



Published in final edited form as:

*Mol Biosyst.* 2008 July ; 4(7): 754–761. doi:10.1039/b801245h.

## Engineering Stochasticity in Gene Expression

Jeffrey J. Tabor<sup>1,4</sup>, Travis S. Bayer<sup>2</sup>, Zachary B. Simpson<sup>1</sup>, Matthew Levy<sup>1</sup>, and Andrew D. Ellington<sup>1,3,\*</sup>

<sup>1</sup>Center for Systems and Synthetic Biology and Institute for Cell and Molecular Biology, University of Texas at Austin

<sup>2</sup>Division of Biology, California Institute of Technology, Pasadena, California 91125

<sup>3</sup>Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, TX 78712.

### Summary

Stochastic fluctuations (noise) in gene expression can cause members of otherwise genetically identical populations to display drastically different phenotypes. An understanding of the sources of noise and the strategies cells employ to function reliably despite noise is proving to be increasingly important in describing the behavior of natural organisms and will be essential for the engineering of synthetic biological systems. Here we describe the design of synthetic constructs, termed ribosome competing RNAs (rcRNAs), as a means to rationally perturb noise in cellular gene expression. We find that noise in gene expression increases in a manner proportional to the ability of an rcRNA to compete for the cellular ribosome pool. We then demonstrate that operons significantly buffer noise between coexpressed genes in a natural cellular background and can even reduce the level of rcRNA enhanced noise. These results demonstrate that engineering exogenous genetic elements can significantly affect the natural noise profile of a living cell and, importantly, that operons are a facile genetic strategy for buffering against noise.

### Keywords

stochasticity; mRNA; ribosome; ribosome binding site; operon

### Introduction

Gene expression is inherently stochastic, due primarily to the small numbers of molecules involved in the process<sup>1</sup>. Noise intrinsic to gene expression is thought to be dictated by fluctuations in mRNA levels, which may arise from fluctuations in promoter states<sup>2-6</sup> followed by bursts of transcription<sup>7</sup> or the random births and deaths of mRNAs themselves<sup>8-11</sup>. Messenger RNA fluctuations are then amplified by random mRNA:ribosome interactions and concomitant bursts in protein production<sup>8; 12-16</sup>. As a consequence, the concatenation of multiple genes into a single expression unit (operon formation) has been predicted to decrease uncorrelated fluctuations in the levels of protein products<sup>11</sup>. Noise has also been shown to result from fluctuations in factors extrinsic to the genes themselves. These factors are thought to include pathway specific modulators<sup>2; 3; 17-19</sup> and global factors of gene expression such as the levels of nucleic acid polymerases or ribosomes<sup>18; 20-23</sup>.

<sup>4</sup>Current address: Department of Pharmaceutical Chemistry, University of California – San Francisco, San Francisco, California 94158

\*To whom correspondence should be addressed: andy.ellington@mail.utexas.edu

Synthetic biology, or the (re)construction of gene networks with defined performance characteristics, has proven a useful paradigm for studying the principles which govern cellular function<sup>24-26</sup>. Several synthetic studies have demonstrated that noise can drive identical gene regulatory networks to encode significant variation across cell populations<sup>6; 27-32</sup>. It is now apparent that the study and construction of genetic regulatory circuits will require strategies for understanding and controlling intracellular noise. To this end, we constructed ribosome-competing RNA constructs (rcRNAs) as a synthetic biological tool to study the effects of rationally manipulating noise in living cells, and to begin to develop methods for engineering noise buffering.

## Results and Discussion

### Ribosome competition increases noise in gene expression

To examine how the expression of exogenous genes might affect noise, we engineered a series of small mRNAs whose predicted function was to compete for ribosomes in *E. coli* (Figure 1a). These RNAs, which we term ribosome competing RNAs (rcRNAs), contained hairpin stems of varying stabilities that occluded a ribosome binding site (RBS) to different degrees (Figure 1b). If rcRNAs were expressed to high levels in cells, they would presumably compete for cellular ribosome pools based on the extent of exposure of the RBS. Natural and synthetic strategies for RBS occlusion have previously been shown to be efficient methods for blocking ribosome association and subsequent translation as well<sup>33-36</sup>.

To examine whether rcRNAs could directly compete for translation in a manner dependent on the availability of the RBS, we performed two *in vitro* assays. First, the predicted secondary structures of rcRNAs were verified by limited hydrolysis of single-stranded regions (in-line probing)<sup>37</sup>. As predicted, an rcRNA variant engineered to have a perfectly base-paired RBS/anti-RBS stem (rc1), showed the greatest double-strandedness in both regions (Figure 1c). Thus, rc1 was predicted to have minimal capacity for translation competition amongst the rcRNA series. As bulges were designed into the anti-RBS stem in different rcRNA variants, both the RBS and anti-RBS regions became less structured, rationally allowing for increased ribosome binding (Figure 1c).

Second, to directly assay the functionality of the rcRNAs, we performed translation competition assays in a reconstituted *E. coli* lysate. A fixed amount of reporter gene (*gfp*) DNA template was added to a coupled *in vitro* transcription/translation reaction and increasing amounts of DNA templates for either a highly structured (rc1) or a largely unstructured (rc6) rcRNA were added to the reaction. Interestingly, addition of either rcRNA template at low levels resulted in an increase in GFP protein abundance. This effect is likely due to stabilization of the *gfp* mRNA transcript through RNase competition. The rc1 template continued to increase GFP abundance when added at a 5:1 ratio relative to the *gfp* template, but resulted in no further increase in GFP production when added at a 10:1 ratio. In sharp contrast, the addition of increasing amounts of the relatively unstructured rc6 template resulted in a strong inhibition of translational capacity (Figure 1d). These results demonstrated that rcRNAs are capable of reducing translation rate in a manner proportional to the availability of the RBS sequence within the engineered helix.

To measure the effect of ribosome competition on noise, we then followed the expression of a *gfp* gene in individual *E. coli* cells using flow cytometry. Individual rcRNA constructs that could compete for ribosomes to different extents were introduced into *E. coli*, and rcRNA expression was driven to high levels by T7 RNA polymerase (**Experimental**). Noise was quantified as the standard deviation in GFP abundance divided by the mean over the cell populations, a metric also known as the coefficient of variation (CV). As predicted, the availability of the rcRNA RBS showed a strong positive correlation with noise, and a strong

negative correlation with GFP abundance (Figure 2a, Table 1). It is worthy to note that though the availability of the RBS from the in-line probing experiments had a largely predictable effect on *in vivo* expression data, the correlation between the two experiments was not perfect. For example, though rc1 appears to have a less available RBS than rc2 and rc3 (Figure 1c), rc1 expression results in lower GFP abundance and a slightly higher noise profile than does the expression of rc2 and rc3 (Table 1). The RBS of rc4, however, was shown to be the most available in the in-line assays, and rc4 indeed resulted in the lowest GFP abundance and the greatest amount of noise *in vivo*. As was the case with the *in vitro* translation reactions above, expression of the highly structured rcRNAs (rc1-3) generally led to increased GFP abundance, while expression of the mostly unstructured rcRNAs (rc4, 5 and 6) decreased GFP.

The *in vivo* rcRNA expression data showed that noise and GFP abundance obeyed an inverse power law relationship (Figure 2b), a scaling previously observed in studies in which the rate of transcription was modulated<sup>17; 18</sup>. Given that mRNAs are known to compete for cellular ribosome pools<sup>38</sup> and that highly expressed mRNAs are capable of sequestering nearly all cellular ribosomes<sup>39</sup>, one interpretation of these results is that rcRNA-induced noise results from probabilistic mRNA:ribosome interactions which become infrequent as ribosomes become rare. It is very important to note, however, that there are many downstream effects resulting from the initial competition for ribosomes which likely contribute to the noise profiles observed in our experiments. For example, competition for ribosome pools would reduce the expression level of all cellular proteins, including RNA polymerases. A reduction in RNA polymerase abundance should itself increase total noise in gene expression, due to a reduction in transcription rate. Though it has previously been shown that noise scales inversely with net expression level<sup>8; 40</sup>, the results of these experiments demonstrate that the same scaling can arise from ribosome competition as a result of the expression of exogenous genes.

Our results are also consistent with the hypothesis that inefficiently translated rcRNAs function as RNase competitors, stabilizing the *gfp* mRNA, increasing GFP abundance and in turn decreasing noise in GFP levels. The intracellular stability of mRNAs is thought to be governed in part by a competition between ribosome binding and RNaseE-induced degradation<sup>41</sup>. We therefore expected that when ribosome competition increases, mRNA levels should decline, increasing the significance of mRNA fluctuations and therefore noise in protein levels. To examine this effect, we quantitated *gfp* mRNA levels in strains expressing different rcRNAs. Consistent with the mRNA stability model, the magnitude of induction of *gfp* mRNA (**Experimental**) was approximately 12-fold lower in *E. coli* expressing an rcRNA that could effectively compete for ribosomes as compared with a control rcRNA with an occluded RBS (Figure 2c). This is likely a combined effect of ribosome competition and lower mRNA half-life. These two factors can feedback on one another, as the observed decline in mRNA levels should further reduce the frequency of mRNA:ribosome interactions, which could in turn reduce transcriptional or metabolic capability and thus mRNA production. The possibility that such a noise amplification cascade exists would argue strongly that there should be mechanisms for the reduction or regulation of noise in cells.

### Operons buffer gene expression noise

To determine if cells have natural ‘noise abatement’ mechanisms, we used our synthetic noise generators to investigate a standing mathematical model which predicts that the concatenation of multiple genes onto a single mRNA (operon formation) buffers relative fluctuations in the levels of coexpressed proteins<sup>11</sup>. This buffering is predicted to occur because mRNA fluctuations contribute strongly to the noise profile of a given gene (as we have also shown) and operons virtually eliminate relative mRNA fluctuations between genes. To examine the model, we constructed a synthetic operon based on the dual fluorescent protein reporter system<sup>17</sup>. In the engineered operon, two fluorescent genes (*cfp* and *yfp*) were arrayed in series, each

preceded by a strong RBS (Figure 3a). The mRNA encoding these genes was transcribed from an inducible promoter ( $P_{tet}$ ) which allowed for examination of the noise profile over a large range of expression levels. The translation of the two genes was made independent by the incorporation of two consecutive strong stop codons at the end of each open reading frame. As a control, an analogous monocistronic version was built in which the two fluorescent reporter genes were transcribed independently from identical promoters and translated from identical ribosome binding sites (Figure 3a).

*E. coli* populations carrying the monocistronic and bicistronic constructs were grown with increasing concentrations of the transcriptional inducer anhydrotetracycline (aTc) and populations were analyzed by multi-channel flow cytometry (Experimental). In agreement with the model <sup>11</sup>, the bicistronic expression platform resulted in significantly less noise than the monocistronic version over a large range of low protein expression levels (Figure 3b). To directly compare the noise profiles of two platforms, the mono and bicistronic constructs were induced to the same level of protein expression (0.012 relative fluorescence units) and noise was quantified. At this low expression level, the operon showed ~70% less noise than the monocistronic construct (Figure 3b).

The mathematical model <sup>11</sup> predicts that the noise buffering properties of the operon will be greatest at low protein expression levels, and that noise in mono and bicistronic systems will converge as protein abundance increases. We investigated this prediction by varying protein expression from each construct from very low to very high levels. In agreement with the model, the noise output of the two constructs converged as protein abundance increased, and that at the highest expression levels the operon had no noise buffering effect (Figure 3b). The contribution of mRNA fluctuations to noise is expected to be greatest at low expression levels, and this was also the regime under which the operon served to most efficiently buffer noise. At higher expression levels, mRNA fluctuations become less significant, diminishing the noise buffering effect of bicistronic encoding. Taken together, the noise trends of the mono and bicistronic constructs are in strong agreement with the predictions of the model over the protein abundances in our experiments.

We then introduced rcRNAs with the mono and bicistronic reporter constructs in order to assay the tolerance of the different expression systems to genetically-encoded noise. To better compare the mono and bicistronic systems, we induced both to the same high expression level, where their respective noise levels were low. This was achieved by inducing cells carrying the bicistronic plasmid to the maximal expression level (Figure 3b) and varying the aTc concentration added to cells carrying the monocistronic plasmid to achieve the same level of fluorescence. In this high expression regime, noise in the monocistronic construct is naturally lower than the bicistronic version (Figure 3b, green arrow). Even so, the introduction of a noisy rcRNA, rc6, resulted in 20% more noise in the monocistronic construct than in the bicistronic version (Figure 3c). That is, the operon was significantly more tolerant to rcRNA induced noise than the monocistronic construct. These results bolster the hypothesis that operons function as genetic noise insulators, and demonstrate their efficacy in buffering artificially enhanced noise resulting from ribosome competition and mRNA degradation.

## Conclusions

The origins of stochasticity in gene expression vary widely between different organisms and even between genes within individual organisms <sup>2; 8; 9</sup>. Random fluctuations in promoter states, transcription events and mRNA deaths are amplified by translation and lead to noise in protein levels across genetically identical cell populations. This noise in turn, can lead to dramatically different phenotypes between individual cells <sup>6; 30-32; 42-46</sup>.

Here we show that noise in gene expression can be rationally engineered and that noise is strongly affected by the introduction of exogenous mRNAs which compete for ribosomes. Previous studies have shown that bulk cellular mRNAs compete for access to the ribosome<sup>38</sup>, and that highly expressed mRNAs are capable of occupying nearly all cellular ribosomes<sup>39</sup>. Our artificial ribosome-depleted states may mimic naturally occurring cellular states that arise when certain genes, such as those involved in stress response, are expressed to high levels or translated preferentially<sup>47-54</sup>. The relationship between competition for translation and noise may prove to be universal for any of a number of factors required for gene expression (e.g., transcription factors), a hypothesis that is directly amenable to experimental evaluation. Indeed, competition for ribosomes may affect the translation and availability of other gene expression machinery and thereby initiate feedback cascades that further exacerbate variance in gene expression.

We also demonstrate that operons strongly buffer against noise in the coexpression of multiple genes, reducing noise up to ~70% (nearly 4-fold) at low expression levels. While the cotranscription of multiple genes on a single mRNA undoubtedly reduces against noise generated by random promoter transitions and mRNA birth and death events<sup>3-5; 10; 40</sup>, we demonstrate here that operons also buffer against noise arising from diminished ribosome pools. The noise buffering properties of the operon were particularly strong at low expression levels in our experiments. These low expression levels are likely more representative of natural protein abundances than the higher expression levels, as the genes were expressed from a multicopy plasmid. This implies that noise buffering may be a relevant property of many natural operons. It is possible, then, that operons represent a reliable strategy by which cells can stoichiometrically couple the expression of multiple genes, even in translationally compromised or stressed environments.

The finding that operons have noise buffering properties has fundamental evolutionary implications. Though operons are one of the most ubiquitous forms of gene organization in nature, they are relatively unstable over evolutionary time. Operons frequently decompose into multiple genetic loci which are regulated by the same transcription factors, but are independently transcribed and translated<sup>55</sup>. The most stable operons share an intriguing feature: they encode genes whose products physically interact<sup>56</sup>. When gene products interact (e.g. in multi-protein complexes), stoichiometric coupling becomes critical, as the over- or underproduction of any single product will squander cellular resources and have deleterious fitness effects, and polycistronic encoding may provide a strong selective advantage against these effects.

Our results are also relevant to many approaches in synthetic biology. The engineering of synthetic biological systems with complex behaviors is proving to be challenging, as noise in the expression of certain gene products thwarts efforts to forward engineer deterministic phenotypes<sup>29</sup>. Moreover, as the size (in DNA base pairs) of the synthetic biological constructs introduced into living cells is increased, competition for cellular resources, including ribosomes, will become greater and noise will concomitantly increase. One consequence of our experiments is to suggest that operons may offer a robust design strategy for those attempting to engineer synthetic pathways and behaviors that require reliable stoichiometric coupling of multiple gene products (such as the efficient production of foreign metabolites)<sup>57</sup>.

Ultimately, these results nicely emphasize the increasing crossover between systems and synthetic biology. It was a synthetic tool (and orthogonal noise generator) that allowed the rational manipulation of the noise inherent in genetic expression. Novel synthetic circuits are frequently at the mercy of the cellular backgrounds in which they are implanted, and it is thus difficult to predict and model their performance. Our synthetic tool can now also be applied to

any synthetic circuit, acting as an perturbant to determine whether and to what extent the cellular machinery is taxed by a synthetic circuit (and vice versa).

## Experimental

### rcRNA plasmid construction

rcRNAs (**Supplementary Table 1**) were embedded between the NcoI and BlnI sites (underlined) in the following 100 nucleotide DNA: 5'-GATGGCAGCTACTAATGCTAGCTAAGATACCATG-31nt\_rcRNA-GCTGAGCAACGAGCTATAGCTACTACTGATAGTCC-3'. 1pmol of the oligonucleotide was amplified by PCR using the flanking primers 44.31F: 5'-ACGAGATCTCAGATCAGACCAGGATGGCAGCTACTAATGCTAGC-3' and 44.31R: 5'-CGGTGAACTCGTCTATGGATCTGGACTATCAGTAGTAGCTATAG-3', the final 20nt of which were complimentary to 100nt oligo. The resulting double stranded DNA was purified using the Qiaquick PCR purification kit (Qiagen, Valencia CA) and digested with NcoI and BlnI (New England Biolabs, Ipswich MA) according to manufacturer instructions. Appropriate sized DNA fragments were isolated by gel purification using 4% NuSieve GTG agarose (Cambrex, Baltimore MD) in TBE, and recovered using a Wizard SV gel purification column (Promega, Madison WI). 1µg of the rcRNA host vector, pACYCDuet-1 (Novagen, Madison WI) was digested with NcoI and BlnI, and gel purified using 2% Seaplaque agarose (Cambrex) in TAE. 60fmol of digested rcRNA DNA was ligated into 20fmol digested pACYCDuet-1 with T4 DNA ligase (Invitrogen, Carlsbad CA) and transformed into Top10 cells (Invitrogen) according to manufacturer instructions. The insert regions were sequence verified and the plasmids were transformed into BL21DE3 (Novagen) along with the appropriate reporter plasmid for all assays. All DNA oligonucleotides and primers were purchased from Integrated DNA Technologies (Coralville, IA), resuspended in H<sub>2</sub>O and used directly. The pACYCDuet-1 host plasmid carries a p15A (~10-12 copy) origin of replication, and a chloramphenicol resistance marker. rcRNAs were expressed by the IPTG inducible T7 RNA polymerase (carried in the genome of BL21DE3) through the IPTG inducible T7 promoter upstream of the first MCS of pACYCDuet-1. All pACYC-hosted rcRNA plasmids were maintained in BL21DE3 using 34ug/mL chloramphenicol.

### Reporter plasmids

The expression plasmid pASKIBA3.GFPm2 was used for all GFP flow cytometry assays. The gene GFPm2 was expressed from the multiple cloning site of the host plasmid pASKIBA3 (IBA, Göttingen Germany) under the control of a Tetracycline Repressor (TetR) controlled promoter. pASKIBA3 carries a marker for ampicillin resistance, a pBR322 (~40 copy number) based origin of replication, and the TetR gene. pASKIBA3.GFPm2 was maintained in BL21DE3 using 50ug/mL ampicillin.

The CFP/YFP expression plasmids were constructed using MIT's registry of standard biological parts (<http://parts.mit.edu>) and DNA isolation, purification and ligation methods as described above. The host plasmid pSB4A3 (MIT's registry of standard biological parts) was used for construction of the monocistronic (J13065) and the bicistronic constructs (J13004). pSB4A3 bears a pSC101 origin of replication (~ 5 copy) and an ampicillin resistance marker. 25ug/mL ampicillin was used to maintain all pSB4A3-based plasmids. For noise assays, J13065 and J13004 were cloned into pASKIBA3 and expressed under control of TetR.

To build the monocistronic construct (J13065), a DNA sequence bearing the P<sub>tet</sub> promoter (R0040), strong RBS (B0034), ECFP coding sequence (E0020) and strong transcriptional terminator (B0015) was cloned into pSB4A3 using the restriction enzymes EcoRI and PstI. A second strong transcriptional terminator (B0015) was cloned downstream of the first. Finally

an EYFP expression cassette carrying the  $P_{tet}$  promoter, strong RBS (B0034), EYFP gene (E0030) and transcription terminator (B0015) were cloned downstream of the tandem transcription terminators using standard biobrick assembly methods (<http://parts.mit.edu>). The ECFP/EYFP monocistronic expression construct (J13065) was cloned into the pASKIBA3 plasmid using the restriction sites PflMI and PstI. The primers jt\_pflmI\_bb\_prefix\_F: 5'-CGTAGACTAGACCACCATCGAATGGGAATTCGCGGCCGCTTCTAG-3' and BBa\_G00101: 5'-ATTACCGCCTTTGAGTGAGC-3' were used to add an upstream PflMI site to the beginning of J13065 while maintaining the downstream PstI site from pSB4A3. The amplified DNA was digested and cloned into pASKIBA3 using methods described above.

The bicistronic CFP/YFP expression construct, J13004 carries the same promoter, RBS, coding sequences and transcriptional terminator as J13065 (Figure 3a). It was constructed by cloning a DNA segment containing R0040, B0034 and E0020 upstream of a segment containing B0034, E0030 and B0015 as above. J13004 was then moved to pASKIBA3 as before.

### GFP flow cytometry assays

*E. coli* cultures were grown overnight at 37°C with shaking at 250 rpm from frozen stocks in LB plus appropriate antibiotics. Three parallel cultures were diluted 1:100 in fresh media with 100µM IPTG as appropriate, grown to early log phase ( $OD_{600} \sim 0.1$ ) and treated with 200ng/mL anhydrotetracycline (aTc) to induce GFP expression. 4 hours later, 1-10µL of cells were diluted into 0.5-1mL phosphate buffered saline (PBS) and scanned for GFP fluorescence. BL21DE3 carrying no plasmid were measured for autofluorescence and subtracted from all data sets. All GFP expression data were collected on a Becton Dickinson FACSCalibur flow cytometer with a 488nm argon excitation laser and a 515-545 emission filter. 30,000 events were collected for each data set and the cell populations were then gated away from other objects by polygonal gates around the dominant forward scatter/side scatter profile. Non-fluorescent objects not representative of the cell population were discarded. Data were analyzed using WinMDI version 2.8 (Joseph Trotter, The Scripps Research Institute).

### CFP/YFP flow cytometry

*E. coli* cultures were grown overnight from frozen stocks as above. Cultures were diluted 1:100 in fresh media with 100µM IPTG when appropriate and grown to  $OD_{600} \sim 0.1-0.2$ . The cultures were then induced with the appropriate amount of anhydrotetracycline (from 2ng/mL to 500ng/mL) and grown shaking at 250rpm and 37°C for 3.75h. 50µL of each culture was then diluted into 2mL of PBS and analyzed on a Becton-Dickinson LSRII cytometer with Trigon Violet and Octagon Blue lasers for CFP and YFP respectively. ECFP fluorescence was measured through a 525/50 bandpass filter while EYFP expression was measured through a 530/30 bandpass filter. BL21DE3 cells carrying no fluorescent proteins were measured for autofluorescence and subtracted from all data sets. 30,000 data points were collected at the LOW flow rate and cells expressing CFP or YFP alone were used to calculate filter bleedthrough, which was used to correct all data sets. Data were exported from the associated FACSDiva software as fcs files and converted to raw, ASCII format with FlowJo7.2 (Tree Star Inc., Ashland, OR) for further analysis.

Cell populations were gated away from other objects by a polygonal gate based on the dominant forward scatter/side scatter profile. Objects with fluorescence values at the lower limit of detection of the cytometer were discarded. The first and last 0.2s of the data were then removed to eliminate cytometer flow rate variability<sup>40</sup>.

### CFP/YFP noise calculations

Intrinsic, extrinsic and total noise in CFP and YFP were quantified using the equations of Elowitz et al.<sup>17</sup> as follows.

$$\eta_{\text{int}}^2 = \frac{\langle (c - y)^2 \rangle}{2\langle c \rangle \langle y \rangle}$$

$$\eta_{\text{ext}}^2 = \frac{\langle cy \rangle - \langle c \rangle \langle y \rangle}{\langle c \rangle \langle y \rangle}$$

$$\eta_{\text{tot}}^2 = \frac{\langle c^2 + y^2 \rangle - 2\langle c \rangle \langle y \rangle}{2\langle c \rangle \langle y \rangle}$$

Where  $\eta$  is noise,  $c$  indicates CFP fluorescence,  $y$  indicates YFP fluorescence and  $\langle \rangle$  indicates a population averaged measurement.

### mRNA quantitation assays

Cells were grown overnight as above, subcultured 1:100 into fresh media with 100 $\mu$ M IPTG to induce rcRNA expression when appropriate. The cultures were grown to early-log phase ( $OD_{600} \sim 0.1$ ) at which point transcription of the *gfpm2* mRNA was induced by the addition of 200ng/mL anhydrotetracycline (aTc) when appropriate. Cells were grown for 4 hours, at which point total RNA was recovered and DNase-treated using an RNeasy kit (Qiagen) according to manufacturer's instructions. cDNA was created for the *gfpm2* mRNA, and 16s ribosomal RNA using MMLV reverse transcriptase (Ambion, Austin TX) according to manufacturer instructions. *Gfpm2* mRNA was specifically reverse transcribed using the primer *GFPm2.rt.R*: 5'- TTTTCGTTGGGATCTTTTCGAA-3' and 16s rRNA using *16s.rt.R*: 5'- ACCGCTGGCAACAAAAGATAA-3'. *GFPm2* cDNA was quantitated on an Applied Biosystems 7900HT real-time PCR machine using the forward primer *GFPm2.rt.F*: 5'- GATGGCCCTGTCCTTTTACCA-3', *GFPm2.rt.R*, and the Taqman probe *GFPm2.rt.p*: 5'- 6FAM/ACAACCATTACCTGTCCACACAATCTGCC/BHQ1 where BHQ1 indicates Black Hole Quencher 1 (Integrated DNA technologies). Cycle threshold (C(T)) values were compared to 16s ribosomal RNA C(T) values. 16s cDNA was amplified using the forward primer *16s.rt.F*: 5'- CGTGTTGTGAAATGTTGGGTAA-3', the reverse primer *16s.rt.R*, and the Taqman probe *16s.rt.p*: 5'- TCCCACAACGAGCGCAACC-3'. ROX (Invitrogen) was used as a passive reference dye for all samples. Average  $\Delta$ C(T) values for each sample were calculated from 3 independent cultures grown in parallel. Fold induction in *gfpm2* mRNA levels from uninduced (aTc minus) to induced (aTc plus) samples was calculated as: fold induction =  $\text{mRNA}_{\text{induced}}/\text{mRNA}_{\text{uninduced}}$

### Translation competition assays

Coupled transcription/translation reactions were performed with the Rapid Translation System RTS 100 *E.coli* HY Kit (Roche, Indianapolis IN) according to manufacturer instructions. 75fmol GFPm2 double stranded DNA template bearing the T7 promoter and ribosome binding site sequence from pACYCDuet-1 was added with increasing molar ratios of rcRNA DNA sequence bearing the same promoter and RBS elements. GFPm2 fluorescence was read on a Tecan SAFIRE fluorescence microplate reader using a clear flat bottomed 96-well plate and an excitation wavelength of 488nm and an emission wavelength of 509nm. The autofluorescence of a reaction without GFPm2 DNA template was measured and subtracted from all samples. All reactions were performed and measured in triplicate.



## In-line probing

In-line RNA structure assays were performed on rcRNA transcripts as described in reference 22 unless otherwise noted. rcRNAs were amplified using primers flanking the T7 promoter and terminator from pACYCDuet-1, transcribed in vitro by a T7 RNA polymerase transcription kit (Epicentre, Madison WI), purified on a 12% denaturing polyacrylamide gel, phosphatased with Antarctic Phosphatase (NEB), radiolabeled on the 5'-most nucleotide, and gel purified as before. In-line degradation was performed at 25°C for 41 hours, and equivalent counts of <sup>32</sup>P containing RNA were added to each lane of a 10% denaturing polyacrylamide gel. The gel was dried under vacuum at 75°C for 1 hour and imaged using a Phosphorimager (Amersham, Piscataway NJ).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

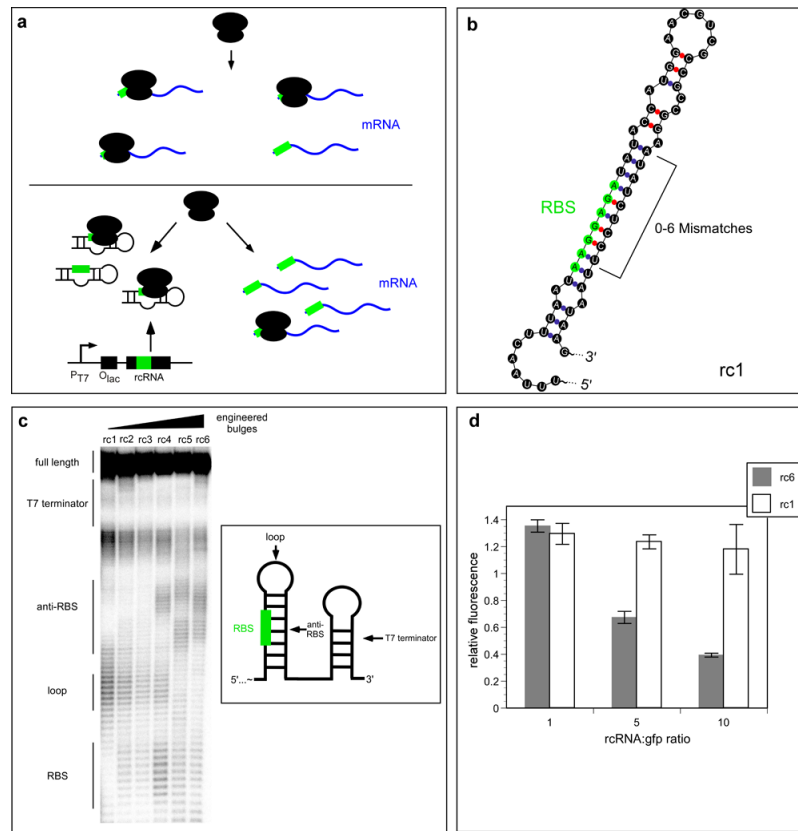
We thank K.E. Griswold for assistance with two-channel flow cytometry, E.A. Davidson for the GFP expression plasmid, J.J. Bull, J.J. Collins, E.M. Marcotte and C.A. Voigt for critical reading of an earlier version of the manuscript. We also thank several anonymous reviewers for insightful commentary and suggestions. This work was supported by NIH grant #9R01GM077040-05, the Welch Foundation and the Texas Higher Education Coordinating Board TDT-003658-0611-2003.

## Bibliographic References

1. McAdams HH, Arkin A. Stochastic mechanisms in gene expression. *Proc Natl Acad Sci U S A* 1997;94:814–9. [PubMed: 9023339]
2. Becskei A, Kaufmann BB, van Oudenaarden A. Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat Genet* 2005;37:937–44. [PubMed: 16086016]
3. Blake WJ, M KA, Cantor CR, Collins JJ. Noise in eukaryotic gene expression. *Nature* 2003;422:633–7. [PubMed: 12687005]
4. Golding I, Paulsson J, Zawilski SM, Cox EC. Real-time kinetics of gene activity in individual bacteria. *Cell* 2005;123:1025–36. [PubMed: 16360033]
5. Raser JM, O'Shea EK. Control of stochasticity in eukaryotic gene expression. *Science* 2004;304:1811–4. [PubMed: 15166317]
6. Blake WJ, Balazsi G, Kohanski MA, Isaacs FJ, Murphy KF, Kuang Y, Cantor CR, Walt DR, Collins JJ. Phenotypic consequences of promoter-mediated transcriptional noise. *Mol Cell* 2006;24:853–65. [PubMed: 17189188]
7. Pedraza JM, Paulsson J. Effects of molecular memory and bursting on fluctuations in gene expression. *Science* 2008;319:339–43. [PubMed: 18202292]
8. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y, Barkai N. Noise in protein expression scales with natural protein abundance. *Nat Genet* 2006;38:636–43. [PubMed: 16715097]
9. Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, Derisi JL, Weissman JS. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature*. 2006
10. Paulsson J. Summing up the noise in gene networks. *Nature* 2004;427:415–8. [PubMed: 14749823]
11. Swain PS. Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *J Mol Biol* 2004;344:965–76. [PubMed: 15544806]
12. Cai L, Friedman N, Xie XS. Stochastic protein expression in individual cells at the single molecule level. *Nature* 2006;440:358–62. [PubMed: 16541077]
13. Yu J, Xiao J, Ren X, Lao K, Xie XS. Probing gene expression in live cells, one protein molecule at a time. *Science* 2006;311:1600–3. [PubMed: 16543458]
14. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A. Regulation of noise in the expression of a single gene. *Nat Genet* 2002;31:69–73. [PubMed: 11967532]

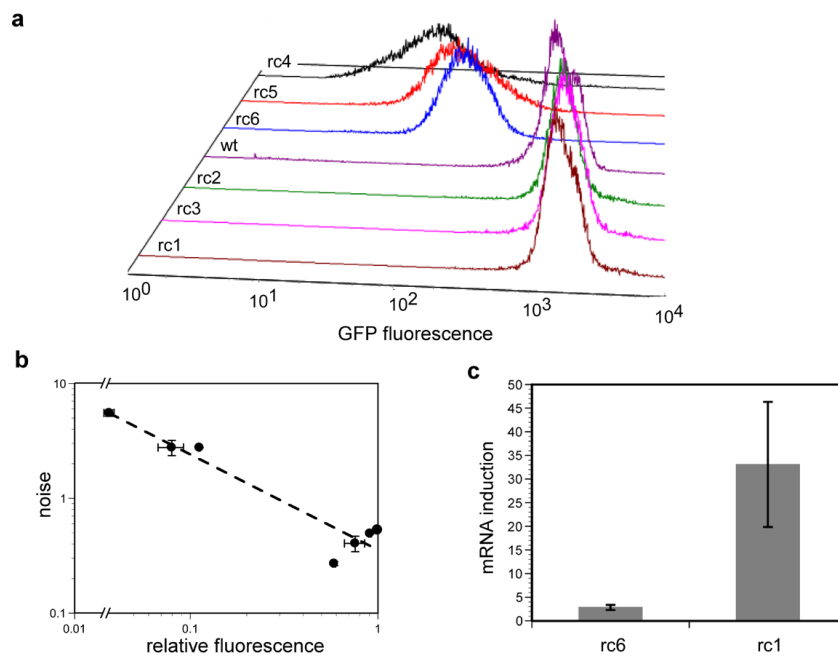
15. Thattai M, van Oudenaarden A. Intrinsic noise in gene regulatory networks. *Proc Natl Acad Sci U S A* 2001;98:8614–9. [PubMed: 11438714]
16. Singh UN. Polyribosomes and unstable messenger RNA: a stochastic model of protein synthesis. *J Theor Biol* 1969;25:444–60. [PubMed: 5387049]
17. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. *Science* 2002;297:1183–6. [PubMed: 12183631]
18. Pedraza JM, van Oudenaarden A. Noise propagation in gene networks. *Science* 2005;307:1965–9. [PubMed: 15790857]
19. Volfson D, Marciniak J, Blake WJ, Ostroff N, Tsimring LS, Hasty J. Origins of extrinsic variability in eukaryotic gene expression. *Nature*. 2005
20. Guido NJ, Lee P, Wang X, Elston TC, Collins JJ. A pathway and genetic factors contributing to elevated gene expression noise in stationary phase. *Biophys J*. 2007
21. Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, Pesce CG, Brent R. Regulated cell-to-cell variation in a cell-fate decision system. *Nature* 2005;437:699–706. [PubMed: 16170311]
22. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. Gene regulation at the single-cell level. *Science* 2005;307:1962–5. [PubMed: 15790856]
23. Chabot JR, Pedraza JM, Luitel P, van Oudenaarden A. Stochastic gene expression out-of-steady-state in the cyanobacterial circadian clock. *Nature* 2007;450:1249–52. [PubMed: 18097413]
24. Drubin DA, Way JC, Silver PA. Designing biological systems. *Genes Dev* 2007;21:242–54. [PubMed: 17289915]
25. Sprinzak D, Elowitz MB. Reconstruction of genetic circuits. *Nature* 2005;438:443–8. [PubMed: 16306982]
26. Yeh BJ, Lim WA. Synthetic biology: lessons from the history of synthetic organic chemistry. *Nat Chem Biol* 2007;3:521–5. [PubMed: 17710092]
27. Guido NJ, Wang X, Adalsteinsson D, McMillen D, Hasty J, Cantor CR, Elston TC, Collins JJ. A bottom-up approach to gene regulation. *Nature* 2006;439:856–60. [PubMed: 16482159]
28. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 2000;403:339–42. [PubMed: 10659857]
29. Elowitz MB, Leibler S. A synthetic oscillatory network of transcriptional regulators. *Nature* 2000;403:335–8. [PubMed: 10659856]
30. Suel GM, Kulkarni RP, Dworkin J, Garcia-Ojalvo J, Elowitz MB. Tunability and noise dependence in differentiation dynamics. *Science* 2007;315:1716–9. [PubMed: 17379809]
31. Suel GM, Garcia-Ojalvo J, Liberman LM, Elowitz MB. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* 2006;440:545–50. [PubMed: 16554821]
32. Weinberger LS, Burnett JC, Toettcher JE, Arkin AP, Schaffer DV. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell* 2005;122:169–82. [PubMed: 16051143]
33. Isaacs FJ, Dwyer DJ, Ding C, Pervouchine DD, Cantor CR, Collins JJ. Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol* 2004;22:841–7. [PubMed: 15208640]
34. Masse E, Gottesman S. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci U S A* 2002;99:4620–5. [PubMed: 11917098]
35. Moller T, Franch T, Udesen C, Gerdes K, Valentin-Hansen P. Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev* 2002;16:1696–706. [PubMed: 12101127]
36. Winkler W, Nahvi A, Breaker RR. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 2002;419:952–6. [PubMed: 12410317]
37. Soukup GA, Breaker RR. Relationship between internucleotide linkage geometry and the stability of RNA. *Rna* 1999;5:1308–25. [PubMed: 10573122]
38. Liang ST, Xu YC, Dennis P, Bremer H. mRNA composition and control of bacterial gene expression. *J Bacteriol* 2000;182:3037–44. [PubMed: 10809680]
39. Xia B, Etchegaray JP, Inouye M. Nonsense mutations in *cspA* cause ribosome trapping leading to complete growth inhibition and cell death at low temperature in *Escherichia coli*. *J Biol Chem* 2001;276:35581–8. [PubMed: 11457833]

40. Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 2006;441:840–6. [PubMed: 16699522]
41. Regnier P, Arraiano CM. Degradation of mRNA in bacteria: emergence of ubiquitous features. *Bioessays* 2000;22:235–44. [PubMed: 10684583]
42. Arkin A, Ross J, McAdams HH. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* 1998;149:1633–48. [PubMed: 9691025]
43. Grossman AD. Genetic networks controlling the initiation of sporulation and the development of genetic competence in *Bacillus subtilis*. *Annu Rev Genet* 1995;29:477–508. [PubMed: 8825484]
44. Kemkemer R, Schrank S, Vogel W, Gruler H, Kaufmann D. Increased noise as an effect of haploinsufficiency of the tumor-suppressor gene neurofibromatosis type 1 in vitro. *Proc Natl Acad Sci U S A* 2002;99:13783–8. [PubMed: 12368469]
45. Maughan H, Nicholson WL. Stochastic processes influence stationary-phase decisions in *Bacillus subtilis*. *J Bacteriol* 2004;186:2212–4. [PubMed: 15028708]
46. Wolf DM, Vazirani VV, Arkin AP. Diversity in times of adversity: probabilistic strategies in microbial survival games. *J Theor Biol* 2005;234:227–53. [PubMed: 15757681]
47. Giuliodori AM, Brandi A, Gualerzi CO, Pon CL. Preferential translation of cold-shock mRNAs during cold adaptation. *RNA* 2004;10:265–276. [PubMed: 14730025]
48. Fiedler F, Mallo GV, Bodeker H, Keim V, Dagorn JC, Iovanna JL. Overexpression of the PC3/TIS21/BTG2 mRNA is part of the stress response induced by acute pancreatitis in rats. *Biochem Biophys Res Commun* 1998;249:562–5. [PubMed: 9712737]
49. Brandi A, Spurio R, Gualerzi CO, Pon CL. Massive presence of the *Escherichia coli* ‘major cold-shock protein’ CspA under non-stress conditions. *Embo J* 1999;18:1653–9. [PubMed: 10075935]
50. Yamanaka K, Inouye M. Induction of CspA, an *E. coli* major cold-shock protein, upon nutritional upshift at 37 degrees C. *Genes Cells* 2001;6:279–90. [PubMed: 11318871]
51. Welch WJ, Kang HS, Beckmann RP, Mizzen LA. Response of mammalian cells to metabolic stress; changes in cell physiology and structure/function of stress proteins. *Curr Top Microbiol Immunol* 1991;167:31–55. [PubMed: 2055098]
52. Wei LY, Roepe PD. Low external pH and osmotic shock increase the expression of human MDR protein. *Biochemistry* 1994;33:7229–38. [PubMed: 7911681]
53. Jones PG, VanBogelen RA, Neidhardt FC. Induction of proteins in response to low temperature in *Escherichia coli*. *J Bacteriol* 1987;169:2092–5. [PubMed: 3553157]
54. Colon-Ramos DA, Shenvi CL, Weitzel DH, Gan EC, Matts R, Cate J, Kornbluth S. Direct ribosomal binding by a cellular inhibitor of translation. *Nat Struct Mol Biol* 2006;13:103–111. [PubMed: 16429152]
55. Lawrence JG. Shared strategies in gene organization among prokaryotes and eukaryotes. *Cell* 2002;110:407–13. [PubMed: 12202031]
56. Dandekar T, Snel B, Huynen M, Bork P. Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem Sci* 1998;23:324–8. [PubMed: 9787636]
57. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MC, Withers ST, Shiba Y, Sarpong R, Keasling JD. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 2006;440:940–3. [PubMed: 16612385]

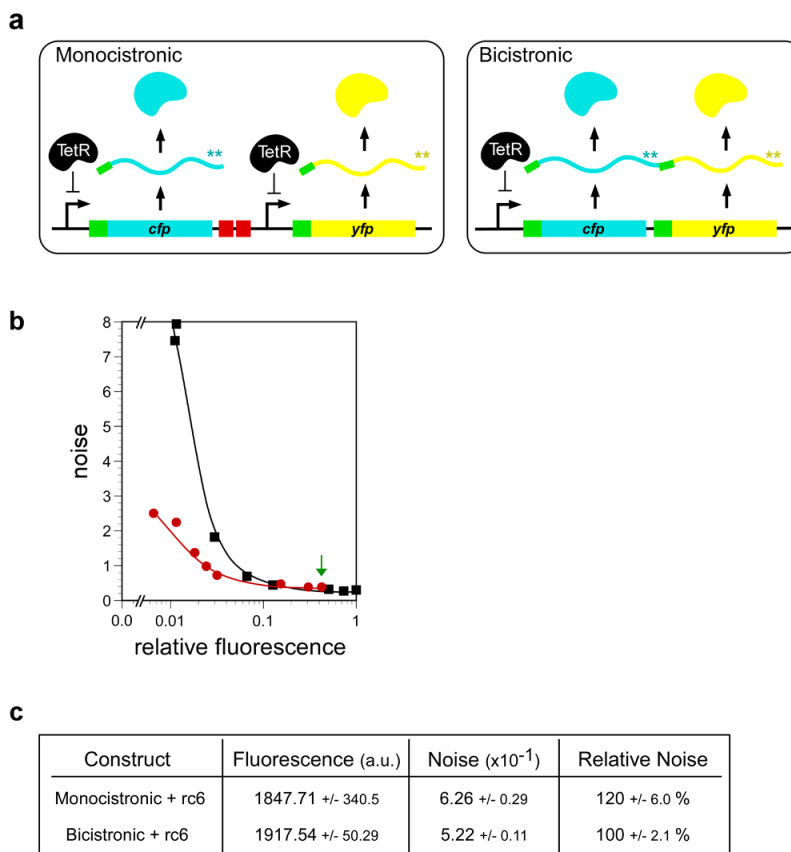


**Figure 1.**

rcRNA strategy for ribosome competition. (a) Cellular mRNAs (blue) associate with ribosomes probabilistically at a rate dependent upon the concentration of the mRNA and ribosomes in that cell. (Top) "Wild-type" gene expression scenario. (Bottom) An exogenous rcRNA competes with bulk cellular mRNAs for translational machinery. A reduction in the number of available ribosomes results in a decreased probability of a given cellular mRNA associating with a ribosome. (b) A series of rcRNAs based on rc1 (shown) were engineered to contain ribosome binding sites (green) of increasingly single-stranded nature. Mutations of 2, 3, 4 or 6 mismatches were made in the antisense stem (bracketed) such that the region in and around the RBS became destabilized. (c) In-line probing of rcRNA structure. Each rcRNA molecule was transcribed *in vitro*. The 5'-most nucleotide was then radioactively labeled with a phosphate group containing a <sup>32</sup>P atom and the RNAs were incubated under conditions which promote spontaneous hydrolysis of the phosphodiester backbone<sup>37</sup>. Nucleotide residues which tend to be single-stranded undergo hydrolysis significantly faster than nucleotides involved in a base pair. Hydrolysis results in two truncated RNAs, a labeled 5'-fragment and the remaining 3' fragment. The RNA population is then separated on a polyacrylamide gel matrix and runs as a pattern of degradation products of different sizes, based on the position of the hydrolysis event relative to the 5'-end of the RNA. Nucleotide residues which are more single stranded show up as more intense bands while residues which are more double stranded show less intense bands. Equivalent counts of <sup>32</sup>P containing RNA were added to each lane of a 10% denaturing polyacrylamide gel. Schematic of the rcRNA transcript is shown at right. (d) *In vitro* translation competition assays. A fixed amount of *gfp* DNA template was added to a coupled *in vitro* transcription/translation reaction along with increasing molar ratios of rcRNA DNA templates. The unpaired rcRNA, rc6, is represented as grey bars while rc1 is represented as white bars. Data are normalized to a *gfp* only control. Error bars represent 95% confidence intervals derived from 3 experiments run in parallel.



**Figure 2.** The effect of ribosome competing RNAs (rcRNAs) on gene expression. (a) Flow cytometric histograms of *E. coli* populations expressing GFP as well as a single rcRNA variant. rcRNA variants expressed (near to far): rc1, rc3, rc2, wild-type (uninduced cells carrying the plasmid for rc1), rc6, rc5, rc4. (b) relative GFP fluorescence versus noise (standard deviation divided by mean protein abundance) from the flow cytometry data in panel 'a'. The data were fit to an equation of the form  $y = ax^b$  where the value of  $b$  is  $-0.83$  (dashed line). Three cultures of each sample were grown in parallel and assayed under the conditions described in the Experimental. Error bars represent 95% confidence intervals. (c) Relative abundance of *gfp* mRNA in *E. coli* treated with 200ng/mL anhydrotetracycline (aTc) as compared to cells treated with no inducer in the presence of a high affinity (rc6) or low affinity (rc1) rcRNA. RNA was prepared and quantitated from three cultures grown in parallel under the conditions described in Experimental. Error bars represent 95% confidence intervals.



**Figure 3.** Noise in monocistronic and bicistronic expression platforms. (a) Schematic for engineered monocistronic (left) or bicistronic (right) CFP/YFP expression platforms. The monocistronic plasmid, pASKJ13065, was assembled such that both genes were expressed from the same TetR-repressed promoter (Ptet), carried the same RBS (green), and were followed by the same transcription terminator (not shown after YFP), but were transcribed independently from the same plasmid. On both plasmids, both genes contained double stop codons at the end of the respective open reading frame (asterisks shown in bicistronic version). The bicistronic plasmid, pASKJ13004, was designed exactly as pASKJ13065 except that both genes were expressed from a single Ptet promoter. (b) Intrinsic noise of mono and bicistronic constructs. *E. coli* carrying the monocistronic (black squares) and bicistronic (red dots) plasmids were induced to different extents by the addition of increasing amounts of the inducer anhydrotetracycline (aTc; between 0 and 500ng/mL) to the media. Reporter gene expression and noise were calculated as described in the Experimental. Relative fluorescence indicates CFP values from which *E. coli* autofluorescence and YFP bleedthrough were corrected. The data used to compare low level expression from the two constructs is the point at which each produces 0.012 relative fluorescence units (2<sup>nd</sup> point from left in each data set). Lines connecting data points are guides to the eye. (c) CFP fluorescence and intrinsic noise data in *E. coli* populations expressing mono or bicistronic CFP and YFP as well as the noisy rcRNA, rc6. High level CFP/YFP expression was matched between cells carrying the two constructs by inducing pASKJ13004 carrying cells with 250ng/mL aTc (equivalent to green arrow in Figure 3b) and pASKJ13065 carrying cells with 37.5 ng/mL aTc. Noise was quantitated from three cultures grown in parallel. Error values represent 95% confidence intervals.

**Table 1**

rcRNA design and function. Name of rcRNA variant, number of bulges rationally designed into the anti-RBS stem, GFP fluorescence and noise in the expression of GFP in from pASKIBA3.GFPm2 in *E.coli* strain BL21DE3 expressing each rcRNA variant are noted. Wild-type refers to an uninduced strain carrying the plasmid for rc1. Noise is quantified here as the standard deviation divided by the mean. 95% confidence intervals are shown for GFP fluorescence and noise.

Construct	Designed bulges	Relative GFP ( $\times 10^{-1}$ )	Noise ( $\times 10^{-1}$ )
Wild-type		5.8 $\pm$ 0.1	2.7 $\pm$ 0.1
rc1	0	10 $\pm$ 0.1	5.3 $\pm$ 0.04
rc2	2	7.6 $\pm$ 0.9	4.1 $\pm$ 0.7
rc3	4	9.1 $\pm$ 0.1	4.9 $\pm$ 0.1
rc4	4	0.4 $\pm$ 0.1	55 $\pm$ 3.5
rc5	6	0.8 $\pm$ 0.1	28 $\pm$ 4.2
rc6	6	1.2 $\pm$ 0.007	28 $\pm$ 0.2