OMICS A Journal of Integrative Biology
Volume 7, Number 3, 2003

Analysis of Noise in Quorum Sensing

CHRIS D. COX,[1,2] GREGORY D. PETERSON,[2,3] MICHAEL S. ALLEN,[2,4] JOSEPH M. LANCASTER,[3] JAMES M. McCOLLUM,[3] DEREK AUSTIN,[3,4] LING YAN,[1,2] GARY S. SAYLER,[2,5] and MICHAEL L. SIMPSON[2,4,6]

ABSTRACT

Noise may play a pivotal role in gene circuit functionality, as demonstrated for the genetic switch in the bacterial phage λ. Like the λ switch, bacterial quorum sensing (QS) systems operate within a population and contain a bistable switching element, making it likely that noise plays a functional role in QS circuit operation. Therefore, a detailed analysis of the noise behavior of QS systems is needed. We have developed a set of tools generally applicable to the analysis of gene circuits, with an emphasis on investigations in the frequency domain (FD), that we apply here to the QS system in the marine bacterium *Vibrio fischeri*. We demonstrate that a tight coupling between exact stochastic simulation and FD analysis provides insights into the structure/function relationships in the QS circuit. Furthermore, we argue that a noise analysis is incomplete without consideration of the power spectral densities (PSDs) of the important molecular output signals. As an example we consider reversible reactions in the QS circuit, and show through analysis and exact stochastic simulation that these circuits make significant and dynamic modifications to the noise spectra. In particular, we demonstrate a "whitening" effect, which occurs as the noise is processed through these reversible reactions.

INTRODUCTION

IT IS NOW UNDERSTOOD that noise often plays a pivotal role in gene circuit functionality. In perhaps the most notable example to date, Arkin, Ross, and McAdams analyzed the coupling between a noise-modulated genetic switch and macroscopic phenotype selection using the λ phage lysis-lysogeny decision circuit (λ switch) as a model system (Arkin et al., 1998). The λ phage infects *E. coli* and chooses between lysogenic and lytic pathways. The lysogenic or lytic outcome in each cell is determined by a critical race (Mcadams and Shapiro, 1995) between the accumulations of two regulatory proteins, cI and Cro, after infection. If there is enough early production of cI, positive feedback drives cI production up, while negative

[1]Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, Tennessee.
[2]Center for Environmental Biotechnology, University of Tennessee, Knoxville, Tennessee.
[3]Department of Electrical and Computer Engineering, University of Tennessee, Knoxville, Tennessee.
[4]Molecular Scale Engineering and Nanoscale Technologies Research Group, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
[5]Department of Microbiology, University of Tennessee, Knoxville, Tennessee.
[6]Department of Materials Science and Engineering, University of Tennessee, Knoxville, Tennessee.

feedback drives Cro production down, leading to the choice of the lysogenic pathway. This happens only when, by chance (i.e. noise), there is early and strong cI production.

The probabilistic nature of the λ switch provides the means for an interesting λ phage population control mechanism. Various studies have shown that the nutritional state of the cell and the level of the phage population around the cell influence the fraction of λ-infected E. coli cells that become lysogenic. These signals influence the λ switch critical race by affecting the early production of cI, which modulates the fractional split of lytic and lysogenic outcomes. This use of noise is reminiscent of dithering, a scheme that adds noise to analog-to-digital converter circuits to provide an advantage in a collective sense. The λ switch only provides one-bit resolution within individual cells. However, in a manner similar to dithering, the noise and the critical race condition allow a nearly continuous selection of the fraction of cells that become lysogenic within a population, thus yielding precise control from very imprecise components. It is important to note that this favorable (for the phage) behavior arises within a population, not within individual cells, and the noise is an essential element in this emergent functionality of this gene circuit. An understanding of this and similar gene circuit architectures requires a comprehensive understanding of the nature of the noise and how it is processed by gene circuits.

Noise analysis in genetic circuits usually is performed using exact stochastic simulation (Gillespie, 1977), or approximate solutions are found using the Langevin approach (Rao et al., 2002). Exact stochastic simulation provides the most accurate analysis as it deals with discrete molecules of each species, microscopic rate constants, and circuit non-linearities. However, even for circuits of only modest complexity, this simulation can be computationally demanding, and does not generally lead to the easy coupling of simulated noise behavior with specific circuit elements, parameters, and operational regimes. On the other hand, the Langevin approach approximates stochastic processes in cells by adding small white noise terms to the ordinary differential equations (ODEs) associated with the chemical reactions (Rao et al., 2002). Although the Langevin equations are much easier to deal with analytically than exact stochastic simulation, there are several important caveats. At low molecular populations the Langevin representation and the associated linearization and approximations may lose validity, and considerable care must be exercised in applying Langevin analysis to systems with pathway bifurcations (e.g., the λ switch described above). A naïve application of this analysis technique in such situations can lead to erroneous results that completely miss essential circuit function (Arkin et al., 1998). However, even though all of these issues affect the quantitative accuracy of the analysis for some circuits in some operational regions, qualitative (i.e., circuit architecture) observations still have some validity and often lead to insights into gene circuit topology.

The noise behavior of relatively simple gene circuits have been solved through the time domain solution of the associated Langevin equations (Ozbudak et al., 2002). However, for more complex circuits with multiple noise sources and signal processing elements, such analysis becomes cumbersome and obscures the intuitive connection between calculated noise behavior and circuit elements, parameters, or operational regimes that originally motivated the use of the Langevin approach instead of (or more likely in addition to) exact stochastic simulation. We recently reported a frequency domain (FD) analysis technique that explicitly retained the spectral features of the noise in the analysis (Simpson, 2003). Applied to negatively autoregulated genetic circuits, the FD analysis showed that feedback impacts both magnitude and frequency composition of the noise, and the response of downstream gene circuits to this noise exhibited a strong dependence on frequency composition (Simpson et al., 2003).

Here we report the use of the FD techniques in combination with optimized exact stochastic simulation tools for the analysis of gene circuit architecture. We have found this coupling between analytical and simulation approaches leads to a deeper understanding of the circuit architecture and strategy, and clearly identifies the most important relationships between circuit structure and function. For example, in addition to the effects on the noise magnitude and frequency composition, the FD analysis of the negatively autoregulated gene circuit showed that the arrangement of decay rates most often encountered in natural gene circuits (e.g., mRNA usually decays much faster than protein, Hill kinetics usually faster than mRNA decay) minimizes noise in the protein concentration, not the mRNA concentration (Simpson et al., 2003). As the regulatory elements are usually proteins and minimal noise in the regulatory system may be advantageous, perhaps this is not just coincidence. It is these types of insights that we seek from a coupled FD/simulation approach.

The bacterial quorum sensing (QS) system is a prime candidate for architectural analysis using FD/ exact stochastic simulation techniques. The QS circuit contains both positive and negative feedback, and a cascade of circuit blocks through which the spectra of the numerous noise sources are shaped. Furthermore, QS operates on a population, and as demonstrated in the λ switch example, noise may play a functional role in such circuits. However, the QS circuit does contain a bistable switching element, so care must be exercised when using FD analysis in the operational regime near the switching point.

RESULTS

Analysis and simulation of the quorum sensing circuit

QS is a cell-cell communication network that enables bacterial populations to collectively regulate their behavior based on cell density, serving as a way for bacteria to coordinate their metabolic efforts for specific functions such as infection onset, antibiotic production, or biofilm formation (Greenberg, 2000; Withers et al., 2001). This intercellular exchange relies on self-generated, low molecular weight diffusible signal molecules called autoinducers (AIs). AI molecules released by a single, free-living bacterium do not accumulate to high enough concentrations to be detected. However, when a sufficient bacterial population density is reached, AI concentrations achieve a threshold level that allows individual bacteria to coordinately activate or repress specific gene expression. QS was first identified in the marine bacterium *Vibrio fischeri*, where it controls expression of bioluminescence (Stevens and Greenberg, 1997). The *V. fischeri* QS network consists of AI molecules called N-acylhomoserine lactones (AHLs). AHLs are synthesized from precursors by a synthase protein, LuxI, and, upon reaching a critical threshold, interact with a transcriptional activating LuxR protein to induce expression of genes responsible for bioluminescence.

The *lux* regulon consists of two divergently transcribed operons, designated operon_L and operon_R, and exhibits complex regulatory behavior (Fig. 1). Operon_L consists of the *luxR* gene, which has a promoter region consisting of three transcriptional start sites designated L_1, L_2 and L_3. Operon_R contains *luxICDABEG*, the genes responsible for AI production (*luxI*) and bioluminescence (*luxCDABEG*). Binding of cyclic AMP–cAMP receptor protein (cAMP-CRP) complex to the crp regulatory region of the DNA positively regulates *luxR* transcrition at sites L_1 and L_2 and negatively regulates operon_R (Dunlap and Greenberg, 1988; Shadel and Baldwin, 1992a). L_3 is independent of cAMP-CRP (Shadel and Baldwin, 1992a). LuxR binds with AI to form a complex (designated $[RAI]_2$ in our model) that binds to the lux regulatory region, thereby positively regulating operon_R (Engebrecht et al., 1983) and all three of the transcriptional start sites of operonL (Shadel and Baldwin, 1992a). However, at very high expression levels, *luxR* has been shown to be capable of negative autoregulation (Dunlap and Greenberg, 1988). While the precise mechanism is unknown, it has been shown to be dependent on LuxR, AI, the lux regulatory region, and a regulatory element contained in the *luxD* gene (Shadel and Baldwin, 1991). The *luxD* element shares 11 of 20 bp with the lux regulatory region and functions as a LuxR binding site (Shadel and Baldwin, 1992b), possibly suggesting the formation of a DNA loop through interaction of $(RAI)_2$ complexes bound at the lux regulatory and the *luxD* element control regions. The loop could interfere with the binding of cAMP-CRP or of RNAp to the promoter of operator_L (Shadel and Baldwin, 1991).

On the basis of the regulatory behavior described above, we have developed a model for the *luxIR* QS circuit in *Vibrio fischeri*[a] that is depicted in Fig. 2 and formulated in terms of chemical reactions in Table 1; the rationale for parameter value selection will be discussed below. In this version of the model, we have assumed that the cell is growing in the absence of glucose such that cAMP-CRP is bound to the crp regulatory region resulting in a relatively high transcription rate of *luxR* (k_{crpR}) and low transcription rate of *luxI* (k_{basalI}). Upon binding of the $(RAI)_2$ complex to the lux operator, transcription of both operons increase (k_{maxR} and k_{maxI}). Transcription produces messenger RNA (mRNA), which is in turn translated to the two proteins LuxR and LuxI. Decay reactions exist for both mRNA and proteins. AI is produced from

[a]The Vibrio fischeri quorum sensing model described in this paper is available in SBML (Systems Biology Markup Language) level 2.0 format at: http://biocomp.ece.utk.edu.

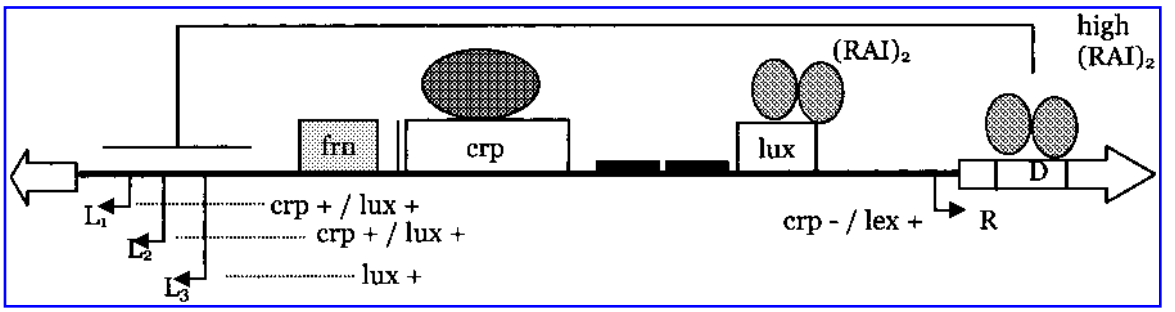


FIG. 1. Regulatory region and behavior of the *lux* regulon.

substrates (assumed not to be limiting) via LuxI. AI binds with LuxR to form the complex RAI, which in turn is assumed to undergo dimerization to form the final regulatory complex $(RAI)_2$. This complex binds to both the lux operator and to the regulatory region of *luxD*. When both DNA sites are occupied, the DNA looping reaction is allowed to occur, which completely stops transcription of LuxR. A pool of AI is assumed to exist external to the cell, which is allowed to freely diffuse in without depleting the amount of AI in solution. Diffusion of AI out of the cell is then modeled as decay.

Values of some parameters are available from the literature. The rate of autoinducer synthesase by LuxI (k_{AI}) is given by Schaefer (Schaefer et al., 1996). We assumed a typical mRNA half-life of 2 min to estimate γ_{mR} and γ_{mI}. (Stephanopoulos, 1998). AI has been shown to diffuse freely through the cell walls of

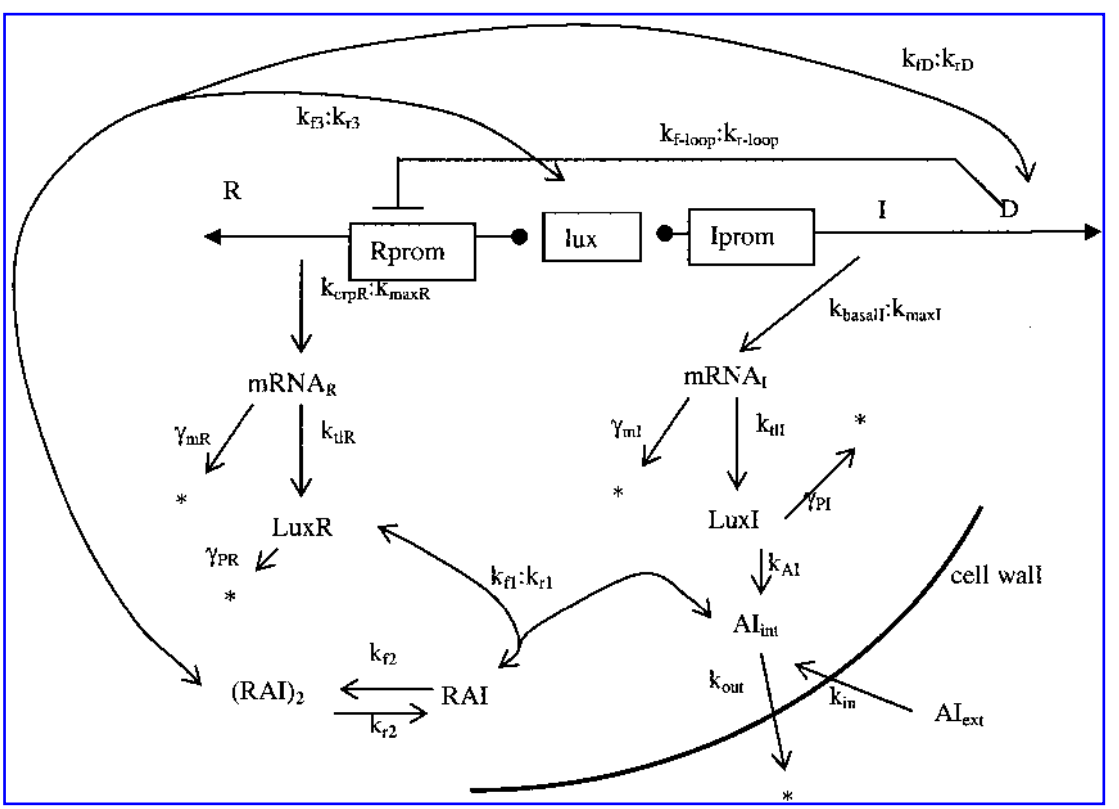


FIG. 2. The schematic representation of quorum sensing model.

Table 1. Model Reactions and Stochastic Rate Constants

Reaction	*Rate constant*
AI + LuxR → RAI	$k_{f1} = 0.1$
RAI → LuxR + AI	$k_{r1} = 2$
2 RAI → RAI2	$k_{f2} = 0.06$
RAI2 → 2 RAI	$k_{r2} = 4$
RAI2 + luxD → luxDcomp	$k_{fD} = 0.01$
luxDcomp → RAI2 + luxD	$k_{rD} = 3$
luxDcomp + ipromR → DNAloop	$k_{f\text{-}loop} = 6$
DNAloop → luxDcomp + ipromR	$k_{r\text{-}loop} = 1$
mRNAR → LuxR + mRNAR	$k_{t1R} = 0.03$
mRNAI → LuxI + mRNAI	$k_{t1I} = 0.03$
mRNAR → *	$\gamma_{mR} = 0.006$
mRNAI → *	$\gamma_{mI} = 0.006$
LuxI → *	$\gamma_{PI} = 0.001$
LuxR → *	$\gamma_{PR} = 0.006$
LuxI → LuxI + AI	$k_{AI} = 0.017$
AI → *	$k_{out} = 7.1$
AIsource → AI + AIsource	k_{in} = variable
luxbox + RAI2 + promI + promR → ipromI + ipromR	$k_{f3} = 0.1$
ipromI + ipromR → luxbox + RAI2 + promI + promR	$k_{r3} = 5$
promR → promR + mRNAR	$k_{crpR} = 0.01$
promI → promI + mRNAI	$k_{basII} = 0.000015$
ipromR → ipromR + mRNAR	$k_{maxR} = 0.06$
ipromI → ipromI + mRNAI	$k_{maxI} = 0.1$

Vibrio fischeri (Kaplan and Greenberg, 1985). Based on these results we estimated the stochastic rate constant for a reaction event that causes AI to diffuse from the cell as:

$$k_{out} = \frac{DA}{V\delta} \tag{1}$$

where D is the diffusion coefficient (5×10^{-10} m^2/sec), A is the external area of the cell (8×10^{-12} cm^2), V is the volume of the cell (1.6×10^{-18} m^3), and δ is the diffusion distance ($=V^{1/3}$). The value obtained (2134 sec^{-1}) was so large that virtually all the computation time would be consumed moving molecules in and out of the cell. It was decided to reduce the magnitude of this constant by a factor of 300. This should have negligible effect on the simulations since the value of the constant is still much greater than the stochastic rate constant for any other reaction directly involving AI. To verify this, the mean and standard deviation of AI populations were compared for simulations of the entire luxIR model and for a model that only included diffusion of AI in and out of the cell; they were not statistically different. A comparison of the mean and standard deviations of AI populations for the latter model using different values of k_{in} and k_{out} while keeping their ratio constant also yielded no significant differences.

Values of the remaining constants were determined by rough calibration to reproduce phenotypic behavior of *Vibrio fischeri* as reported in the literature using realistic constraints on parameter values. For example, the burst rate ($=k_{tl}/\gamma_m$) was constrained to be between 5 and 40 (Kennell and Riezman, 1977). These behaviors include (a) a rapid response (<30 min) to exogenously added AI (Kaplan and Greenberg, 1985), (b) an increase by almost 3 orders of magnitude of O_R gene products in response to a 20× increase in exogenous AI (Kaplan and Greenberg, 1985), and (c) induction of *luxR* at moderate AI levels but repression at high levels (Shadel and Baldwin, 1991, 1992b). Quasi steady-state analysis was used to identify kinetic parameters yielding desired behavior. Assumptions of quasi steady state are valid for reversible equilibrium reactions when these reactions occur much more frequently than the production and decay reactions for the proteins. We will further assume that AI is maintained at a constant level within the cell by diffusion from an external source (such as other cells). The model is formulated by considering the five possible states that

exist considering whether the lux box operator and the *luxD* regulatory element are occupied by $(RAI)_2$ or not: (1) both empty; (2) lux box empty, *luxD* occupied; (3) lux box occupied, *luxD* empty; (4) both filled; and (5) DNA loop formation. We will let S_i represent the fraction of time that the regulon is in each of the five states such that:

$$\sum_{i=1}^{5} S_i = 1 \tag{2}$$

The conditional probability of S_5 given that both regulatory regions are occupied is given by $k_{f\text{-}loop}/(k_{f\text{-}loop} + k_{r\text{-}loop})$ so that:

$$S_5 = \frac{k_{f\text{-}loop}}{k_{f\text{-}loop} + k_{r\text{-}loop}} \ (S_4 + S_5) \tag{3}$$

In a like manner, if we consider the conditional probability of the lux box being occupied given that there is no DNA loop we can write the equation:

$$S_3 + S_4 = (1 - S_5)\Omega_{lux} \tag{4}$$

where Ω_{lux} is the probability of occupation of the lux box neglecting the possibility of DNA loop formation and is given by:

$$\Omega_{lux} = \frac{1}{1 + \dfrac{K_d'}{[R]^2[AI_{\text{int}}]^2}} \tag{5}$$

where $[AI_{int}]$ is the average population number of AI internal to the cell and

$$K_d' = \frac{k_{r3}k_{r2}k_{r1}^2}{k_{f3}k_{f2}k_{f1}^2} = K_d \frac{k_{r2}k_{r1}^2}{k_{f2}k_{f1}^2} \tag{6}$$

The probability of the *luxD* site being occupied is handled in a parallel manner:

$$S_2 + S_4 = (1 - S_5)\Omega_{luxD} \tag{7}$$

$$\Omega_{luxD} = \frac{1}{1 + \dfrac{K_{dD}'}{[R]^2[AI_{\text{int}}]^2}} \tag{8}$$

where

$$K_{dD}' = \frac{k_{rD}k_{r2}k_{r1}^2}{k_{fD}k_{f2}k_{f1}^2}$$

Finally, a quasi-steady state balance on S1 yields:

$$S_1 = \frac{S_2k_{rD} + S_3k_{r3}}{\dfrac{k_{r2}k_{r1}^2}{k_{r1}k_{r2}^2}\ [R]^2[AI]^2(k_{f3} + k_{fD})} \tag{9}$$

Now the average *luxR* and *luxI* transcription rates for any population of R and AI_{int} can be written:

$$\left.\frac{d[mRNAR]}{dt}\right|_{transcription} = k_{crpR}(S_1 + S_2) + k_{\max R}(S_3 + S_4) \tag{10}$$

$$\left.\frac{d[mRNAI]}{dt}\right|_{transcription} = k_{basalI}(S_1 + S_2) + k_{\max I}(S_3 + S_4 + S_5) \tag{11}$$

Steady-state balances on R and I can now be written to yield:

$$R = \frac{k_{tlR}}{\gamma_{mR}\gamma_{pR}}\ [k_{crpR}(S_1 + S_2) + k_{\max R}(S_3 + S_4)] \tag{12}$$

$$I = \frac{k_{tlI}}{\gamma_{mI}\gamma_{pI}}\ [k_{basalI}(S_1 + S_2) + k_{\max I}(S_3 + S_4 + S_5)] \tag{13}$$

Equations 2–13 can be solved simultaneously to yield the steady-state solution. A data set by Kaplan and Greenburg (1985) showing the increase in bioluminescence as a function of exogenous AI was used to cal-

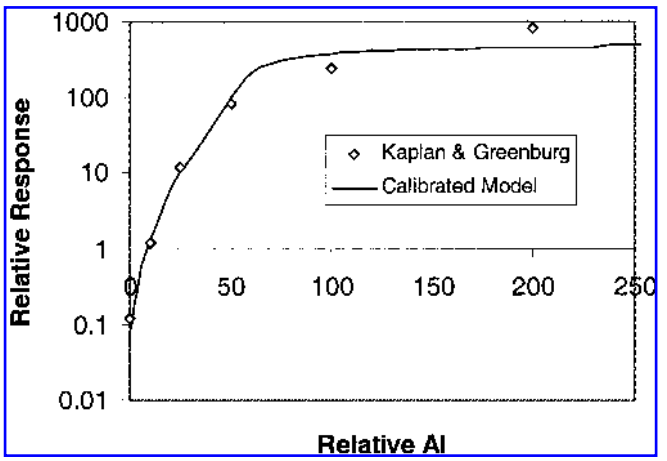


FIG. 3. Calibration of the steady-state analytical model to experimental data.

ibrate the model by adjusting the parameters to fit data. The parameters obtained are the same as the stochastic constants listed in Table 1 except the value of k_{f2} for the deterministic model was half of the stochastic version of the constant to account for reaction stoichiometery (Gillespie, 1977). Bioluminescence was used as an indicator of *luxI* transcription rates, since the genes are on the same operon. The steady-state analytical solution of the model and data are compared in Figure 3. At this point in time, only the relative response of the regulon is compared, since the set of parameters determined were not unique. Experiments underway to measure average LuxR levels within the cell will provide the data needed to further constrain the choices of parameters.

Another model calibration criteria was that both positive and negative autoregulation of LuxR has been described in the literature. The transcriptional rates of LuxR and LuxI calculated by equations (10) and (11) are shown in Figure 4. Transcription of *luxR* initially increases with AI population, but levels out and decreases slightly as a result of negative autoregulation via the formation of the DNA loop. The arrow indicates the occurrence of a fivefold over-expression of LuxR, resulting in a further decrease in *luxR* transcription rate, as has been reported in the literature (Dunlap and Greenberg, 1988; Shadel and Baldwin, 1991). The transcription rate of *luxI* is proportional to $S_3+S_4 + S_5$ and increases with AI. The transcription rate of *luxR* is proportional to $S_3 + S_4$ and any time spent in S_5 reduces its transcription rate and lowers the steady-state population of LuxR. Therefore, negative autoregulation of LuxR greatly extends the dynamic range of the operon_R response, because otherwise the induction rate would equal Ω_{lux}, which is proportional to$[\text{LuxR}]^2$.

Stochastic simulations were performed using ESS[b] (Exact Stochastic Simulator) an implementation of the Gibson and Bruck (2000) optimization of Gillespie's algorithm (Gillespie, 1977). A range of values (25 to 2000) for k_{in} was used to yield various AI levels within the cell. The results from stochastic simulation are compared to the deterministic steady-state calculations in Figure 5. Some differences between the deterministic and stochastic models are expected because each of the states is discrete, having a value of either 0 or 1. Nevertheless the two modeling approaches give nearly identical results. The stochastic model displays the same dynamic range that was seen in the deterministic model. This is an important result because is demonstrates that the model parameters determined using a deterministic model are useful in stochastic models, with appropriate adjustment. This is especially important for calibration, since calibration via stochastic simulation would be tedious beyond measure.

Stochastic simulations provide information concerning the noise behavior of the system. The *luxIR* system becomes up regulated when the $(RAI)_2$ complex binds to the lux box. Positive feedback systems are

[b]Open source code for the Exact Stochastic Simulator (ESS) is available at www.biospice.org or http://biocomp.ece.utk.edu.

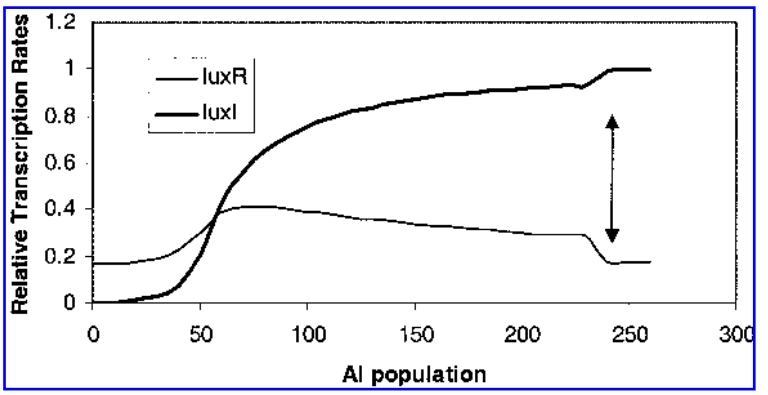


FIG. 4. Relative transcription rates as a function of AI population calculated by the deterministic model.

characterized by a threshold that must be crossed for the system to transition from a lower stable state to a higher one. The threshold in the *luxIR* system can be considered to be the concentration of $(RAI)_2$ complex required to initiate the state transition. An increase in AI population contributes to crossing this threshold in two ways. Most obviously, it increases the mean population of $(RAI)_2$ via its interactions with LuxR. A more subtle, equally important effect, is that the noise in $(RAI)_2$ also increases as the threshold is approached. These effects can be seen in Figure 6. The standard deviation of the $(RAI)_2$ population increases as it approaches the threshold. Once the threshold is reached, the production of LuxI increases rapidly with small increases in AI.

Frequency domain analysis: tools for gene circuit architectural analysis

The FD approach is equivalent to the Langevin approach (Gillespie, 2000) and shares the same limitations and caveats. However, in FD analysis the noise is dealt with using its power spectral density (PSD; the frequency distribution of the noise) and the signal processing elements are dealt with in terms of their transfer functions, $H(f) = \partial o(f)/\partial i(f)$, where o and i are output and input signals (molecular concentrations or synthesis rates) respectively, and f is the frequency in Hz (Simpson et al., 2003). For a circuit with l nodes (chemical species) with g noise sources, the noise PSDs at all nodes within the gene circuit are given by

$$[S_{out}(f)] = [|H(f)|^2][S_{source}(f)] \tag{14}$$

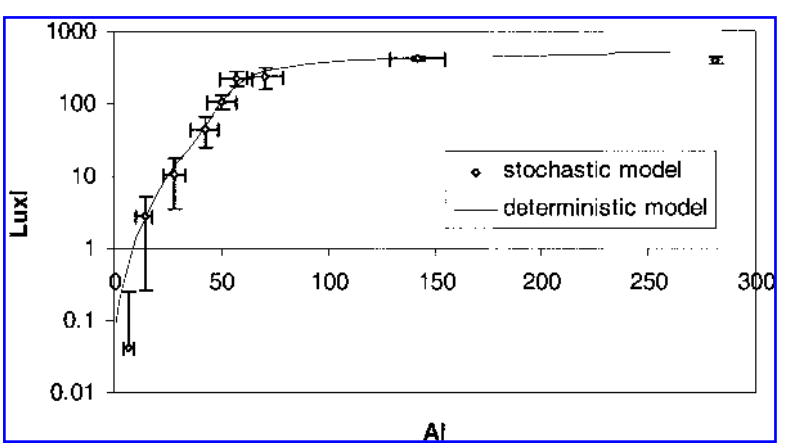


FIG. 5. Comparison of stochastic and deterministic solutions to *luxIR* quorum sensing model. The error bars for the stochastic model results represent one standard deviation.

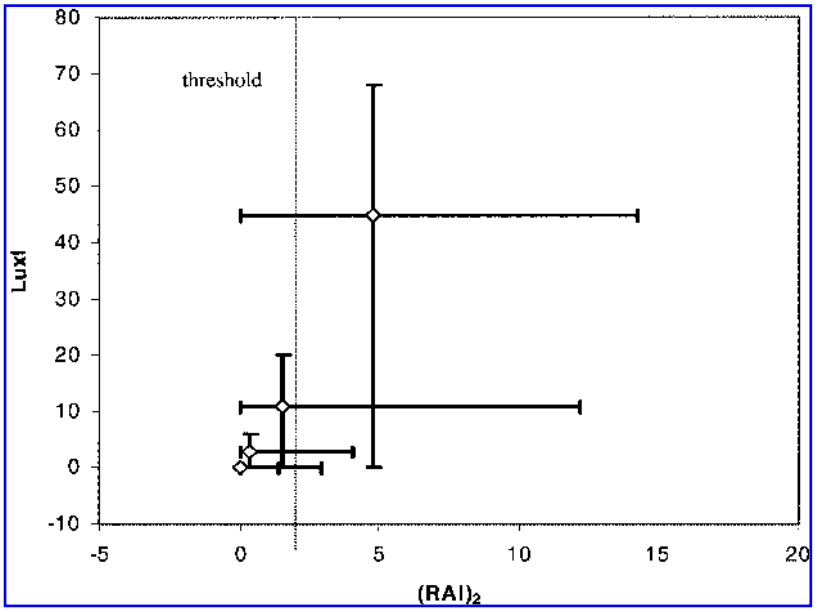


FIG. 6. Effect of noise in reaching the threshold in positive feedback circuits. Error bars represent 10 and 90 percentile values.

where $[S_{out}(f)]$ is an $l \times 1$ matrix of the PSDs of the total noise in the concentration of each chemical species; $[S_{source}(f)]$ is an $g \times 1$ matrix of the PSDs of the noise sources; and $[|H(f)|^2]$ is an $l \times g$ matrix of the power gains between node k (source) to node j (output). The elements of the noise power gain matrix are given by

$$H_{j,k}(f)H^*_{j,k}(f) = H^2_{const}\frac{\prod_{n=1}^{n_{j,k}}\left(1+\left(\frac{f}{f_{zn\_j,k}}\right)^2\right)}{\prod_{m=1}^{m_{j,k}}\left(1+\left(\frac{f}{f_{pm\_j,k}}\right)^2\right)} \tag{15}$$

where $H_{j,k}(0) = \partial o_j(0)/\partial i_k(0)$, H_{const}, is a frequency independent term (usually $H_{const} = H_{j,k}(0)$), $H^*_{j,k}(f)$ is the complex conjugate of $H_{j,k}(f)$, and $f_{z1\ldots n}$ and $f_{p1\ldots m}$ are the critical frequencies associated with the zeros and poles of the transfer functions from a source at location k to an output at location j. This assumes no poles or zeros at the origin. Poles or zeros at the origin are easily added by the placement of the proper term(s) in the numerator and/or denominator of equation (15). For later convenience, the order of the poles and zeros are arranged from lowest ($n,m = 1$) to highest ($n,m = n_{j,k}, m_{j,k}$) frequency. The poles arise from molecular decay or dilution processes, and the zeros from poles in the back propagating path in feedback circuits (Simpson et al., 2003).

The noise sources in genetic circuits are most often found in pairs at the point of molecular synthesis (or polymerization or complex formation) and the associated point of decay (or dilution or polymer/complex dissolution) due to the random timing and discrete nature of these events. For example, the single gene circuit of Figure 7 has four distinct noise sources associated with mRNA synthesis, mRNA decay, protein synthesis, and protein decay or dilution. Analysis of these noise sources shows them to be well described by shot noise with a wide band (compared to the frequency limitations of the circuit) white spectrum (Simpson et al., 2003). For a rate of a particular molecular event (i.e., synthesis, decay, dimerization, etc.) of k_T, the single-sided (positive frequency only) PSD of the noise source, $S_{source}(f)$, is (Simpson et al., 2003)

$$S_{source}(f) = 2k_T \tag{16}$$

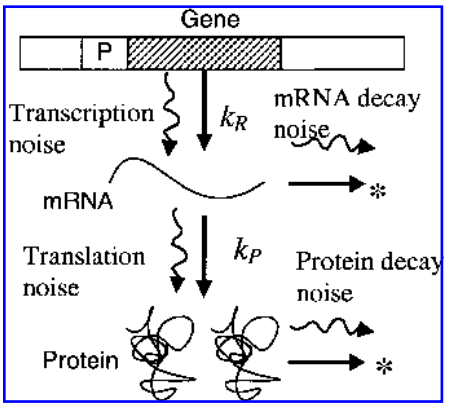


FIG. 7. Single gene circuit showing the noise sources associated with mRNA and protein synthesis and decay.

At steady state, production and decay (polymerization and dissolution, etc) are in equilibrium, so the total PSD for the pair of noise sources is

$$S_{TT\_eq}(f) = 4k_T \tag{17}$$

and the noise can be modeled as random sources in parallel with the molecular processes generating the noise (Fig. 7).

Below we describe the application of FD techniques to the architectural analysis of an important sub-circuit of the QS system. Through this architectural analysis we seek an assessment of the most important (i.e., dominant) noise sources and frequency shaping elements in the sub-circuit; the location and effect of feedback elements; the parameter values that most affect critical circuit functions; and the role of noise (i.e., as a critical functional element as in the λ switch, or as a detriment to circuit operation to be minimized). Furthermore, we look to develop a deeper understanding of the role of this sub-circuit in the overall circuit strategy coupled to feasible explanations of the selective pressures that led to the evolution of the circuit architecture.

In principle it would be possible to perform the analysis using the graphical representation in Figure 2 or even just the list of chemical reactions and associated differential equations that constitute the circuit. However, as we are dealing with signal processing concepts, we adopt an analogous electronic circuit representation, as shown in Figure 8. Sources of molecular synthesis (including polymerization or complex formation or dissolution) are modelled as either independent or dependent current sources; molecular decay and dilution are modelled as resistors; and the molecular storage elements (i.e., the interior of the cell) are modelled as capacitors (one for each molecular species). We use the circuit representation to determine the transfer functions that go into equations (14) and (15).

As shown in Figure 8, we have divided the circuit into three sections: (1) transcription/translation (noise analysis previously reported [Thattai and van Oudenaarden, 2001; Ozbudak et al., 2002; Simpson et al., 2003]); (2) binding of the autoinducer (AI) and dimerization; and (3) feedback. Reversible reaction circuits, like those of section (2), are ubiquitous elements in gene circuits and they play an important role in the function and noise behavior of the QS system. As such, we present the details of the general analysis of such circuits below.

Consider the simple reversible reaction in Figure 9. Species A is synthesized at a rate of k_p, is converted to species B at rate k_f[A], and decays at rate γ_D[A]. Species B is converted back into A at a rate k_r[B]. The transfer functions are found by circuit analysis to be approximately (we have assumed that $k_r \geq \gamma_D$):[c]

[c]Note that we have employed the concepts of the Miller effect and pole splitting, which are commonly used in the analysis of electronic circuits [21]. If $\gamma_D > k_r$, these two terms should be exchanged in the denominators of the transfer functions.

$$H_{A,k_p}(f) = \frac{1}{\gamma_D} \frac{\left(1 + i\frac{2\pi f}{k_r}\right)}{\left(1 + i\frac{2\pi f\left(1+\frac{k_f}{k_r}\right)}{\gamma_D}\right)\left(1 + i\frac{2\pi f}{k_r\left(1+\frac{k_f}{k_r}\right)}\right)} \quad \text{(a)}$$

$$H_{B,k_p}(f) = \frac{k_f}{k_r\gamma_D} \frac{1}{\left(1 + i\frac{2\pi f\left(1+\frac{k_f}{k_r}\right)}{\gamma_D}\right)\left(1 + i\frac{2\pi f}{k_r\left(1+\frac{k_f}{k_r}\right)}\right)} \quad \text{(b)}$$

$$H_{A,k_{f,r}}(f) = \frac{1}{k_r\gamma_D} \frac{i2\pi f}{\left(1 + i\frac{2\pi f\left(1+\frac{k_f}{k_r}\right)}{\gamma_D}\right)\left(1 + i\frac{2\pi f}{k_r\left(1+\frac{k_f}{k_r}\right)}\right)} \quad \text{(c)}$$

$$H_{B,k_{f,r}}(f) = \frac{1}{k_r} \frac{-\left(1 + i\frac{2\pi f}{\gamma_D}\right)}{\left(1 + i\frac{2\pi f\left(1+\frac{k_f}{k_r}\right)}{\gamma_D}\right)\left(1 + i\frac{2\pi f}{k_r\left(1+\frac{k_f}{k_r}\right)}\right)} \quad \text{(d)} \quad (18)$$

A simple examination of the transfer functions yields some important information. First, the zeros in the transfer function come from the feedback inherent in reversible reactions. Furthermore, notice that the addition of the reversible reaction to the k_p/γ_D circuit (which is equivalent to the protein output of the single gene circuit [Thattai and van Oudenaarden, 2001; Ozbudak et al., 2002; Simpson et al., 2003]) adds a pole and a zero to this output signal. So, the reversible reaction changes the transient response and the noise spectrum at A, but does not change the steady state value of A. The transfer function from k_p to B has the same two poles as that from k_p to A, but does not have the zero. So, the high frequency components of the noise associated with k_p are more heavily filtered at B than at A.

A second round of insights is gained by noticing that the low frequency pole of the transfer functions is divided by the gain of the circuit ($\Delta\langle B\rangle/\Delta\langle A\rangle = k_f/k_r$[d]) plus one, while the high frequency pole is multiplied by this factor. This phenomenon is know as pole splitting (Gray, 2001), and as the gain is increased,

[d]For $\gamma_D > k_r$ the poles move by a factor of $1 + \frac{kf}{\gamma D}$.

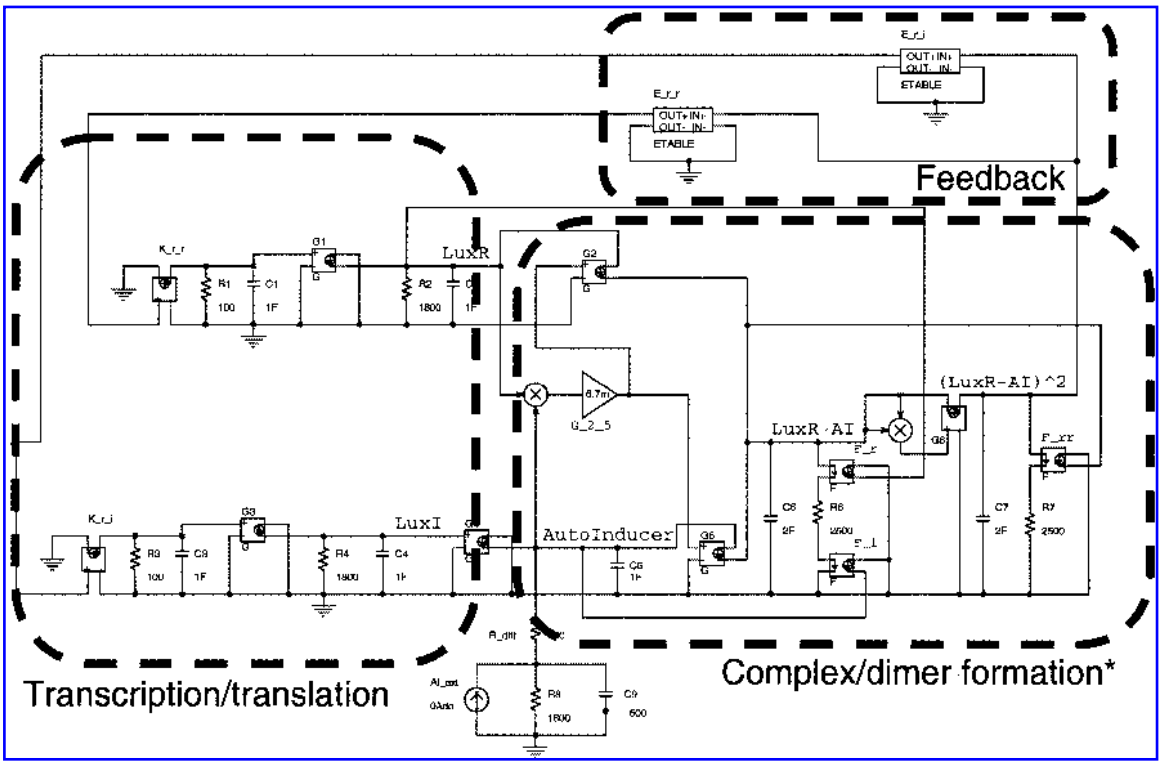


FIG. 8. Simplified analogous electronic circuit representation of the LuxI/R quorum sensing system. This circuit only includes the components of the positive feedback loop.

the low-frequency pole moves closer to the origin ($f = 0$), while the other pole moves higher in frequency. For reasons that will become obvious shortly, it is instructive to consider two special cases: (1) species A is most likely to decay before cycling through the reversible reaction ($\gamma_D >> k_f$) and the gain is low ($k_r >> k_f$); and (2) species A is likely to cycle through the reversible reaction several times before decay ($k_f >> \gamma_D$) and the gain is high ($k_f >> k_r$). For the first special case, the circuit at A has reduced to simply the k_p/γ_D circuit, and there is both very little signal and very little noise in B. The noise in A is band limited (bandwidth $= 2\pi/\gamma_D$) and mostly comes from the k_p/γ_D noise sources (Fig. 10a).

In the second case, noise in A comes predominantly from the noise associated with the reversible reaction sources and it is very broadband, while the PSD of the noise in B has two components: (1) the PSD of the noise in B associated with k_p has a large magnitude, but is very limited in bandwidth; and (2) the

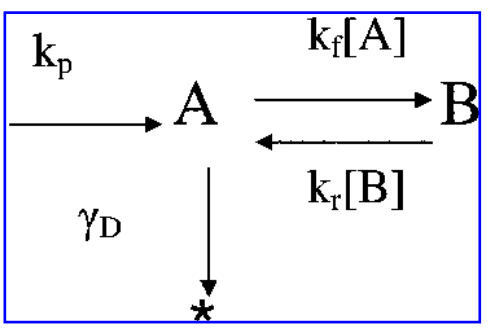


FIG. 9. A reversible reaction where Species A is synthesized at a rate of k_p, is converted to species B at rate $k_f[A]$, and decays at rate $\gamma_D[A]$. Species B is converted back into A at a rate $k_r[B]$.

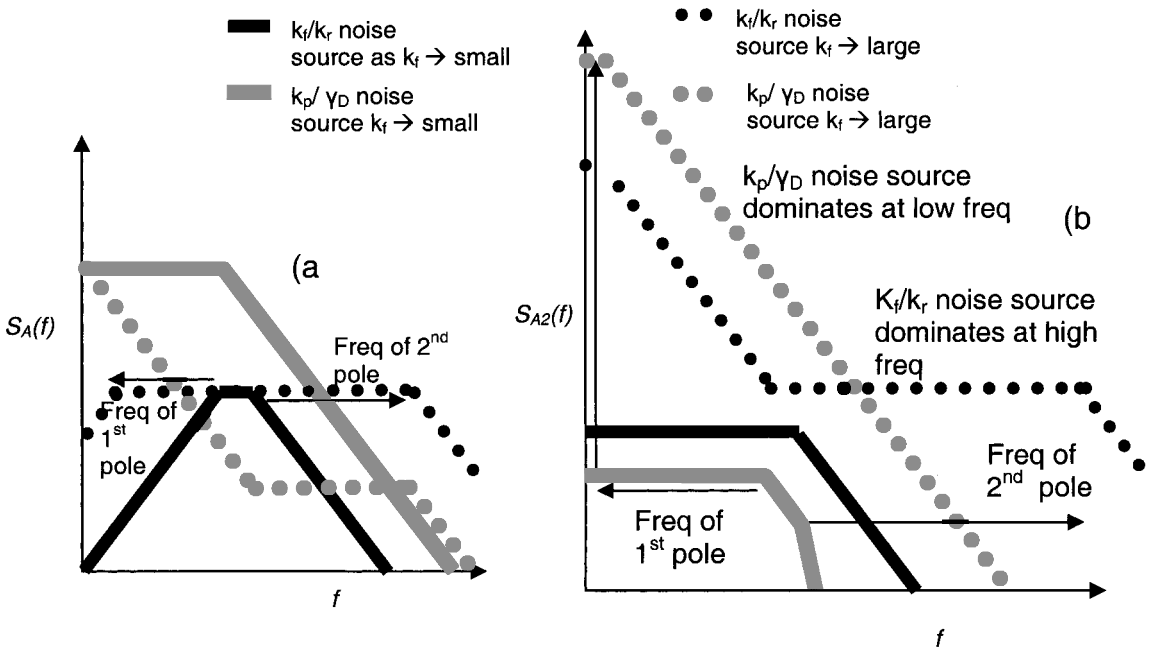


FIG. 10. The noise power spectral density for (**a**) species A and (**b**) species A dimer. The arrows show the pole splitting and the change in noise spectrum as the small signal k_f increases due to increasing [A]. The noise in species A whitens as k_f increases, while the noise in the dimer becomes more band-limited.

PSD of the noise in *B* associated with cycling through the reversible reaction is relatively small in magnitude (greatly attenuated by the low frequency pole), but has a very broad spectral distribution as the low-frequency pole and the zero cancel at frequencies above the zero (Fig. 10b).

These two special cases are of particular interest in the QS systems as the binding of the autoinducer (AI) and dimerization of this complex (Figs. 1, 2, and 8) proceed from the first special case to the second as [AI] increases. In the binding of AI, the forward reaction takes place at a rate of (Table 1) k_{f1}[AI][R]. From the point of view of R, the linearized small signal value of k_f in the above analysis would be k_{f1}[AI]. For small [AI] the first special case above applies because $k_f \rightarrow 0$. Conversely as [AI] increases, so does the small signal k_f, and the noise behavior approaches that of the second special case. The same is true for the dimerization reaction, where the small signal k_f value grows from essentially zero to larger values as the [RAI] increases. A final element in this chain is the reversible binding of the $(RAI)_2$ to the operator in the *lux* promoter (Fig. 2). These three reversible reactions are cascaded, and as described earlier, each reversible reaction operates to separate the two most closely spaced poles. At a node where two reversible reactions are connected, there are two possible paths (Fig. 11). At a given node (C in Fig. 11) if the most likely event is to proceed into the forward path of the next reversible reaction in the chain, then the noise at this node (C) "whitens," and the band-limited noise appears at the output of the next reversible reaction (D) in the chain. For the reactions described here, where the linearized (i.e., small signal) k_f terms are a function of molecular concentrations, this situation evolves as [AI] and then in turn [RAI] and $[(RAI)_2]$ increase, giving rise to a dynamic noise behavior as demonstrated through the simulation results (using ESS) shown in Figure 12.

DISCUSSION

Figure 6 clearly demonstrates that the role of noise in the QS circuit is to include the entire population in the positive feedback function. As the [AI] increases toward the threshold value (i.e., the concentration where the QS circuit switches into the induced state), the noise in $(RAI)_2$ increases. At an individual cell level, this would only result in some cells reaching the threshold early, while others reach it later. How-

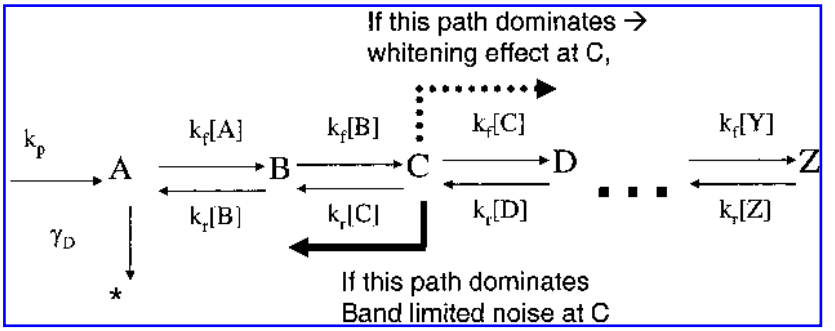


FIG. 11. The noise behavior in cascaded reversible reaction circuits depends on the most likely fate of molecules at a given node.

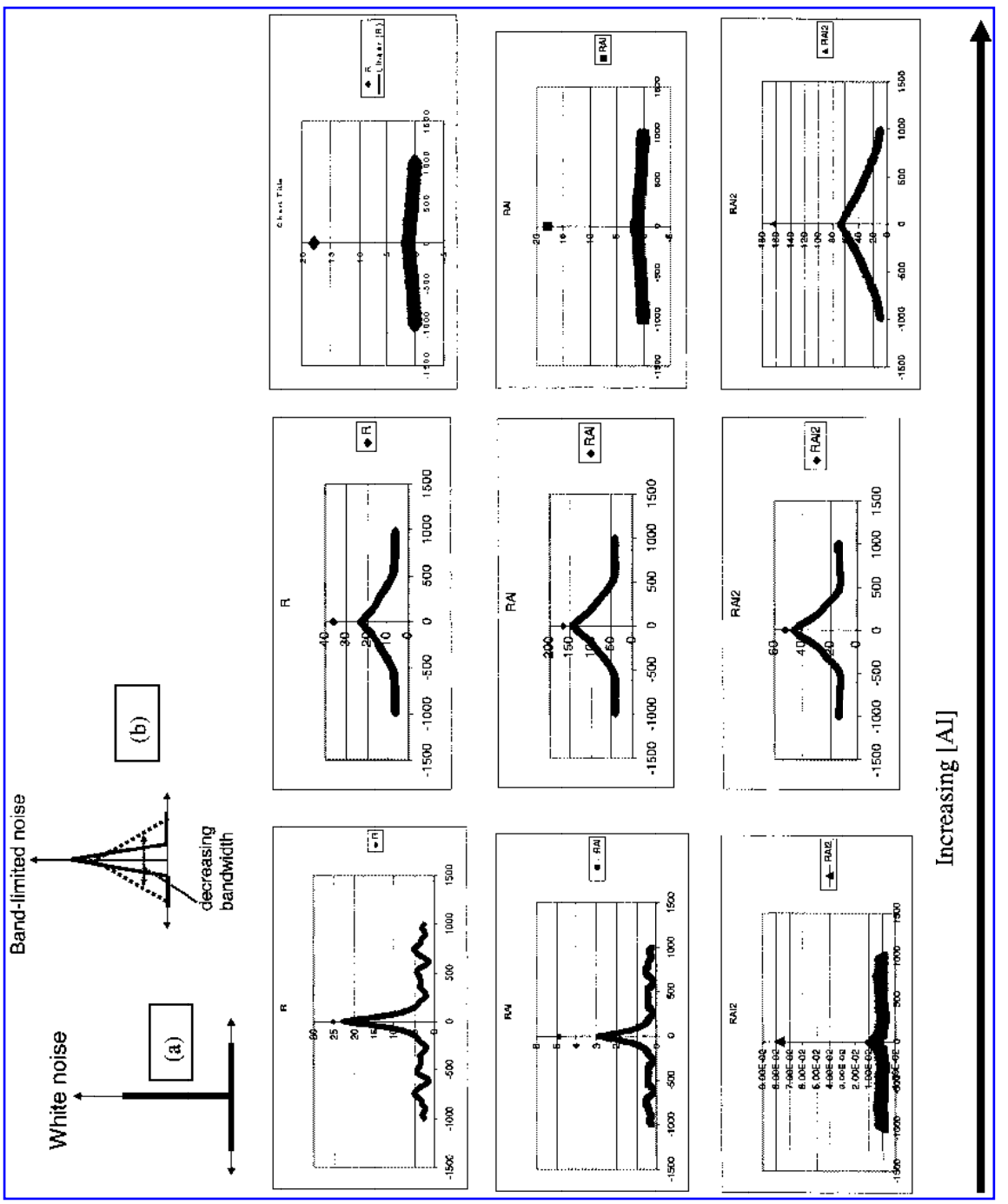
ever, since these early switching cells reinforce their behavior throughout the population by production of additional AI (and there is no counter balancing penalty for those cells with low $[(RAI)_2]$), the noise is a functional element in the QS positive feedback system. On the surface this does not seem to be as striking an example of noise functionality as that seen in the λ switch, as the noise would seem to do nothing that increasing the strength of the positive feedback would not provide. However, we hypothesize that there is a benefit obtained by distributing this additional positive feedback, in the form of noise, throughout the population instead of within the individual cells. We outline this hypothesis below.

Consider a case where the AI is actively removed from the population, either by degradation (perhaps by a gram-positive competitor) or by some other means. In this case, [AI] under represents the magnitude of the population, and the QS circuit does not switch as it would if AI were removed only by passive diffusion. However, as the population increases, more cells will switch due to noise. At relatively low population levels, cells that have switched are likely to be well removed from each other, and they will not remain in the induced state as the background [AI] is too low. However, at higher population levels the possibility of small groups of locally clustered cells all switching into the induced state is increased. If this group is large enough, the additional local production of AI may allow these cells to remain in the induced state. Over time more of these local groups are formed, perhaps eventually leading to self-sustained induction of the entire population. Thus the noise, acting as positive feedback distributed throughout the population, may act to ensure the operation of the QS circuit even if other organisms or processes are actively interfering with the signaling system.

The rationale for the use of frequency domain analysis is that the power spectral densities of the noise terms are critically important in gene circuit noise behavior. It was previously demonstrated that autoregulated gene circuits may optimize noise performance by shifting noise into frequency regimes where it has a smaller effect on total circuit performance (Simpson et al., 2003). The analysis of the reversible reactions of the QS system reported here show a significant *and* dynamic remodeling of the noise spectra. How this may ultimately impact the total operation of the QS circuit is still under investigation. However, the analysis of this gene circuit will not be complete without a careful consideration of this complex noise behavior and perhaps other implications of the dynamic modulation of the circuit frequency response. Furthermore, as similar reversible reactions are ubiquitous in gene circuit architectures, a detailed understanding of these mechanisms as circuit components is fundamental to the development of gene circuit analysis, modeling, simulation, and design disciplines.

Finally, a common caveat to the use of the Langevin equations, on which FD analysis is based, is that

FIG. 12. Autocorrelation (AC) functions of noise in the reversible reactions of the QS system simulated using ESS. Example AC functions for purely white noise (**a**) and for band-limited noise (**b**). As the bandwidth decreases, the width of the AC function increases. (**c**)–(**e**), (**f**)–(**h**), and (**i**)–(**k**) show the evolution of the AC function of the noise in R, RAI, and $(RAI)_2$, respectively as [AI] increases. Notice the growth of white noise (impulse at the origin) for R and RAI as forward path of the subsequent reaction begins to dominate.


White noise
(a)
Band-limited noise
(b)
decreasing bandwidth
R
RAI
RAI2
Increasing [AI]

they lose validity as the number of molecules of a given species becomes small. A comparison of simulated (using ESS) and calculated (using equations (18[a]–[d]) PSDs for a dimerization reaction are shown in Figure 13. Note that the excellent agreement between calculated and simulated results was obtained for an average population of only 10 molecules in the monomer species. This suggests that at least the architectural insights provided by FD analysis hold validity even for small molecular concentrations.

METHODS

Exact stochastic simulator

The simple, deterministic approach to the description of chemical reactions uses the reaction rate and stoichiometry of each of the constituent reactions to derive a system of ordinary differential equations. When modeling systems with large populations of each chemical species, this approach can describe exactly the evolution of the system over time. For systems containing smaller populations of molecules, this approach no longer provides adequate fidelity. In this regime, a stochastic approach is employed instead.

The chemical master equation (CME) can be used for stochastic simulations of chemical reactions and provides an exact representation of any gas-phase chemical system that is well-stirred and in thermal equilibrium (McQuarri, 1967; Gillespie, 1977, 1992). The CME describes the probability distribution over time of each species using a set of discrete, time-dependent difference and differential equations. At any given time, the state of the chemical system is completely described by the current population of each species. Hence, the CME is a Markov process. The chemical reactions are assumed to occur according to a Poisson process, with the rate of each reaction (also known as its propensity) determined by a stochastic rate constant scaled by the number of available reacting molecules as shown below.

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{X} + \mathrm{Y} \qquad r_1 = k_1[A][B] \tag{19}$$

$$2\mathrm{C} + 3\mathrm{D} \rightarrow \mathrm{X} + 2\mathrm{Z} \qquad r_2 = k_2[C]([C] - 1)[D]([D] - 1)([D] - 2) \tag{20}$$

In the first reaction, propensity is the product of the stochastic rate constant and the populations of each of the reacting species. Similarly, in the second reaction the propensity is the product of the stochastic rate

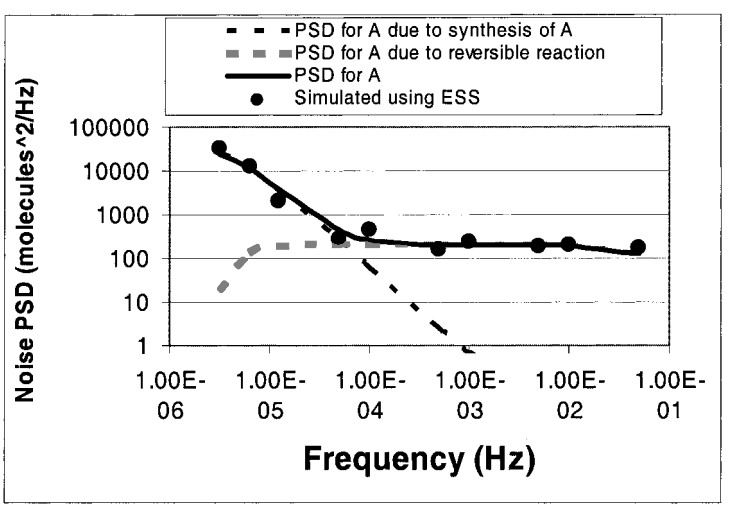


FIG. 13. The calculated and simulated (using ESS) noise power spectral density in species A in a dimerization reaction like that shown in Figure 9. For dimerization, species B becomes A_2, and the forward reaction rate becomes $k_f[\mathrm{A}]^2$. In the above, $k_f = 0.02$, $k_r = 0.01$, $\gamma_D = 0.001$, and $k_p = 0.01$. In the analysis, the contributions of noise from molecular synthesis and the reversible reaction are shown separately and as a sum. Notice the excellent agreement between the simulated and calculated PSDs even though the average [A] is 10 molecules.

constant and the populations of each of the reacting species. Note that for reactants with stoichiometries greater than one the reaction consumes more that one reactant molecule; hence, we must account for the reduction in the populations of available reactants. Each state of the system consists of the populations of each of the chemical species; given this state the propensity of each reaction can be easily determined. One can then construct a Markov chain with each state determined by the population counts, the transitions determined by the set of reactions, and using the propensity values as the reaction rates.

As the potential type and number of chemical species increases, the state space of the Markov chain grows very quickly, making it quite difficult to solve analytically or numerically. Although researchers have demonstrated that smaller systems can be solved analytically or numerically (Elf, 2003), such approaches are often not practical and the CME is simulated using Monte Carlo techniques.

Gillespie introduced two techniques for the exact stochastic simulation of the CME (Gillespie, 1976), the Direct Method and the First Reaction Method. The Direct Method exploits the fact that the reactions follow a Poisson process by adding all the reaction propensity values to find the overall reaction rate of the system $a = \sum_i a_i$. This overall propensity is the parameter for an exponential distribution used to compute the time of the next reaction. To identify which reaction occurred, a uniform random variable is sampled, with each reaction having likelihood a_i/a of being selected. After finding the next reaction, the populations of each of the species populations are updated in accordance with the reaction stoichiometry and the propensity values are calculated for the new state. This process is repeated until the end of the simulation is reached.

The First Reaction Method (Gillespie, 1976) computes a set of exponentially distributed random variables using the propensity values for each of the potential chemical reactions as the rate constant for the distribution. Each of these sampled values represents the length of time until the next occurrence of its respective reaction. One then selects the first reaction that will occur, and updates the system state as with the Direct Method. Although one may wish to then apply the second reaction when the first reaction is completed, it would not be mathematically correct to do so because the first reaction will alter the propensity values of the system reactions. As with the Direct Method, after each reaction the propensity values are recalculated before proceeding to the next reaction.

Gibson and Bruck (Gibson and Bruck, 2000) developed an optimization of the Gillespie method by noting that the propensity values for reactions are only changed if one of the reactant populations is changed. Moreover, the reaction time samples drawn using Gillespie's First Reaction Method remain accurate as long as the reaction propensity values do not change. Hence, Gibson and Bruck employ dependency graphs and improved data structures in their Next Reaction Method to significantly improve performance over the First Reaction Method.

Our Exact Stochastic Simulator (ESS) reads SBML level 2 models and simulates biological models using the Gillespie algorithm with the Gibson and Bruck enhancements. ESS allows reactions to specify rates with constants or Hill Repression Kinetics. However, while the use of Hill Kinetics may significantly speed computation, the noise implications of this representation have not been fully explored. Accordingly, Hill Kinetics within ESS should be used with due caution.

ACKNOWLEDGMENTS

This material is based upon work supported by the Defense Advanced Research Projects Agency Bio-Computation Program and the National Science Foundation under grant no. 0130843. We are indebted to M.J. Doktycz and M.J. Roberts for several fruitful discussions related to the topics presented here. This work was partially performed at the Oak Ridge National Laboratory, managed by UT-Battelle, LLC for the U.S. DOE under contract no. DE-AC05-00OR22725.

REFERENCES

ARKIN, A., ROSS, J., and McADAMS, H.H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. Genetics **149,** 1633–1648.

DUNLAP, P.V., and GREENBERG, E.P. (1988). Control of *Vibrio-Fischeri* Lux gene-transcription by a cyclic–AMP receptor protein Luxr protein regulatory circuit. Journal of Bacteriology **170,** 4040–4046.

ELF, J., LOTSTEDT, P., and SJOBERG, P. (2003). Problems of high dimension in molecular biology. Presented at the 19th GAMM Seminar, Leipzig.

ENGEBRECHT, J., NEALSON, K., and SILVERMAN, M. (1983). Bacterial bioluminescence—isolation and genetic analysis of functions from *Vibrio fischeri*. Cell **32,** 773–781.

GIBSON, M.A., and BRUCK, J. (2000). Efficient exact stochastic simulation of chemical systems with many species and many channels. Journal of Physical Chemistry A **104,** 1876–1889.

GILLESPIE, D.T. (1976). General method for numerically simulating stochastic time evolution of coupled chemical reactions. Journal of Computational Physics **22,** 403–434.

GILLESPIE, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. Journal of Physical Chemistry **81,** 2340–2361.

GILLESPIE, D.T. (1992). A rigorous derivation of the chemical master equation. Physica A **188,** 404–425.

GILLESPIE, D.T. (2000). The chemical Langevin equation. Journal of Chemical Physics **113,** 297–306.

GRAY, P.R., HURST, P.J., LEWIS, S.H. et al. (2001). *Analysis and Design of Analog Integrated Circuits* (Wiley, New York).

GREENBERG, E.P. (2000). Acyl-homoserine lactone quorum sensing in bacteria. Journal of Microbiology **38,** 117–121.

KAPLAN, H.B., and GREENBERG, E.P. (1985). Diffusion of autoinducer is involved in regulation of the *Vibrio fischeri* luminescence system. Journal of Bacteriology **163,** 1210–1214.

KENNELL, D., and RIEZMAN, H. (1977). Transcription and translation initiation frequencies of *Escherichia coli* Lac operon. Journal of Molecular Biology **114,** 1–21.

McADAMS, H.H., and SHAPIRO, L. (1995). Circuit simulation of genetic networks. Science **269,** 650–656.

MCQUARRI, D.A. (1967). Stochastic approach to chemical kinetics. Journal of Applied Probability **4,** 413.

OZBUDAK, E.M., THATTAI, M., KURTSER, I., et al. (2002). Regulation of noise in the expression of a single gene. Nature Genetics **31,** 69–73.

RAO, C.V., WOLF, D.M., and ARKIN, A.P. (2002). Control, exploitation and tolerance of intracellular noise. Nature **420,** 231–237.

SCHAEFER, A.L., VAL, D.L., HANZELKA, B.L. et al. (1996). Generation of cell-to-cell signals in quorum sensing: acyl homoserine lactone synthase activity of a purified *Vibrio fescheri* LuxI protein. Proceedings of the National Academy of Sciences USA **93,** 9505–9509.

SHADEL, G.S., and BALDWIN, T.O. (1991). The *Vibrio fischeri* Luxr protein is capable of bidirectional stimulation of transcription and both positive and negative regulation of the Luxr gene. Journal of Bacteriology **173,** 568–574.

SHADEL, G.S., and BALDWIN, T.O. (1992a). Positive autoregulation of the *Vibrio fischeri* Luxr gene—Luxr and autoinducer activate camp-catabolite gene activator protein complex-independent and complex-dependent Luxr transcription. Journal of Biological Chemistry **267,** 7696–7702.

SHADEL, G.S., and BALDWIN, T.O. (1992b). Identification of a distantly located regulatory element in the Luxd gene required for negative autoregulation of the *Vibrio fischeri* Luxr gene. Journal of Biological Chemistry **267,** 7690–7695.

SIMPSON, M.L., COX, C.D., and SAYLER, G.S. (2003). Frequency domain analysis of noise in autoregulated gene circuits. Proceedings of the National Academy of Sciences USA **100,** 4551–4556.

STEPHANOPOULOS, G.N., ARISTIDOU, A.D., and NIELSEN, J. (1998). *Metabolic Engineering* (Academic Press, San Diego).

STEVENS, A.M., and GREENBERG, E.P. (1997). Quorum sensing in *Vibrio fischeri*: essential elements for activation of the luminescence genes. Journal of Bacteriology **179,** 557–562.

THATTAI, M., and VAN OUDENAARDEN, A. (2001). Intrinsic noise in gene regulatory networks. Proceedings of the National Academy of Sciences USA **98,** 8614–8619.

WITHERS, H., SWIFT, S., and WILLIAMS, P. (2001). Quorum sensing as an integral component of gene regulatory networks in gram-negative bacteria. Current Opinion in Microbiology **4,** 186–193.

Address reprint requests to:
Dr. Michael L. Simpson
Oak Ridge National Laboratory
P.O. Box 2008, M.S. 6006
Oak Ridge, TN 37831-6006

E-mail: simpsonML1@ornl.gov