripts with different 5′-UTR sequences on the mRNA. In particular, pTAC had a 27-nt leader sequence, whereas pBAD produced a 5-nt leader (**Supplementary Fig. 1**). To test the importance of these leader sequences, we used the pLlacO-1 promoter[25], which is induced by IPTG, but in which the operator has been moved upstream of the transcription site (**Supplementary Fig. 1**). The resulting transfer function

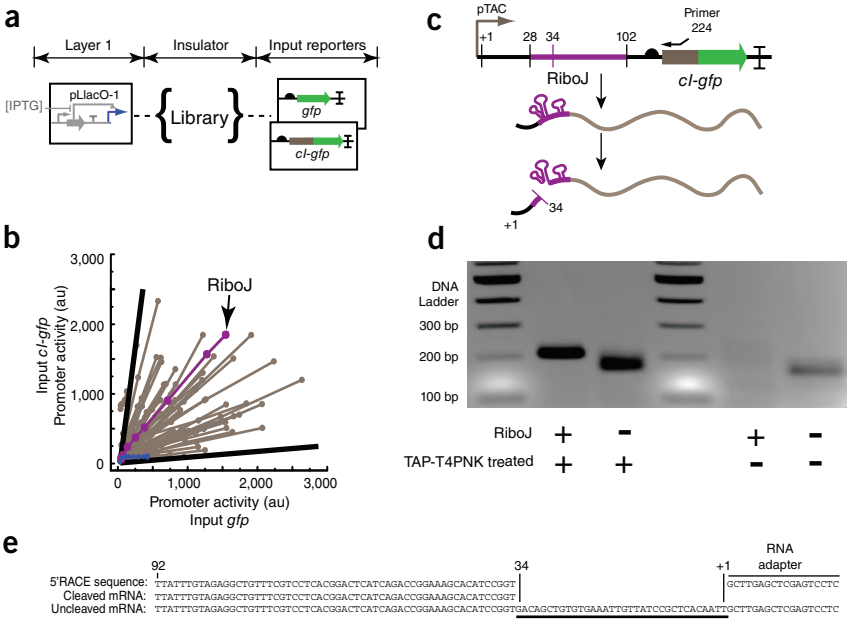

Figure 2 Screening the library of insulator parts. (**a**) The screen is based on the comparison of expression between *gfp* and *cI-gfp*. Each of the 54 putative insulators is inserted into both of these contexts, thus requiring 108 constructs. (**b**) Screening data are shown for the full library of insulators. Each line is a different insulator and each point is the fluorescence from the pair of constructs in part **a** at 0, 5, 50, 100, 300 or 1,000 μM IPTG. The solid black lines mark the upper (slope, 6.5) and lower (slope, 0.12) bounds of the distribution of insulators (pJ5J and DG131aRBS, respectively, in **Supplementary Table 2**). The blue curve corresponds to the behavior of pLlacO-1 without an insulator. (**c**) Under the pTAC promoter, RiboJ cleaves after the +34 nucleotide, thus removing the 5′ sequence from the promoter (+1 to +28). The primer used for the 5′RACE experiments is shown. (**d**) The agarose gel result for the cleaved mRNA and its controls. The gel result is for the amplified DNA samples that were reverse-transcribed from mRNA templates according to the 5′RACE protocol (Online Methods). In the presence of RiboJ, the TAP-T4PNK-untreated sample has two faint bands (lane 5), but the TAP-T4PNK-treated sample has one heavy dark band (lane 1). In the absence of RiboJ, the TAP-T4PNK-treated (lane 6) and TAP-T4PNK-untreated (lane 2) samples, respectively, have a weak or dark band. (**e**) 5′RACE sequencing data. All sequences are the complementary DNA sequences of RNA. The sequencing result of 5′RACE is for the DNA sample in lane 1 of **d** and read from an internal primer of the mRNA to the 5′-end. ‘RNA adapter’ is the reverse complementary DNA sequence of the RNA adaptor. The underlined sequence is removed from the mRNA by the ribozyme after the defined “G” site.

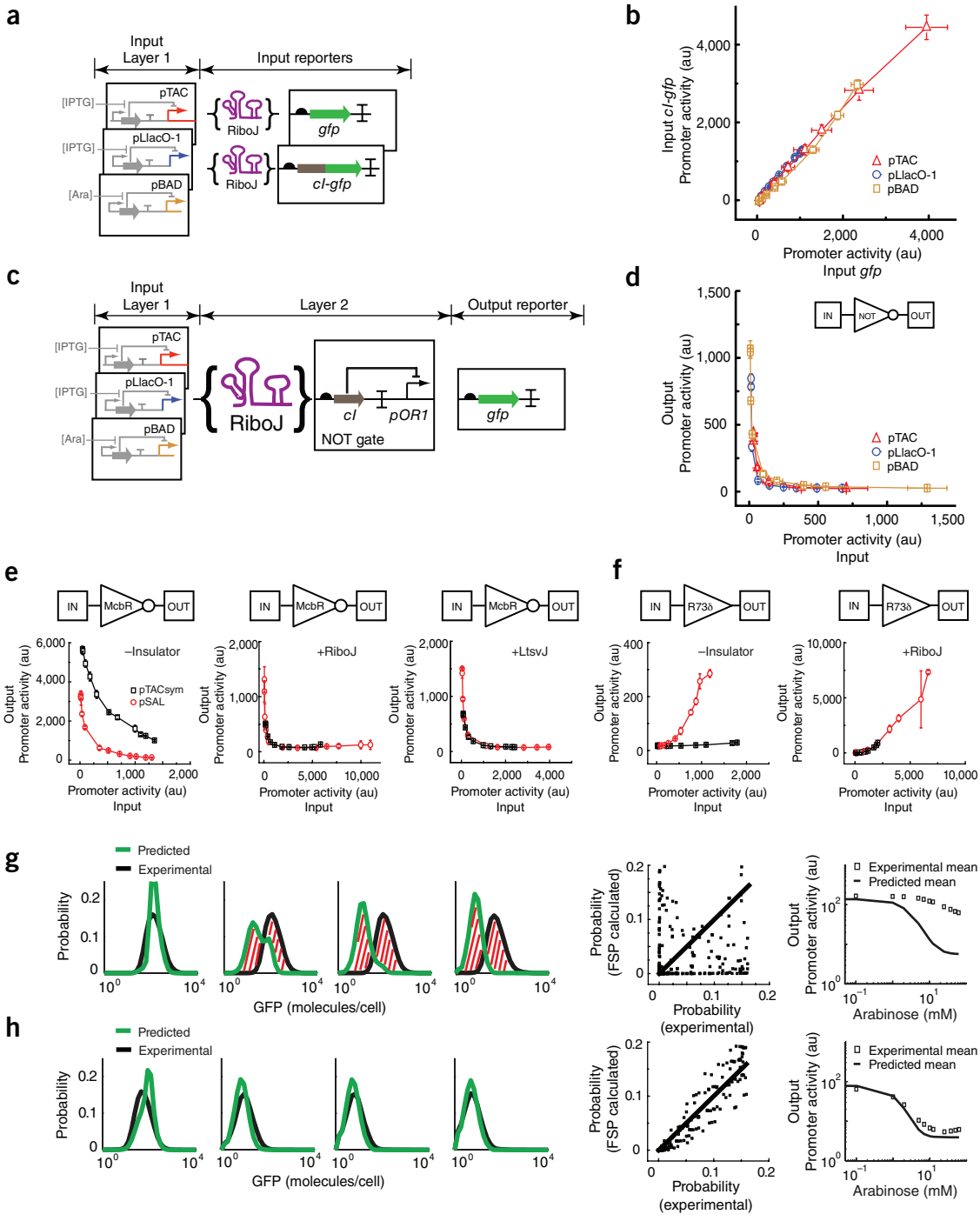

Figure 3 RiboJ and other insulators insulate the transfer function of the NOT and BUFFER gates. (**a**) The constructs from **Figure 1b** rebuilt to contain the RiboJ sequence between the output promoters of the inducible systems and the *gfp* and *cI-gfp* reporters. (**b**) The ratio between GFP and CI-GFP expression collapse onto a single curve for each inducible system. The concentrations of inducer are identical to those in **Figure 1c**. (**c**) The RiboJ insulator inserted upstream of the NOT gate for the three inducible systems. (**d**) The collapse of the data onto a single transfer function. The experiments are identical to those shown in **Figure 1c**. Error bars, mean ± s.d. of at least three experiments performed on different days. (**e**) The transfer functions of McbR NOT gate with and without the RiboJ and LtsvJ insulators. The transfer functions were measures using the pTACsym (red circles) and pSAL (black square) inducible systems. Their inducer concentrations are: pTACsym, 0, 0.1, 1, 5, 10, 20, 50, 100, 200, 500, 1,000, 2,000 μM IPTG and pSAL, 0, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100, 200 μM salicylate. The reporter gene of the two new gates is *yfp*. (**f**) The transfer functions of R73δ BUFFER gate with and without RiboJ. (**g**) The experimental data are compared to the predicted distribution when RiboJ is not included in the circuit. The inducer concentrations are (left to right): 0.1, 5, 10, 50 mM Ara. The hashed red lines show the difference between the predicted and measured distributions. Right: quantitative comparisons of distributions where the probability is divided into 30 bins for each curve. The data fit poorly to a line with slope of unity (R^2 = –0.30). (**h**) Same as **g**, but inclusion of RiboJ improves the predictability of the assembled circuit. The comparison between the predicted and experimental follows a linear fit (R^2 = 0.89). Far right, for **g** and **h**, the experimental mean and predicted mean of the probability distributions.

Table 1 Performance of ribozymes obtained via part mining

Name[a]	Sequence[b]	Slope[c] pBAD	Slope[c] pTAC
LtsvJ	AGTACGTCTGAGCGTGATACCCGCTCACTGAAGATGGCCCGGTAGGGCCGAAACGTACCTCTACAAATAATTTTGTTTAA	0.6 ± 0.1	0.6 ± 0.1
SccJ	AGATGCTGTAGTGGGATGTGTGTCTCACCTGAAGAGTACAAAAGTCCGAAACGGTATCCTCTACAAATAATTTTGTTTAA	1.1 ± 0.1	1.2 ± 0.2
RiboJ	AGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAA	1.4 ± 0.3	1.3 ± 0.3
SarJ	AGACTGTCGCCGGATGTGTATCCGACCTGACGATGGCCCAAAAGGGCCGAAACAGTCCTCTACAAATAATTTTGTTTAA	2.2 ± 0.2	2.4 ± 0.2
PlmJ	AGTCATAAGTCTGGGCTAAGCCCACTGATGAGTCGCTGAAATGCGACGAAACTTATGACCTCTACAAATAATTTTGTTTAA	1.6 ± 0.2	1.8 ± 0.1
VtmoJ	AGTCCGTAGTGGATGTGTATCCACTCTGATGAGTCCGAAAGGACGAAACGGACCTCTACAAATAATTTTGTTTAA	1.0 ± 0.1	1.3 ± 0.1
ChmJ	AGAAGAGGTCGGCACCTGACGTCGGTGTCCTGATGAAGATCCATGACAGGATCGAAACCTCTTCCTCTACAAATAATTTTGTTTAA	1.9 ± 0.1	2.4 ± 0.3
ScvmJ	AGTACTGTCGCCAGACGTGGACCCGGCCTGATGAGTCCGAAAGGACGAAACAGTACCTCTACAAATAATTTTGTTTAA	0.8 ± 0.1	1.4 ± 0.1
SltJ	AGGACGTATGAGACTGACTGAAACGCCGTCTCACTGATGAGGCCATGGCAGGCCGAAACGTCCCTCTACAAATAATTTTGTTTAA	2.2 ± 0.5	1.4 ± 0.2
PlmvJ	AGGAAGAGTCTGTGCTAAGCACACTGACGAGTCTCTGAGATGAGACGAAACTCTTCCCTCTACAAATAATTTTGTTTAA	2.0 ± 0.5	5.9 ± 1.2

[a]The names are based on the ribozyme names from references 47,48. Each insulator also contains an additional hairpin to expose the RBS. [b]The green residues form stem I, the blue residues form stem II and the red residues are the conserved catalytic core. The pink residues represent the stem loop of the added hairpin. The sequence downstream of the hairpin is: actagaaggag-gaaaaaaatg for *gfp*; actagaaccatgaaaaaagtg for *cI-gfp* (**Supplementary Fig. 2**). [c]The slope of the expression of *gfp* and *cI-gfp* genes under the control of two promoters. The slopes are calculated as the average of at least three experiments on different days. The average error is 14%. The detailed measurements are provided in **Supplementary Figure 9**.

was very different from that of either pBAD or pTAC in that it sharply deactivated as a function of input promoter activity, but rapidly saturated at a high output activity (**Fig. 1c**, blue). With the pLlacO-1 promoter, there was a nonlinear relationship between CI-GFP and GFP (**Fig. 1d,** inset). Additional experiments focusing on perturbations of the sequences around the part junction were performed to further elucidate the source of the interference (**Supplementary Figs. 2–4**). Constructs were tested where we systematically made mutations near the junction, truncated the *cI* gene and varied the 5′-UTR. These experiments revealed that the interference was not due to the promoter sequence near the transcriptional start site (**Supplementary Fig. 2**), or to the function of CI (**Supplementary Fig. 3**). The 5′-UTR (**Supplementary Figs. 4**) was required for part-junction interference, and the interference was at the level of transcription, as measured using qRT-PCR (**Supplementary Fig. 5**).

Based on these observations, we collated a library of 54 different 5′-UTR sequences that we thought might function as insulators (**Supplementary Table 1**) and screened this library to identify insulator parts that disrupted part-junction interference. We screened for insulators that either removed the saturation of CI-GFP (that is, linearized the curve), or prevented the transcribed promoter sequence from affecting translation (that is, fixed the ratio of CI-GFP to GFP). Candidate insulators were first inserted between pLlacO-1 and *cI-gfp* or *gfp* and tested for their ability to linearize the curve (**Fig. 2a**). Most of the spacers in the library (48 out of 54) rescued the linear relationship between CI-GFP and GFP (**Fig. 2b**). However, the slopes of the lines varied, ranging from 0.12 to 6.5 (**Supplementary Table 2**).

An insulator (RiboJ) that linearized the ratio of CI-GFP and GFP expression and fixed the CI-GFP:GFP ratio was selected for further study. The insulating capability of RiboJ is directly attributable to its biochemical function. RiboJ comprises the sTRSV- ribozyme[26] with an additional 23-nt hairpin[18] immediately downstream to help expose the RBS[19]. We define the insulator part RiboJ as being the DNA sequence containing both the ribozyme and hairpin. sTRSV autocatalytically cleaved the mRNA at a defined residue, which had the effect of cleaving extraneous RNA leaders that arose from transcription from promoters with different start sites (**Fig. 2c**). We performed rapid amplification of cDNA 5′-ends (5′RACE) (**Fig. 2d**) to identify the unknown sequence at the 5′-end of the mRNA. This confirmed that the 5′-region had been removed, consistent with the action of an active ribozyme (**Fig. 2e**).

The ratio of CI-GFP to GFP expression was measured for different inducible promoters containing RiboJ (**Fig. 3a**) and compared to the curves generated for control promoters lacking RiboJ (**Fig. 1d**). For each promoter, RiboJ collapsed these ratios to a single promoter-independent curve (**Fig. 3b**). The ratio generated by pTAC decreased from 2.6 to 1.1 and the ratio from pBAD increased from 0.3 to 1.2. The pLlacO-1 promoter recovered a linear relationship between CI-GFP and GFP and the ratio became 1.3. These results indicate that the insulator sequence eliminated the impact of a part-junction sequence that was added by a promoter to the 5′-UTR. Each input promoter added a different sequence to the 5′-UTR which was cleaved off by the ribozyme (**Fig. 2c**).

Next, we tested the ability of RiboJ to eliminate the interference mediated by the input promoter such that its identify was decoupled from the function of the NOT gate (**Fig. 3c**). All of the transfer functions collapsed onto the same curve (**Fig. 3d**). However, when we inactivated the cleavage activity of RiboJ, the NOT gate's transfer functions generated by pTAC and pBAD diverged again (**Supplementary Fig. 6**). Circuit dynamics were measured and the influence of RiboJ was found to be independent of time (**Supplementary Figs. 7–8**). Thus, with the inclusion of an insulator, the transfer function became an intrinsic property of the circuit, meaning that the output promoter depended only on the activity of the input promoter parts, and not the sequence at the part junction. This was in sharp contrast with the performance of the circuit without the insulator (**Fig. 1c**). This function held for other genetic circuits we tested, including a NOT gate based on a different repressor (McbR) and a BUFFER gate based on a phage activator (R73δ) (**Fig. 3e,f**). These data were gathered using different inducible promoters (pSal and pTACsym) and a different reporter (yellow fluorescent protein), underscoring the broad utility of the insulator.

Genetic programs require the functional connection of multiple genetic circuits. Reusing the same 75-bp insulator for each circuit in a program could result in evolutionary instability[27]. The reuse of DNA sequences in multiple locations of a design can lead to homologous recombination. To expand the number of available insulators, nine additional ribozymes were identified in the NCBI sequence database and screened for their insulating capability (**Table 1**). Each ribozyme-based insulator was tested for its ability to produce the same ratio of CI-GFP to GFP, whether under the control of pTAC or pBAD (**Table 1**). As with RiboJ, we added an additional hairpin downstream of each ribozyme. Out of this library, five additional insulators were identified: LtsvJ[28], SccJ[29], SarJ[30], PlmJ[31] and VtmoJ[32] (**Supplementary Fig. 9**). These insulators are statistically equally effective. The remaining four insulators were not selected because residual interference remained when the input promoter was changed. The LtsvJ insulator was shown to have the same effect on the McbR-based NOT gate as RiboJ (**Fig. 3e**).

Efforts are underway to automate the assembly of complex, multicircuit programs[15,33]. This requires mathematical models that can quickly predict the combined behavior of circuits when they are connected in series. Cellular regulation is notoriously noisy, and the signal that links the circuits (promoter activity) will fluctuate from cell to cell. When characterizing circuit transfer functions with flow cytometry, it is possible to measure the precise statistics of these cell-to-cell fluctuations within a population. These statistics are now recognized to contain valuable information about gene regulation phenomena[34]. Ideally, one would fully utilize this information to constrain models and predict how signals propagate from one circuit to the next.

The Gillespie stochastic simulation algorithm[35] (SSA) has been broadly applied to chemical and biological systems[36]. This approach requires a large number of individual simulations to get sufficiently precise statistics. It also requires many kinetic rate parameters and is not suited for stiff problems. The stochastic analysis that we employ here is a modified finite state projection (FSP) algorithm, which directly solves the chemical master equation to obtain the probability density vector[37]. This probability density can be directly compared to the population distribution obtained by cytometry[38]. For small circuits, the FSP algorithm is faster and more accurate than the SSA. However, as the number of genes and transcription factors increases, the FSP analysis time grows exponentially. As an example, suppose that a simple two-species circuit can be analyzed in one minute and that each additional species increases computational time by a factor of ten. At that scaling, a four-species analysis would take over an hour to complete, and our largest genetic program to date[39], which consists of 12 regulatory proteins, would require over 10^8 hours to solve!

Two changes were made to accelerate the FSP algorithm. First, we use a coarse grid approach similar to finite element analysis, which allows us to approximate the FSP solution with relatively small-system ordinary differential equations. This approximation allows us to solve three-species reactions much more efficiently (in seconds rather than tens of minutes). Second, we use a spectral method[40] that allows us to break long cascades of chemical reactions into several smaller cliques of three species apiece. So long as there is no feedback from a downstream clique into an upstream one, this modification allows us to break the exponential dependence on the number of species in favor of a linear dependence on the number of cliques. With both approximations, a linear chain of 12 regulatory proteins can now be solved as 10 overlapping 3-species cliques, and can take less than a minute to complete. In principle, the approach can scale to analyze programs consisting of hundreds of individual circuits. Details of the modified FSP algorithm are provided in **Supplementary Figures 10–14**.

Here, we applied the modified version of the FSP algorithm to (i) parameterize the circuit and inducible systems and (ii) predict their collective behavior when connected to form a program (**Supplementary Figs. 11–14**)[37,38]. This computational approach was tested on a connection of the pBAD inducible system to the NOT gate. The NOT gate was characterized in the model with and without RiboJ using the pTAC inducible system to control the input promoter activity. The FSP algorithm was combined with a genetic algorithm to find a set of parameters that fit the data (**Supplementary Figs. 11–13**). These parameters were then fixed and used to predict the combined behavior of the two circuits when they are connected together (**Supplementary Table 3**). In the absence of RiboJ, the model was unable to predict the behavior of the combined circuit (**Fig. 3g**). However, when RiboJ was included, the prediction closely matches experiments (**Fig. 3h**).

We have identified a problem that underpins difficulties in building gene circuits. The challenge is that the transfer function of the circuit should remain consistent when measured in isolation and in the context of a larger program. The problem occurs because of the interference that arises at the junction of parts. Genetic systems have been published that contain >100 parts[39,41,42]. As the number of parts in a system grows, the potential complications emerging from part-junction interference will grow proportionally. This extends beyond the class of interference described here. The fraction of DNA sequences that encodes some functional information is large when considering all biochemical interactions (e.g., protein-binding operators, methylation sites, RNase sites, promoters), including sequence degeneracy. Even if the probability of a functional sequence arising is small, it may approach unity when considering a large sequence with many part-junctions. Several approaches can mitigate this effect. First, computational methods based on bioinformatics and biophysics can actively scan a design to identify and eliminate putative interference[43–45]. Second, the inclusion of insulators such as those described in this work enables classes of interference to be reduced without relying on predictions. Third, strict measurement reporting and carefully engineered context independence could improve the predictability of part combinations[18,46].

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

C.A.V. is supported by Life Technologies, Defense Advanced Research Projects Agency Chronicle of Lineage Indicative of Origins (DARPA CLIO, N66001-12-C-4018), Office of Naval Research (N00014-10-1-0245), National Science Foundation (NSF) (CCF-0943385), National Institutes of Health (AI067699) and the NSF Synthetic Biology Engineering Research Center (SynBERC, SA5284-11210).

AUTHOR CONTRIBUTIONS

C.A.V. conceived of and supervised the project. C.L. designed and performed the experiments. B.S. performed experiments with the McbR repressor. Y.-J.C. and C.L. performed the q-PCR experiments. B.M. and C.L. analyzed data. C.L., B.M., Y.-J.C. and C.A.V. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/nbt.2401.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Kim, P.M. & Tidor, B. Limitations of quantitative gene regulation models: a case study. *Genome Res.* **13**, 2391–2395 (2003).
2. Del Vecchio, D., Ninfa, A.J. & Sontag, E.D. Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.* **4**, 161 (2008).
3. Grünberg, R. & Serrano, L. Strategies for protein synthetic biology. *Nucleic Acids Res.* **38**, 2663–2675 (2010).
4. Tan, C., Marguet, P. & You, L. Emergent bistability by a growth-modulating positive feedback circuit. *Nat. Chem. Biol.* **5**, 842–848 (2009).
5. Ellis, T., Wang, X. & Collins, J.J. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat. Biotechnol.* **27**, 465–471 (2009).
6. Anderson, J.C., Voigt, C.A. & Arkin, A.P. Environmental signal integration by a modular AND gate. *Mol. Syst. Biol.* **3**, 133 (2007).
7. Wang, B., Kitney, R.I., Joly, N. & Buck, M. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nat. Commun.* **2**, 508 (2011).
8. Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl. Acad. Sci. USA* **102**, 3581–3586 (2005).
9. Pedraza, J.M. & van Oudenaarden, A. Noise propagation in gene networks. *Science* **307**, 1965–1969 (2005).
10. Endy, D. Foundations for engineering biology. *Nature* **438**, 449–453 (2005).
11. Voigt, C.A. Genetic parts to program bacteria. *Curr. Opin. Biotechnol.* **17**, 548–557 (2006).
12. Bio FAB Group *et al.* Engineering life: building a fab for biology. *Sci. Am.* **294**, 44–51 (2006).
13. Shetty, R.P., Endy, D. & Knight, T.F. Jr Engineering BioBrick vectors from BioBrick parts. *J. Biol. Eng.* **2**, 5 (2008).
14. Anderson, J.C. *et al.* BglBricks: A flexible standard for biological part assembly. *J. Biol. Eng.* **4**, 1 (2010).

15. Clancy, K. & Voigt, C.A. Programming cells: towards an automated 'Genetic Compiler'. *Curr. Opin. Biotechnol.* **21**, 572–581 (2010).
16. Arkin, A. Setting the standard in synthetic biology. *Nat. Biotechnol.* **26**, 771–774 (2008).
17. Rhodius, V.A., Mutalik, V.K. & Gross, C.A. Predicting the strength of UP-elements and full-length *E. coli* σE promoters. *Nucleic Acids Res.* 10.1093/nar/gkr1190 (2012).
18. Davis, J.H., Rubin, A.J. & Sauer, R.T. Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res.* **39**, 1131–1141 (2011).
19. Carrier, T.A. & Keasling, J.D. Engineering mRNA stability in *E. coli* by the addition of synthetic hairpins using a 5′ cassette system. *Biotechnol. Bioeng.* **55**, 577–580 (1997).
20. Qi, L., Haurwitz, R.E., Shao, W., Doudna, J.A. & Arkin, A.P. RNA processing enables predictable programming of gene expression. *Nat. Biotechnol.* advance online publication, doi:10.1038/nbt.2355 (16 September 2012).
21. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
22. Elowitz, M.B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
23. Yokobayashi, Y., Weiss, R. & Arnold, F.H. Directed evolution of a genetic circuit. *Proc. Natl. Acad. Sci. USA* **99**, 16587–16591 (2002).
24. Lou, C. *et al.* Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch. *Mol. Syst. Biol.* **6**, 350 (2010).
25. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1–I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210 (1997).
26. Buzayan, J.M., Gerlach, W.L. & Bruening, G. Satellite tobacco ringspot virus RNA: A subset of the RNA sequence is sufficient for autolytic processing. *Proc. Natl. Acad. Sci. USA* **83**, 8859–8862 (1986).
27. Galdzicki, M., Rodriguez, C., Chandran, D., Sauro, H.M. & Gennari, J.H. Standard Biological Parts Knowledgebase. *PLoS ONE* **6**, e17005 (2011).
28. Forster, A.C. & Symons, R.H. Self-cleavage of virusoid RNA is performed by the proposed 55-nucleotide active site. *Cell* **50**, 9–16 (1987).
29. Di Serio, F., Daròs, J.A., Ragozzino, A. & Flores, R. A 451-nucleotide circular RNA from cherry with hammerhead ribozymes in its strands of both polarities. *J. Virol.* **71**, 6603–6610 (1997).
30. Kaper, J.M., Tousignant, M.E. & Steger, G. Nucleotide sequence predicts circularity and self-cleavage of 300-ribonucleotide satellite of arabis mosaic virus. *Biochem. Biophys. Res. Commun.* **154**, 318–325 (1988).
31. Hernández, C. & Flores, R. Plus and minus RNAs of peach latent mosaic viroid self-cleave in vitro via hammerhead structures. *Proc. Natl. Acad. Sci. USA* **89**, 3711–3715 (1992).
32. Roossinck, M.J., Sleat, D. & Palukaitis, P. Satellite RNAs of plant viruses: structures and biological effects. *Microbiol. Rev.* **56**, 265–279 (1992).
33. Beal, J., Lu, T. & Weiss, R. Automatic compilation from high-level biologically-oriented programming language to genetic regulatory networks. *PLoS ONE* **6**, e22490 (2011).
34. Munsky, B., Neuert, G. & van Oudenaarden, A. Using gene expression noise to understand gene regulation. *Science* **336**, 183–187 (2012).
35. Gillespie, D.T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977).
36. Arkin, A., Ross, J. & McAdams, H.H. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics* **149**, 1633–1648 (1998).
37. Munsky, B. & Khammash, M. The finite state projection algorithm for the solution of the chemical master equation. *J. Chem. Phys.* **124**, 044104 (2006).
38. Munsky, B., Trinh, B. & Khammash, M. Listening to the noise: random fluctuations reveal gene network parameters. *Mol. Syst. Biol.* **5**, 318 (2009).
39. Moon, T.S., Lou, C., Tamsir, A., Stanton B.C. & Voigt, C.A. Genetic programs constructed from layered logic gates in single cells. *Nature* advance online publication, doi:10.1038/nature11516 (7 October 2012).
40. Walczak, A.M., Mugler, A. & Wiggins, C.H. A stochastic spectral analysis of transcriptional regulatory cascades. *Proc. Natl. Acad. Sci. USA* **106**, 6529–6534 (2009).
41. Ro, D.-K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
42. Temme, K., Zhao, D. & Voigt, C.A. Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. *Proc. Natl. Acad. Sci. USA* **109**, 7085–7090 (2012).
43. Rhodius, V.A. & Mutalik, V.K. Predicting strength and function for promoters of the *Escherichia coli* alternative sigma factor, sigmaE. *Proc. Natl. Acad. Sci. USA* **107**, 2854–2859 (2010).
44. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
45. Delcher, A.L., Harmon, D., Kasif, S., White, O. & Salzberg, S.L. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27**, 4636–4641 (1999).
46. Kelly, J.R. *et al.* Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.***3**, 4 (2009).
47. Khvorova, A., Lescoute, A., Westhof, E. & Jayasena, S.D. Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nat. Struct. Biol.* **10**, 708–712 (2003).
48. Nelson, J.A., Shepotinovskaya, I. & Uhlenbeck, O.C. Hammerheads derived from sTRSV show enhanced cleavage and ligation rate constants. *Biochemistry* **44**, 14577–14585 (2005).

ONLINE METHODS

Media and strains. Luria-Bertani (LB) media (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) (Fisher Scientific) is used for plasmid construction and strain maintenance. All of the induction studies were done using M9 minimal media supplemented with salts (6.8 g/L Na_2PO_4, 3g/L KH_2PO_4, 0.5 g/L NaCl, 1 g/L NH_4Cl, Sigma), 2 mM $MgSO_4$ (Fischer Scientific), 100 μM $CaCl_2$ (Sigma), 0.4% glucose (Sigma) and 0.2% Casamino acids (Acros). Both the LB and M9 media were supplemented with 20 μg/ml chloramphenicol (Acros) and appropriate concentration of isopropyl–D-1-thiogalactopyranoside (IPTG) (USB Corporation), sodium salicylate (Sigma) or L-arabinose (Sigma). All studies were done using *E. coli* DH10B. The sequence of McbR was codon optimized and obtained from Geneart. All the maps of plasmids are shown in **Supplementary Figures 15–21**.

Cell growth. The cells were grown according to the following protocol before assaying their fluorescence. First, cells were inoculated from single colonies on LB agar plates and grown overnight in 4 ml LB in Falcon tubes at 37 °C with shaking (250 r.p.m.). Then, the overnight cultures were diluted 200-fold into prewarmed M9 media in a 96-well plate. The plates were then incubated at 37 °C in a Safire plate spectrophotometer (Tecan) with high orbital shaking, in which OD_{600} measurements were recorded every 5 min. Once the diluted cultures reached an OD_{600} of 0.12–0.14 (~3 h), the cultures were diluted 700-fold and loaded into prewarmed M9 media containing inducer in 96-well plates. These new plates were incubated at 37 °C in a Digital Thermostatic Shaker (Elmi) with 1,000 r.p.m. shaking for 6 h to maintain exponential growth. Finally, after the 6-h induction, a sample of each culture was transferred to a new plate containing 150 μl PBS and 2 mg/ml kanamycin to stop protein expression.

Fluorescence measurement. The fluorescence distribution of each sample was measured using a LSRII flow cytometer (BD Biosciences) with appropriate voltage settings (FSC:516, SSC:286, FITC:680, B:650). Each distribution contains at least 50,000 events and is gated by the forward and side scattering using FlowJo (v7.6). The geometric mean of each sample is calculated. The autofluorescence of *E. coli* DH10B cells containing pSB4C5 plasmid alone was measured and subtracted from the mean value to generate the fluorescence values reported.

5′-terminal mRNA analysis. To validate the self-cleaving activity of the RiboJ spacer, we sequenced the 5′-terminal end of the mRNA encoding *cI* under the control of pTAC. The mRNA was extracted and reverse-transcribed by rapid amplification of cRNA end (5′ RACE) protocol (Invitrogen 5′ RACE System for rapid amplification of cRNA ends, Version 2.0). Total RNA of each sample was isolated and purified using the RiboPure bacteria kit (Ambion). 14 μg total RNA of each sample was sequentially treated with T4 polynucleotide kinase (T4 PNK) (Invitrogen) and tobacco acid pyrophosphatase (TAP) (Epicentre) before ligation with the unique RNA adaptor (5′-GAGGACUCGAGCUCAAGC-3′). Treatment with the T4 PNK is important, because the 5′-hydroxyl terminus of the self-cleaved mRNA cannot ligate with the RNA adaptor until a phosphate is added by the T4 PNK. TAP facilitates converting the 5′-triphosphate to single-phosphate of the uncleaved mRNA, thus promoting the ligation between mRNA and the adaptor. After ligation, reverse transcription and subsequent amplification were performed by SuperScript III (Life Technologies) with a gene-specific primer (5′-ATCCCCATCTTGTCTGCGACAG-3). This PCR product was used as a template for a nested PCR with DNA adaptor (the same sequence with the RNA adaptor except U replaced by T) and another gene-specific primer (5′- TCCTGGGATAAGCCAAGTTC-3′), whose PCR product was purified and subsequently sequenced.

doi:10.1038/nbt.2401