Frequency domain analysis of noise in autoregulated gene circuits

Michael L. Simpson[†‡§¶], Chris D. Cox[‡‖], and Gary S. Sayler[‡††]

[†]Molecular Scale Engineering and Nanoscale Technologies Research Group, Oak Ridge National Laboratory, P.O. Box 2008, MS 6006, Oak Ridge, TN 37831-6006; and [‡]Center for Environmental Biotechnology and Departments of [§]Materials Science and Engineering, [‖]Civil and Environmental Engineering, and [††]Microbiology, University of Tennessee, Knoxville, TN 37996

Edited by Charles S. Peskin, New York University, New York, NY, and approved February 6, 2003 (received for review October 10, 2002)

We describe a frequency domain technique for the analysis of intrinsic noise within negatively autoregulated gene circuits. This approach is based on the transfer function around the feedback loop (loop transmission) and the equivalent noise bandwidth of the system. The loop transmission, *T*, is shown to be a determining factor of the dynamics and the noise behavior of autoregulated gene circuits, and this *T*-based technique provides a simple and flexible method for the analysis of noise arising from any source within the gene circuit. We show that negative feedback not only reduces the variance of the noise in the protein concentration, but also shifts this noise to higher frequencies where it may have a negligible effect on the noise behavior of following gene circuits within a cascade. This predicted effect is demonstrated through the exact stochastic simulation of a two-gene cascade. The analysis elucidates important aspects of gene circuit structure that control functionality, and may provide some insights into selective pressures leading to this structure. The resulting analytical relationships have a simple form, making them especially useful as synthetic gene circuit design equations. With the exception of the linearization of Hill kinetics, this technique is general and may be applied to the analysis or design of networks of higher complexity. This utility is demonstrated through the exact stochastic simulation of an autoregulated two-gene cascade operating near instability.

There has been considerable interest in the modeling and simulation of genetic circuits (1–6). This modeling has aided in the design of synthetic genetic circuits (7–9), which may prove helpful in understanding natural complexity, or may find applications in biosensors or other whole cell devices (10). The stochastic properties of these systems are of particular interest in modeling (1, 11–14), but also in some experimental studies (14–17). The molecules of interest, small signaling molecules, mRNA, and the resulting proteins are often present in low abundance within individual cells, giving rise to wide spatial and temporal variations in the concentration of these molecules.

Several strategies for analyzing or simulating the stochastic properties of genetic circuits have been reported (1, 11–15). Most often the results are given as signal-to-noise ratio, noise strength, stability parameters, or a time history of molecular concentration. Lost or hidden within such results are the frequency domain features of the noise, which are quite important as noise cascades through subsequent circuits. Maintaining the frequency domain features is especially important in autoregulated gene circuit analysis because feedback impacts both magnitude and frequency composition of the noise. There are electronic circuit-processing schemes that optimize performance by shifting noise into a frequency regime where it has a smaller impact on total system performance (e.g., sigma-delta analog-to-digital converters). Similar schemes could be used in engineered genetic circuits and may have evolved within natural genetic circuits. A complete analysis requires that the frequency composition of the noise be preserved.

Here, we apply frequency domain techniques by using the loop transmission concept to elucidate the noise performance of genetic circuits while maintaining critical frequency information. The assumption of system linearity is the most significant limitation of the methods presented here. However, linearity in a limited neighborhood around a steady state condition has been assumed in other analyses, and results that agree well with exact stochastic simulations fully accounting for circuit nonlinearity were obtained for at least some regions of operation (12, 14). This analysis reveals important relationships between circuit parameters that have implications for the noise performance of autoregulated circuits. The resulting equations have simple forms, making them useful as design equations, and lead to straightforward interpretations and to possible explanations for the range and ordering of parameter values found in natural genetic circuits.

Open Loop Gene Circuit

Fig. 1*a* is a schematic diagram of an unregulated single gene system. This discrete stochastic system is most accurately represented by a chemical master equation that defines the time evolution of the probabilities of finding the system in particular discrete states (1). However, the Langevin approach, which models the system with coupled continuous differential equations with additive noise terms, is often used in the analysis of these systems (1, 14). Although it is an approximate analysis, which loses validity when the number of mRNA or protein molecules is small, this approach is often solved with much greater analytical ease than other representations.

Noise properties of the systems in Fig. 1 were recently analyzed and modeled (12), and this was followed by a Langevin analysis and experimental validation of the model (14). The Langevin equations for this system are

$$\frac{dr}{dt} = -\gamma_R r + k_R(t) + \eta_R \qquad [1]$$

$$\frac{dp}{dt} = -\gamma_P p + k_P r + \eta_P, \qquad [2]$$

where r and p are mRNA and protein concentrations, γ_R and γ_P are mRNA and protein decay rate constants, k_R is the transcription rate, k_P is the translation rate constant, and η_R and η_p are white noise sources (14). At steady state, the average mRNA ($\langle r \rangle$) and protein ($\langle p \rangle$) populations found from Eqs. **1** and **2** are k_R/γ_R and $k_R k_R/\gamma_R\gamma_P$, respectively.

The random timing and discrete nature of each biosynthetic and molecular decay event generate four distinct noise sources associated with mRNA synthesis, mRNA decay, protein synthesis, and protein decay. Each of these is analogous to electronic shot noise arising from discrete charge carriers crossing a semiconductor junction at random times. Shot noise has been extensively analyzed and is usually considered in terms of its frequency composition, or power spectral density (PSD). The single-sided (i.e., positive frequency only) PSD is given by $2qI_{\text{ave}}$ (16), where q is the discrete unit

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: PSD, power spectral density.

[¶]To whom correspondence should be addressed. E-mail: simpsonML1@ornl.gov.

BIOPHYSICS

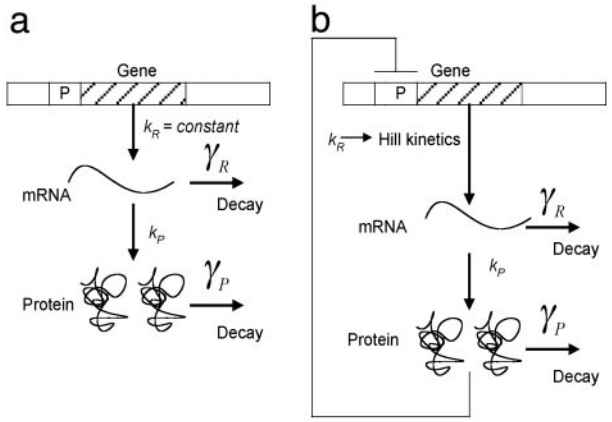

Fig. 1. Model of single gene expression. (*a*) mRNA molecules are synthesized from the template DNA strand at rate k_R, and proteins are translated at a rate of k_P off of each mRNA molecule. γ_R and γ_P are the decay rates for mRNA and protein respectively. (*b*) The same as *a* except that the protein is negatively autoregulated such that $k_R = k_{R_{max}}/[1 + (p/k_d)^n]$, where $k_{R_{max}}$ is the maximum rate of transcription, k_d is the concentration of protein where $k_R = k_{R_{max}}/2$, and n is the Hill coefficient.

of the current carrier and I_{ave} is the average current. We infer the power spectral density (we use the term "power" here in analogy to electronic noise) for mRNA and protein synthesis and decay from the shot noise relationship. The average rate of mRNA synthesis is k_R ($=I_{ave}$) produced in discrete units of one molecule ($=q$), giving a single-sided PSD of $2k_R$·[(molecules)^2/(s^2·Hz)]. Ideally the shot noise PSD would be constant with frequency (i.e., a white spectrum), although any real PSD has a limited extent in frequency. The dead time between mRNA synthesis events [e.g., short interval of blockage as an RNA polymerase open complex clears the promoter region (11)] band-limits this noise. However, other frequency response-determining processes, such as protein and mRNA decay, are the more limiting, and a white spectrum can be assumed with no loss of accuracy.

The average decay and synthesis rates are equivalent at steady state, giving a PSD for the mRNA decay noise source that is also $2k_R$. If the population of molecules is large enough, synthesis and decay processes are essentially uncorrelated, and the combined single-sided PSD for both mRNA noise sources, S_{RR}, is

$$S_{RR} = 2k_R + 2k_R = 4k_R. \quad [3]$$

By using a similar argument,

$$S_{PP} = 2\,\frac{k_R k_P}{\gamma_R} + 2\,\frac{k_R k_P}{\gamma_R} = 4\,\frac{k_R k_P}{\gamma_R}, \quad [4]$$

where S_{PP} is the combined single-sided PSD for both protein noise sources. The factors of four on the right hand side of Eqs. **3** and **4** arise from the selection of a single-sided PSD (a factor of two) and a second factor of two from the summation of the two uncorrelated noise terms associated with synthesis and decay. This lack of correlation between synthesis and decay noise terms requires that synthesis event timing does not affect decay event timing. At very low populations, this assumption may not be valid.

The output (i.e., protein concentration) PSD is found by summation of noise source PSDs as modified by the gene circuit (see Eq. **A-2** in *Appendix*). The frequency domain transfer functions ($A(f)$ in *Appendix*) for sources located at the point of protein ($A_P(f)$) and mRNA ($A_R(f)$) productions are found by Fourier transform and solution of Eqs. **1** and **2** to obtain

$$A_P(f) = \frac{A_P(0)}{\left(1 + i\,\frac{f}{f_{\text{protein}}}\right)}, \quad A_P(0) = \frac{1}{\gamma_P} \quad [5]$$

$$A_R(f) = \frac{A_R(0)}{\left(1 + i\,\frac{f}{f_{\text{mRNA}}}\right)\left(1 + i\,\frac{f}{f_{\text{protein}}}\right)}, \quad A_R(0) = \frac{k_P}{\gamma_R \gamma_P} = \frac{b}{\gamma_P} \quad [6]$$

where the critical frequencies associated with mRNA ($f_{\text{mRNA}} = \gamma_R/2\pi$) and protein decay ($f_{\text{protein}} = \gamma_P/2\pi$) are known as the poles of the system and define the frequency behavior of the transfer function. $A_R(f) \approx A_R(0)$ when f is small compared with either pole frequency. At frequencies above the first pole (assuming the poles are well spaced), $|A_R|$ goes approximately as $1/f$ and the phase shift asymptotically approaches $-90°$. After the second pole, $|A_R|$ goes approximately as $1/f^2$ and the phase shift asymptotically approaches $-180°$. The term b ($=k_P/\gamma_R$) in Eq. **6** is the average number of proteins produced from each mRNA transcript [or "burst" rate (11)]. Because the effect of a source PSD on the output PSD goes down as $1/f^2$ after the first pole, the second pole is often neglected. Because mRNA typically decays much faster than the protein ($\gamma_R \gg \gamma_P$), we neglect the mRNA pole in this noise analysis. With this simplification, the noise bandwidth (Δf_N) for both noise sources can be approximated as (see *Appendix*)

$$\Delta f_N \approx \frac{\pi}{2} f_{\text{protein}} = \frac{\gamma_P}{4}, \quad [7]$$

and the variance of the output (protein population) noise follows directly (see Eq. **A-5** in *Appendix*) as

$$\begin{aligned}\sigma^2_{n\_\text{out}} &= (S_{PP}|A_P(0)|^2 + S_{RR}|A_R(0)|^2)\Delta f_N \\ &= \left(\frac{4k_R k_P}{\gamma_R \gamma_P^2} + \frac{4k_R k_P^2}{\gamma_R^2 \gamma_P^2}\right)\frac{\gamma_P}{4} = \left(\frac{k_P k_R}{\gamma_P \gamma_R} + \frac{k_P^2 k_R}{\gamma_P \gamma_R^2}\right) \\ &= \frac{k_R k_P}{\gamma_R \gamma_P}\left(1 + \frac{k_P}{\gamma_R}\right) = \langle p\rangle(1 + b), \end{aligned} \quad [8]$$

where we have used the relationships for $\langle p\rangle$ and b defined above.

The noise figures of merit are

$$\frac{\sigma^2_{n\_\text{out}}}{\langle p\rangle} = 1 + b, \quad [9]$$

$$\frac{\langle p\rangle}{\sigma_{n\_\text{out}}} = \sqrt{\frac{\langle p\rangle}{(1 + b)}} = \sqrt{\frac{k_R b}{\gamma_P(1 + b)}}, \quad [10]$$

Eq. **9** gives the noise strength and Eq. **10** gives the output signal-to-noise ratio with results that are in agreement with previous analysis (12, 14). With these definitions and foundation in the frequency domain analysis of noise in gene circuits, we now turn to the case of autoregulated gene circuits.

Autoregulated Gene Circuit

Feedback is applied by allowing the output signal to modulate the level of the input signal (Fig. 1*b*). A Hill repression function of the form $k_R = k_{R_{max}}/[1 + (p/k_d)^n]$ often describes the behavior of autoregulated gene circuits, where $k_{R_{max}}$ is the maximum rate of transcription, k_d is the protein population where $k_R = k_{R_{max}}/2$, and n is known as the Hill coefficient. To allow linear analysis, we do a Taylor series expansion of the Hill function around $\langle p\rangle$ to obtain $k_R = k_0 - \alpha p$ with $\alpha = dk_R/dp|_{p=\langle p\rangle}$ and $k_0 = k_{R_{max}}/[1 +$

$(\langle p\rangle/k_d)^n] + \alpha\langle p\rangle$. The higher order terms in the Taylor series have been neglected because we assume only small excursions around $\langle p\rangle$. This linear approximation is most accurate for small variations around $\langle p\rangle \approx k_d$. At low protein population, the continuous representation and the small excursion around $\langle p\rangle$ assumptions in the series expansion are not valid. The Hill expression assumes a change in p produces an instantaneous change in k_R, and this assumption is usually followed in gene circuit analysis (8, 12, 14). Whereas the repression dynamics are likely to be much faster than other processes in the circuit, the frequency response of repression may play a role in second order effects. Although it will be neglected in portions of the analysis to follow, in general we will assume that α has a frequency dependency.

To deal with feedback, we introduce the loop transmission concept (see *Appendix*). The loop transmission, T, is the transfer function around the loop and may be thought of as a measure of the resistance of the feedback loop to variation. T is calculated by introducing a perturbation (Δ) at any point within the circuit (e.g., a small change in transcription rate) and measuring the response (ρ) that returns to the same point (e.g., a reactionary change in transcription rate). $T(f)$ is given by $\rho(f)/\Delta(f)$, and the feedback is negative if $T(0)$ is negative (i.e., has a phase of $\pm 180°$). $T(f)$ for the gene circuit of Fig. 1*b* is

$$\begin{aligned} T(f) &= A_R(f)\alpha(f) \\ &= \frac{A_R(0)\alpha(0)}{\left(1 + i\,\frac{f}{f_{\text{protein}}}\right)\left(1 + i\,\frac{f}{f_{\text{mRNA}}}\right)\left(1 + i\,\frac{f}{f_\alpha}\right)} \\ &= \frac{T(0)}{\left(1 + i\,\frac{2\pi f}{\gamma_P}\right)\left(1 + i\,\frac{2\pi f}{\gamma_R}\right)\left(1 + i\,\frac{f}{f_\alpha}\right)}, \end{aligned} \qquad [11]$$

where $T(0)$ is $-b\alpha(0)/\gamma_P$, and $A_R(f)$, f_{protein}, and f_{mRNA} are as described in the previous section. We assume a single pole (f_α) to account for repression dynamics. Inspection of the loop transmission allows interesting observations about this circuit. Because each pole will add $-90°$ of phase shift when fully in effect, two poles in T well before f_c (see *Appendix*), will produce an oscillator or a latch, but not a stable system, because the third pole will add enough phase shift to create positive feedback. The typically large difference between mRNA and protein decay rates allows stable negative feedback systems with a relatively large $|T(0)|$. This finding is significant, because most of the benefits of negative feedback systems are accentuated at larger magnitudes of T as shown in the analysis to follow.

Because the steady-state value of the synthesis and decay terms are altered by feedback, Eqs. **3** and **4** must be modified as follows

$$S_{RR}(f) = 4(k_0 - \alpha(0)\langle p\rangle) = 4\left(k_0 - \frac{\alpha(0)bk_0}{\gamma_P\left(1 + \frac{\alpha(0)b}{\gamma_P}\right)}\right) = S_{RR} \qquad [12]$$

$$S_{PP}(f) = \frac{4bk_0}{1 + \frac{\alpha(0)b}{\gamma_P}} = S_{PP}, \qquad [13]$$

where the autoregulated steady state value of $\langle p\rangle = bk_0/\{\gamma_P[1 + (\alpha(0)b/\gamma_P)]\}$ was used. Assuming a reasonably large phase margin, we approximate this gene circuit as a single dominant pole system (see *Appendix*) with f_{protein} as the lowest frequency pole in T. The variance of the output (protein population) noise is found by using Eq **A-5** in *Appendix*

Table 1. Forward gain and feedback terms used to calculate the closed-loop gains in Eq. 14

	Forward gain ($A(0)$)	Feedback ($\beta(0)$)
mRNA	b/γ_P	$\alpha(0)$
Protein	$1/\gamma_P$	$\alpha(0)b$

$$\begin{aligned} \sigma^2_{n\_\text{out}} &= \frac{\pi(1 + T(0))f_{\text{protein}}}{2}\left(S_{pp}|A_{cl\_P}(0)|^2 + S_{RR}|A_{cl\_R}(0)|^2\right) \\ &= \frac{k_0 b(b+1)}{\gamma_P(1 + |T(0)|)^2}, \end{aligned} \qquad [14]$$

where A_{cl_R} and A_{cl_P} are the closed-loop gains for noise sources located at mRNA and protein synthesis points. The $T(0)$ in Eq. **11** and the forward gain and feedback terms (see *Appendix*) given in table 1 were used to calculate $A_{cl\_R}(0)$ and $A_{cl\_P}(0)$. As described in *Appendix*, a noise bandwidth of

$$\Delta f_N \approx \frac{\pi}{2}(1 + |T(0)|)f_{\text{protein}} \qquad [15]$$

was used in Eq. **14**. The noise figures of merit for the autoregulated system are

$$\frac{\sigma^2_{n\_\text{out}}}{\langle p\rangle} = \frac{(b+1)}{(1 + |T(0)|)} = \frac{\left(\frac{\sigma^2_{\text{total}}}{\langle p\rangle}\right)_{\text{unregulated}}}{(1 + |T(0)|)} \qquad [16]$$

$$\frac{\langle p\rangle}{\sigma_{n\_\text{out}}} = \sqrt{\frac{bk_0}{\gamma_P(1+b)}}. \qquad [17]$$

Discussion

To a large degree, loop transmission determines the noise performance of autoregulated systems. Eq. **16** indicates that the noise strength of the negatively autoregulated system is reduced by a factor of $1/(1 + |T(0)|)$ compared with the unregulated case. However, autoregulation has an additional but more subtle effect on noise behavior that is not explicitly shown by other analyses. The noise bandwidth of the autoregulated circuit is increased by a factor of $1 + |T(0)|$ compared with the unregulated case. Thus, not only are the variance and standard deviation of the noise reduced by feedback, the noise that remains is *shifted to higher frequencies*.

This spectral shift becomes especially important as noise cascades through complex networks of genes. As an example, consider an autoregulated protein (p_1) that controls the transcription rate for a second protein (p_2) that is not autoregulated (Fig. 2*a*). Depending on the value of γ_{P2}, the noise shifted to higher frequencies through the action of negative feedback in the first circuit may have a negligible effect on the total noise of p_2. This result leads to up to an additional factor of $1 + |T(0)|$ decrease in the effect of the noise in the p_1 concentration on the noise strength of p_2 as is shown graphically in Fig. 2*b* and illustrated through simulations in Fig. 3. In this figure, the ratio of the simulated noise strengths for two cases (p_1 without feedback/p_1 with feedback) of the circuit of Fig. 2*a* is plotted as a function of γ_{P2}. This noise strength ratio is shown both for total p_2 noise and with the intrinsic p_2 noise removed so that only the effect of the p_1 noise is seen. For this case, when $\gamma_{P2} \ll \gamma_{P1}$, the p_2 noise strength is seen to be reduced by a factor that approaches $(1 + |T(0)|)^2$ compared with the unregulated case. However, for larger values of γ_{P2} the higher frequency noise content of the p_1 circuit with feedback is not filtered from the p_2 noise content, and the advantage of feedback reduces to the factor of $1 + |T(0)|$ shown in Eq. **16**.

BIOPHYSICS

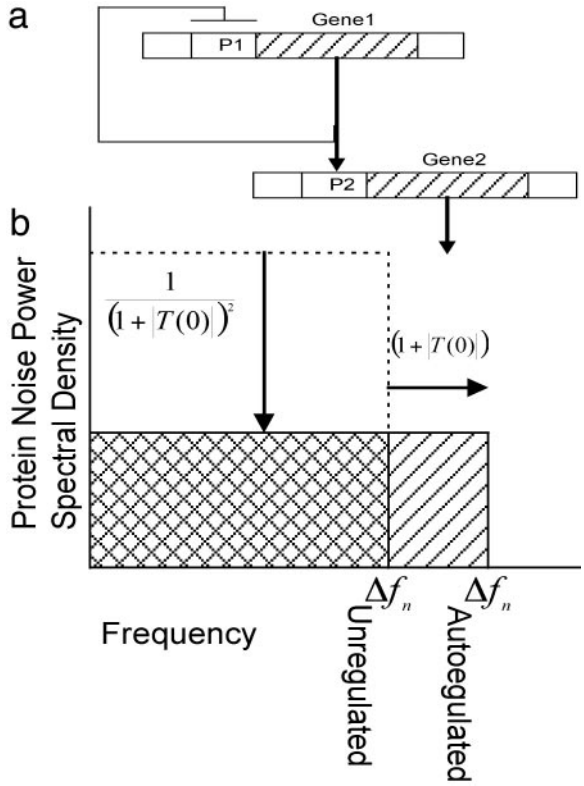

Fig. 2. Cascaded gene circuits. (*a*) Schematic representation of a two-gene cascade where the expression of the second gene is positively regulated by the concentration of the first protein such that $k_{R2} = k_{R_{max2}}/[1 + (k_{d2}/p_1)^n]$. The first gene circuit is negatively autoregulated as described in Fig. 1*b*. (*b*) Graphical representation of the change in the p_1 noise PSD caused by feedback in the p_1 circuit. If properly filtered by the second gene circuit in the cascade, the noise between the unregulated and regulated noise bandwidths may be rejected.

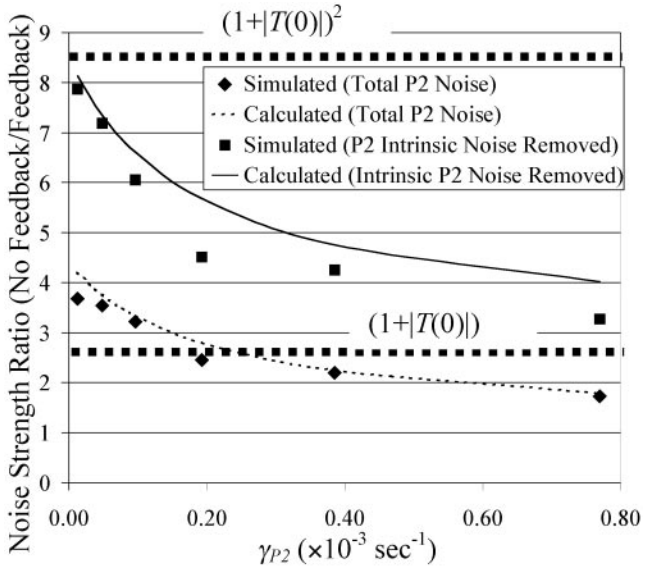

Fig. 3. Calculated and simulated ratio of p_2 noise strengths (p_1 circuit: without feedback (i.e., k_{R1} = constant)/with feedback) of the gene cascade in Fig. 2*a*. This ratio is plotted both for total p_2 noise and with the p_2 intrinsic noise removed, and shows that feedback may decrease the effect of p_1 noise by as much as $(1 + |T(0)|)^2$. The calculations were performed by using Eq. **16** for the regulated p_1 noise strength and Eq. **9** for the intrinsic p_2 and unregulated p_1 noise strengths. The total p_2 noise strength was found by adding the intrinsic p_2 noise strength to the p_1 noise strength multiplied by the noise power gain, $(dp_2/dp_1)^2|_{p_1 = \langle p_1 \rangle}$, and modifying the noise bandwidth to reflect the additional p_2 pole. The parameters for this circuit are as follows: $b_1 = 8$, $b_2 = 4$, $\gamma_{p1} = 0.0001925$/s, $\gamma_{R1} = 0.00289$/s, $\gamma_{R2} = 0.00578$/s, $k_{d1} = 800$, $k_{d2} = 600$, $k_{r_{max1}} = 0.0231$/s, $k_{r_{max2}} = 0.05$/s, $n_1 = n_2 = 7$. For the no feedback case, k_{R1} was set to a constant value of 0.0167/s, giving the same value of $\langle p_1 \rangle$, 700, for both feedback and no feedback cases. $\langle p_2 \rangle$ varied with the value of γ_{p2}, but was the same for feedback and no feedback cases. The intrinsic noise of p_2 was determined in simulation by setting $k_{R2} = 0.0367$/s whereas all other parameters remained unchanged. The simulation method is described in *Appendix*.

It is instructive to consider how far the reduction of noise by feedback may be taken. Within a single gene circuit, it is difficult to reach large values of T using physiologically relevant biosynthesis and decay rates. However, applying feedback around cascaded circuits (Fig. 4*a*) may provide large $|T|$, resulting in a large reduction in PSD and a significant modification of the noise spectrum. The cost is additional protein poles and reduced phase margin. With low phase margin approaching instability, there will be considerable peaking in the closed-loop frequency response near f_c (defined in *Appendix*). This frequency peaking will increase the variance of the noise, but this excess noise will be concentrated at higher frequencies as shown in the simulation results of Fig. 4*b*. This excess noise will have no impact on the PSD at lower frequencies, and, if filtered appropriately by downstream gene circuits, total noise can be significantly reduced through this shifting of noise power into the frequency peaking regime as demonstrated by the results of Fig. 4. For this case, the limit of this strategy is set by the stability criteria that phase margin must be greater than zero for the full range of parameters.

Fig. 4*b* further illustrates that the analytical methods presented here are easily extended to gene circuits of greater complexity such as the gene circuit cascade in Fig. 4*a*. Whereas the dominant single pole simplification cannot be used for such low phase margin systems, Eqs. **A-1** and **A-2** in *Appendix* can be applied to the analysis even as instability is approached. The PSDs calculated by using Eqs. **A-1** and **A-2** are shown in Fig. 4*b* and found to compare quite well with the PSDs derived from stochastic simulations for both open- and closed-loop cases.

As the analysis and simulations described above demonstrate, gene circuit structure heavily influences functionality. It is not just the magnitude and dynamics of feedback, but also the order of poles that has implications for the noise characteristics of autoregulated gene circuits. As described in *Appendix*, poles in the portion of the system in the feedback path from the output back to the noise source (the β term in *Appendix*) become zeros in the closed-loop system. In contrast to poles, for zeros, the magnitude of the closed-loop gain increases with f, and the phase asymptotically approaches +90°. If the protein concentration is considered the output, the zeros are at f_{mRNA} and f_α, and having the dominant pole in the system provided by protein decay minimizes this noise. The mRNA noise source has a zero at f_α, indicating that this noise source is minimized if f_α is the highest frequency pole in the system. Thus, to minimize protein noise, the sequence of poles would be protein decay, followed by mRNA decay, followed by f_α. For mRNA as an output signal, these closed-loop zeros would be in different locations, and a different ordering of poles would be needed to minimize mRNA noise. However, the regulatory elements are usually proteins, and perhaps it is not just coincidence that autoregulated gene circuit parameters usually follow a sequence that minimizes protein, not mRNA, noise. This pole sequence relationship is an important consideration for synthetic gene circuit design where the protein decay rate constant may be reduced (pole frequency raised) through genetic manipulations to provide proper circuit function (8).

In previous experimental studies, reporter gene activity from a series of individual cells was measured by using techniques such as

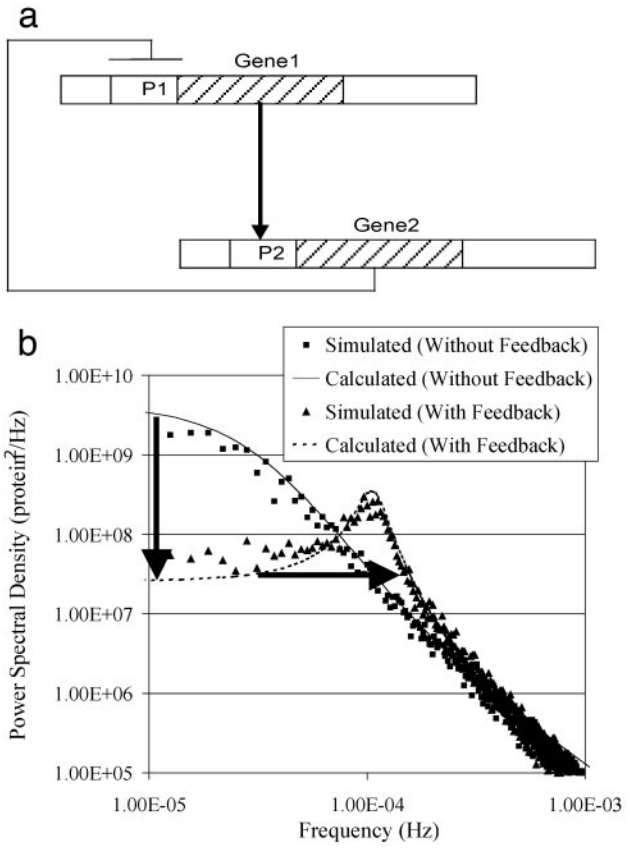

Fig. 4. Global feedback in a gene cascade. (*a*) Schematic representation of a two-gene cascade with both single-gene circuits within the feedback path such that $k_{R1} = k_{R_{max1}}/[1 + (p_2/k_{d1})^n]$ and $k_{R2} = k_{R_{max2}}/[1 + (k_{d2}/p_1)^n]$. (*b*) The calculated and simulated noise PSD for the circuit in *a* both with and without (i.e., k_{R1} = constant) feedback. The calculations were performed by using Eq. **A-1** and **A-2** from *Appendix*. The parameters for this circuit are as follows: $k_{r_{max1}} = k_{r_{max2}} = 0.0231$/s, $b_1 = b_2 = 13$, $\gamma_{p1} = \gamma_{p2} = 0.0001925$/s, $\gamma_{R1} = \gamma_{R2} = 0.00578$/s, $k_{d1} = k_{d2} = 350$, $n_1 = n_2 = 7$. For the no feedback case, k_{R1} was set to a constant value of 0.00449/s, giving the same value of $\langle p_2 \rangle$, 490, for both feedback and no feedback cases, allowing direct comparison of the PSDs. The simulation method is described in *Appendix*.

flow cytometry (14) or fluorescence microscopy (15). It is assumed that this phenotypic noise has a statistical equivalence to that obtained from a time sequence of reporter gene activity measured in an individual cell. However, the PSD is the Fourier transform of the autocorrelation function and can be found only by using time sequence data from an individual cell. Because very low frequency (less than mHz) measurements are needed to obtain a complete picture of the PSD, observations of several hours (or longer) and long cell cycle times may be required. However, some important frequency domain features (e.g., frequency peaking) could be seen with shorter observation times.

Limitations of the analysis technique presented here should be noted. As described earlier, at low molecular populations the Langevin representation, the shot noise analysis, and the Taylor series expansion of the Hill functions are not valid. Also, noise swings that create large excursions from $\langle p \rangle$ lead to inaccuracies in the analysis due to nonlinearity and saturation of the Hill function. Additionally, cell division and the random partitioning of molecules between daughter cells are not included here. Furthermore, the biochemical models used here are constructed as birth and death processes with first order rates. Higher order interactions between biochemical species will introduce more complex frequency dependencies, which may invalidate some

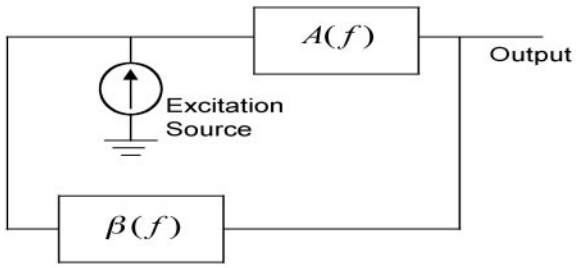

Fig. 5. Partitioning of a circuit to allow analysis by using the feedback formalism presented here. For a noise source at any location within the circuit, the transfer functions $A(f)$ and $\beta(f)$ are defined. $A(f)$ is the transfer function from the noise source location to the output (i.e., location where noise behavior is to be calculated) and $\beta(f)$ is the transfer function from this output back to the noise source location.

assumptions (e.g., single dominant pole). However, as long as the system can be linearized within a limited neighborhood of operation, the equations in *Appendix* can be used to determine the frequency domain behavior of the noise even as the circuit approaches instability. Whereas all of these issues affect the quantitative accuracy of the analysis for some circuits in some operational regions, the qualitative (i.e., circuit architecture) observations described above still have some validity.

Perhaps analysis tools like those presented here are best used to examine gene circuit structure–function relationships. These relationships may elucidate the selective pressures that influenced natural genetic circuit topology, or in synthetic biology applications, provide guidance in the selection of operational regimes to explore either through simulation or experiment. The frequency domain approach demonstrates the central role of loop transmission in noise behavior, the shifting of noise to higher frequency regimes through autoregulation, the tradeoff between noise reduction through negative autoregulation and excessive high frequency noise as phase margin is diminished, and the noise optimizing ordering of parameters that control dynamics. This understanding of simpler components allows an examination of the structure–function relationship at higher levels of gene circuit ordering. It has been suggested that there is an element of convergent evolution between complex engineered devices and biological systems (18), and a rich history of electronic system design may provide clues for understanding gene network topology. Autoregulated gene circuits almost certainly play an important role in gene networks (19), and the history of the development of feedback configurations for electronic systems may be a resource for developing an understanding of gene circuits and the selective pressure that drove their evolution. The extent to which this analogy is useful can be determined only by a great deal more analysis, modeling, and experimentation. Conversely, it is possible that new strategies for engineered system design may emerge from this examination. Certainly Black's development of electronic feedback amplifiers in the 1920s (20) may have benefited from an understanding of autoregulated gene function, and it is possible that presently unidentified gene network topologies may impact future system design strategies.

Appendix

Loop Transmission Analysis. Loop transmission analysis is used extensively to determine stability, closed-loop response, transient response, and noise behavior of linear systems. This analysis starts by casting the circuit into the form shown in Fig. 5. $A(f)$ (the forward gain) is the frequency-dependent transfer function from the excitation source (e_i) under consideration to the defined output of the circuit [i.e., $\partial \text{Output}/\partial e_i(f)$], whereas $\beta(f)$ is the frequency-dependent transfer function from this output back to the excitation source. The loop transmission, $T(f)$, is the transfer function around

the loop and is equal to $A(f)\beta(f)$. Excitation sources may be located anywhere within the circuit (e.g., different noise sources) leading to different partitioning of circuit components into either $A(f)$ or $\beta(f)$. Whereas location affects the closed-loop gain of a given excitation source, it does not change the loop transmission. Stability of the feedback circuit can be determined by inspection of the phase, $\varphi(f_c)$, of $T(f_c)$, where the crossover frequency, f_c, is defined by the relationship $|T(f_c)| = 1$. If $\varphi(f_c) > -180°$, then the system will reach a stable steady state. Otherwise, the output will oscillate or latch (i.e., reach a fixed state that does not respond to changes in the excitation signal). The phase margin of the system, φ_{PM}, is $\varphi(f_c) + 180°$, with larger phase margin leading to more stable system response.

The closed-loop gain, A_{cl}, is given by

$$A_{cl}(f) = \frac{1}{\beta(f)}\left(\frac{-T(f)}{1 - T(f)}\right), \qquad \textbf{[A-1]}$$

and the poles of $\beta(f)$ become zeros in the closed-loop response. For open-loop circuits ($\beta(f) = 0$) Eq. **A-1** reduces to $A(f)$. For multiple noise sources at m different locations within the circuit, the PSD of the output noise, S_{nn_o}, is

$$S_{nn\_o} = \sum_{j=1}^{m} A_{cl\_j}(f)A^*_{cl\_j}(f)S_{nn\_j}, \qquad \textbf{[A-2]}$$

where A_{cl_j} and S_{nn_j} are the closed-loop gain and PSD of the noise source at location j.

If the second pole in the system is approximately a factor of 2 or more greater than f_c ($\varphi_{PM} \approx 60°$) and $\beta(f)$ does not contain the first pole in the system, Eq. **A-1** can be approximated as a single dominant pole system given by

$$A_{cl}(f) = \frac{1}{\beta(0)}\left(\frac{-T(0)}{1 - T(0)}\right)\frac{1}{\left(1 + i\,\dfrac{f}{(1 + |T(0)|)f_{p1}}\right)}, \qquad \textbf{[A-3]}$$

where f_{p1} is the first (i.e., dominant) pole in T. Eq. **A-3** demonstrates that an important effect of negative feedback is an increase in system bandwidth. The noise bandwidth of the system is given by

$$\Delta f_N = \frac{\int_0^\infty A^*_{cl}(f)A_{cl}(f)df}{|A_{cl}(0)|^2}. \qquad \textbf{[A-4]}$$

For a single pole system $\Delta f_N = (\pi/2)f_p$, where f_p is the pole frequency. For this single dominant pole system, the noise variance at the output due to white noise sources within the system is given by

$$\sigma^2_{out} = \frac{\pi(1 + T(0))f_{p1}}{2}\sum_{j=1}^{m} S_{nn\_j}|A_{cl\_j}(0)|^2. \qquad \textbf{[A-5]}$$

Simulations. Stochastic simulations shown in Figs. 3 and 4 were conducted by using Gillespie's algorithm (21) for two-gene circuits in which the second gene was under positive regulation of the first. The system was assumed to be at quasi steady-state for $t > 50/k_{min}$ where k_{min} is the smallest rate constant in the simulation. The means and variances of gene product populations were calculated by sampling the quasi steady-state time series at a period of 10 s until the change in mean was less than 5×10^{-4} over a time span of 400,000 s. PSDs of the time series were calculated via fast Fourier transformation of 32,000 autocorrelation values with a sampling time of 10 s.

Parameter Selection for Example Gene Circuits. The parameters for the example gene circuits were chosen both to demonstrate key points of the analysis and to fit within a physiologically relevant range. The burst rates used in the examples ($b = 4$ to 13) were near that for *lacA* ($b = 5$) and well below the $b = 40$ of *lacZ* (22). The mRNA and protein decay rates where selected to yield half-lives of 2–4 min and 1 h, respectively (23). The wide variation in Hill kinetic parameters $k_{r_{max}}$, n, and k_d observed in natural systems reflects the stoichiometric and kinetic structure of repression and induction mechanisms. The parameters chosen here were selected from ranges used in other analyses of this type (12, 14) and to produce suitably large $|T(0)|$ that clearly demonstrate the effects of loop transmission on noise behavior. For the nonfeedback cases, values of k_R were selected to approximate the mean populations of p_1 for the corresponding feedback cases. The transcription rate used to calculate the intrinsic noise in the second gene in Fig. 3 was calculated assuming Hill kinetics and the mean population of p_1.

We thank M. J. Roberts, M. J. Doktycz, G. D. Peterson, J. M. Lancaster, M. S. Allen, and J. T. Fleming for several fruitful discussions related to the topics presented here. We gratefully acknowledge funding support from the National Science Foundation and the Defense Advanced Research Projects Agency through National Science Foundation Grant EIA-0130843. This work was partially performed at the Oak Ridge National Laboratory, managed by the University of Tennessee-Battelle, LLC, for the U.S. Department of Energy under Contract DE-AC05-00OR22725.

1. Rao, C. V., Wolf, D. M. & Arkin, A. P. (2002) *Nature* **420,** 231–237.
2. De Jong, H. (2002) *J. Comput. Biol.* **9,** 67–103.
3. Endy, D. & Brent, R. (2001) *Nature* **409,** 391–395.
4. Hasty, J., McMillen, D., Isaacs, F. & Collins, J. J. (2001) *Nat. Rev. Genet.* **2,** 268–279.
5. McAdams, H. H. & Arkin, A. (1998) *Annu. Rev. Biophys. Biomol. Struct.* **27,** 199–224.
6. Smolen, P., Baxter, D. A. & Byrne, J. H. (2000) *Bull. Math. Biol.* **62,** 247–292.
7. Hasty, J., McMillen, D. & Collins, J. J. (2002) *Nature* **420,** 224–230.
8. Elowitz, M. B. & Leibler, S. (2000) *Nature* **403,** 335–338.
9. Gardner, T. S., Cantor, C. R. & Collins J. J. (2000) *Nature* **403,** 339–342.
10. Simpson, M. L., Sayler, G. S., Fleming, J. T. & Applegate, B. (2001) *Trends Biotechnol.* **19,** 317–323.
11. McAdams, H. H. & Arkin, A. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 814–819.
12. Thattai, M. & van Oudenaarden, A. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 8614–8619.
13. Swain, P. S., Elowitz, M. B. & Siggia, E. D. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 12795–12800.
14. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. (2002) *Nat. Genet.* **31,** 69–73.
15. Becskei, A. & Serrano, L. (2000) *Nature* **405,** 590–593.
16. Schwartz, M. (1959) *Information Transmission, Modulation & Noise* (McGraw–Hill, New York).
17. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. (2002) *Science* **297,** 1183–1186.
18. Csete, M. E. & Doyle, J. C. (2002) *Science* **295,** 1664–1669.
19. Thieffry, D., Huerta, A. M., Perez-Rueda, E. & Collado-Vides, J. (1998) *Bioessays* **20,** 433–440.
20. Black, H. S. (1934) *Bell Syst. Tech. J.* **13,** 3.
21. Gillespie, D. T. (1977) *J. Phys. Chem.* **81,** 2340–2361.
22. Kennell, D. & Riezman, H. (1977) *J. Mol. Biol.* **114,** 1–21.
23. Stephanopoulos, G. N., Aristidou, A. A. & Nielsen, J. (1998) *Metabolic Engineering* (Academic, New York).