Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

Living with Noisy Genes: How Cells Function Reliably with Inherent Variability in Gene Expression

Narendra Maheshri[1] and Erin K. O'Shea[2]

[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; email: narendra@mit.edu

[2]Howard Hughes Medical Institute, FAS Center for Systems Biology, Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138; email: erin_oshea@harvard.edu

Annu. Rev. Biophys. Biomol. Struct. 2007. 36:413–34

The *Annual Review of Biophysics and Biomolecular Structure* is online at biophys.annualreviews.org

This article's doi:
10.1146/annurev.biophys.36.040306.132705

Copyright © 2007 by Annual Reviews.
All rights reserved

1056-8700/07/0609-0413$20.00

Key Words

intrinsic noise, extrinsic noise, gene networks

Abstract

Within a population of genetically identical cells there can be significant variation, or noise, in gene expression. Yet even with this inherent variability, cells function reliably. This review focuses on our understanding of noise at the level of both single genes and genetic regulatory networks, emphasizing comparisons between theoretical models and experimental results whenever possible. To highlight the importance of noise, we particularly emphasize examples in which a stochastic description of gene expression leads to a qualitatively different outcome than a deterministic one.

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

Contents

INTRODUCTION 414
BACKGROUND 414
Stochastic Chemical Kinetics as a Random Walk 414
Experimental Measurement and Classification of Noise 417
GENE EXPRESSION: MOLECULES, STATISTICS, AND MECHANISMS 418
In Vivo Single-Molecule Measurements of Gene Expression and the Statistics of Intrinsic Noise 418
Steady-State Protein Distributions of Single Genes and the Nature of Extrinsic Noise 421
Single-Cell Tracking of Gene Expression and the Dynamics of Noise 421
A Genome-Wide View of Noise in Single Genes 423
PUTTING IT ALL TOGETHER: SIMPLE NETWORKS 424
Linear Networks and the Importance of Correlated Noise Sources 424
Negative Feedback 426
Positive Feedback 427
Noise and Oscillations 428
OUTLOOK 430

INTRODUCTION

The basic unit of life, a single cell, performs thousands of chemical transformations simultaneously. Because small numbers of each molecular species are involved, the outcome of intracellular reactions is subject to thermal noise. In addition, these reactions occur in a crowded, dynamic microenvironment that plays a crucial role in determining reaction rate. Hence, the already difficult problem of understanding the design and structure of the thousands of chemical transformations that occur at the cellular level is compounded by the fact that many of these transformations are not completely predictable.

Gene expression noise: related to the variability in the resulting protein, a dimensionless quantity defined as standard deviation of protein number divided by the mean protein number

Gene expression is a set of chemical transformations that is fundamental to all life. Since the demonstration of a functional role for stochastic gene expression in phage λ (2) and a subsequent review by the same authors (39), there has been an explosion of studies focused on investigating the origins and consequences of noise in gene expression. Investigators have made progress in this area by first establishing the molecular mechanisms of noise generation at the single gene level, then building on this knowledge to test and predict its effects on larger regulatory networks. We adopt a similar approach in this review, beginning with recent single-molecule studies that directly measure expression noise and confirm stochastic models, then working toward an understanding of how this noise shapes, and is shaped by, genetic regulatory networks. In light of this topic's multidisciplinary nature, we emphasize connections between theoretical and experimental work, paying special attention to examples in which a stochastic description yields a qualitatively different outcome from a deterministic one.

BACKGROUND

Stochastic Chemical Kinetics as a Random Walk

To design and interpret experiments analyzing noise in gene expression, it is essential to have some background in the mathematics of stochastic processes. This brief introductory section is designed for readers new to this area and provides references for further reading.

Random walks. Imagine a pedestrian executing a one-dimensional random walk along a line where he or she randomly steps forward or backward. We can discretize the line and label each position as a unique state. The set of all possible states is then the state space of our walk, here an infinite one-dimensional

state space. Because the probability of moving forward or backward depends only on the current state of the pedestrian, this random walk is said to be memoryless, or a Markov process.

We want to analyze the probability that our pedestrian is in a particular state at some time, given that he or she starts in some initial state:

$$P(n, t+\tau) = \sum_i W(n, t+\tau | i, t) P(i, t). \quad 1.$$

The probability, $P(n, t + \tau)$, that the pedestrian is in state n at time $t + \tau$ depends on where he or she was previously, at time t (assuming each step takes time τ), multiplied by the transition probability (W) of walking to state n from state i during the period lasting from t to $t + \tau$. Obviously, in his or her one-dimensional random walk, the pedestrian can reach state n in one step only from state $n - 1$ or $n + 1$, and the transition probabilities are constant functions of time. Incorporating this fact, and now taking the continuous time limit, we arrive at a master equation for this process:

$$\begin{aligned}\frac{dP}{dt} &= W'(n|n-1)P(n-1) \\ &\quad + W'(n|n+1)P(n+1) \\ &\quad - [W'(n|n-1) + W'(n|n+1)]P(n).\end{aligned} \quad 2.$$

The master equation in Equation 2 describes how the discrete probability distribution of the pedestrian's position evolves continuously in time. The transition probabilities W are now replaced by transition rates W'. In general, master equations are difficult to solve. If the spacing between steps becomes very small, the master equation can be simplified to a more tractable diffusion equation that describes the movement of our pedestrian in a continuous state space with a diffusion coefficient, D, proportional to the magnitude of the transition rates:

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial n^2}. \quad 3.$$

Equation 3 is the Fokker-Planck equation for this system and describes the time evolution of the probability distribution of our pedestrian's position. If the state space is unbounded, the solution to this equation is a normal distribution centered at the pedestrian's initial position, whose variance increases with time.

Application to chemical kinetics. We can think of a chemical reaction as a random walk in the state space of non-negative integers. To see this, let us consider the following master equation that describes a birth-death process of a single molecule:

$$\begin{aligned}\frac{dP(X)}{dt} &= kP(X-1) + \gamma(X+1)P(X+1) \\ &\quad - [k + \gamma X]P(X).\end{aligned} \quad 4.$$

Each state corresponds to the number of molecules. The transition probabilities, k and γX, are the mesoscopic production and degradation rates of X, respectively (see Reference 21 for the precise relationship between mesoscopic and macroscopic reaction rates). As the average number of molecules becomes large (35) or, alternatively, when the system volume becomes large at constant concentration of x—van Kampen's Ω expansion (66)—we can take an appropriate continuous limit of Equation 4 and recover the following Fokker-Planck equation:

$$\begin{aligned}\frac{\partial P(x,t)}{\partial t} &= \frac{\partial}{\partial x}((k' - \gamma x)P) \\ &\quad + \frac{\partial^2}{\partial x^2}\left(\frac{(k' + \gamma x)}{2}P\right); \\ k' &= k/V, \quad x = X/V.\end{aligned} \quad 5.$$

Equation 5 is a diffusion equation with a convective term. The convective term describes the deterministic reaction, driving the probability distribution to the macroscopic steady-state value, whereas the diffusive term describes the noise that causes the distribution to spread around this value. Equation 5 represents the well-known Ornstein-Uhlenbeck process, which arises repeatedly in applications. The steady-state solution can be solved exactly and is a normal distribution, centered on the macroscopic steady state with

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

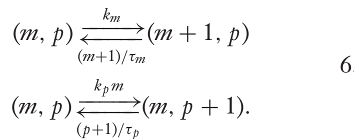
Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

CV: coefficient of variation

a variance proportional to the diffusive term (50). The noise in x, η_x, is a dimensionless quantity defined as the standard deviation of the distribution of x divided by the mean value of x, the coefficient of variation (CV).

Often the transition probabilities (the mesoscopic reaction rates) in the chemical master equation are nonlinear functions of X. If we restrict ourselves to finding the steady-state probability distribution, we can linearize the Fokker-Planck equation around the macroscopic steady state, recovering Equation 5, the linearized noise approximation. The multidimensional case (treated in Reference 21), in which X represents a vector of different molecular species, has a multivariate Gaussian as the steady-state solution.

Gene expression is a sequential process: Fluctuations at the promoter can be transmitted to mRNA levels and ultimately protein levels. Below we consider the birth-death processes of transcription and translation, in which the two-dimensional state space includes both mRNA (m) and protein (p) numbers. The transcription rate, k_m, is constant; the translation rate ($k_p m$) depends on mRNA number; and the degradation of both mRNA and protein is first order with rates τ_m^{-1} and τ_p^{-1}:



$$(m, p) \underset{(m+1)/\tau_m}{\overset{k_m}{\rightleftarrows}} (m+1, p)$$
$$(m, p) \underset{(p+1)/\tau_p}{\overset{k_p m}{\rightleftarrows}} (m, p+1). \qquad 6.$$

In a sequential birth-death process (in which any event depends on previous events, but not subsequent events), one can factorize the solution to the linearized noise approximation to yield a useful representation of how the protein noise, η_p (see above), depends on various reaction parameters (44):

$$\eta_p^2 \equiv \frac{\sigma_p^2}{\langle p \rangle^2} = \frac{1}{\langle p \rangle H_{pp}} + \eta_m^2 \frac{H_{pm}^2}{H_{pp}^2} \frac{H_{pp}/\tau_p}{H_{mm}/\tau_m + H_{pp}/\tau_p}. \qquad 7.$$

The H_{xy} terms measure how the steady-state balance of births and deaths of x depends on y. They are closely connected to the order of the chemical reactions. Because both death processes are first order, H_{mm} and H_{pp} are 1. Similarly, because the protein birth process has a first-order dependence on the mRNA number, $H_{pm} = 1$. This leads to

$$\eta_p^2 = \frac{1}{\langle p \rangle} + \frac{1}{\langle m \rangle} \frac{\tau_m}{\tau_m + \tau_p} = \frac{1}{\langle p \rangle}\left(1 + k_p \tau_p \frac{\tau_m}{\tau_m + \tau_p}\right). \qquad 8.$$

The extent to which protein noise is dependent on mRNA noise depends on the relative lifetimes of mRNA and protein. Because mRNA lifetimes are small compared with protein lifetimes ($\tau_m \ll \tau_p$), Equation 8 further simplifies to

$$\eta_p^2 = \frac{1}{\langle p \rangle}(1 + k_p \tau_m). \qquad 9.$$

The $k_p \tau_m$ term represents translational bursts of proteins produced during the short life of each mRNA. In addition, Equation 9 is the basis for understanding why the protein noise (η_p) due to stochastic chemical events decreases by the square root of 2 when gene expression and protein abundance double (64) (see below for more detail).

The Langevin approach is an alternative approach to deriving a Fokker-Planck equation from a master equation. The Langevin equation for production and first-order degradation of one species is identical to the deterministic differential equation except for the addition of a noise term, ξ:

$$\frac{dx}{dt} = k - \gamma x + \xi. \qquad 10.$$

If ξ is Gaussian-distributed white noise, Equation 10 also represents an Ornstein-Uhlenbeck process and is identical to the Fokker-Planck equation (Equation 5). The Langevin approach is desirable when one wishes to add noise heuristically into the deterministic equations based on physical reasoning (50). For a review on numerical

techniques developed to solve the equations presented here, see Reference 65.

Experimental Measurement and Classification of Noise

A cell can be conceptually described by a set of state values (e.g., size, shape, and protein number). In gene expression studies, the relevant state variable is usually protein (or mRNA) number and is typically measured at the single-cell level using a fluorescent protein. The distribution of this state variable, and hence the noise, is then ascertained using either a flow cytometer or fluorescence microscopy (**Figure 1*a***).

Any large dynamical system can be artificially divided into modules. The noise in the

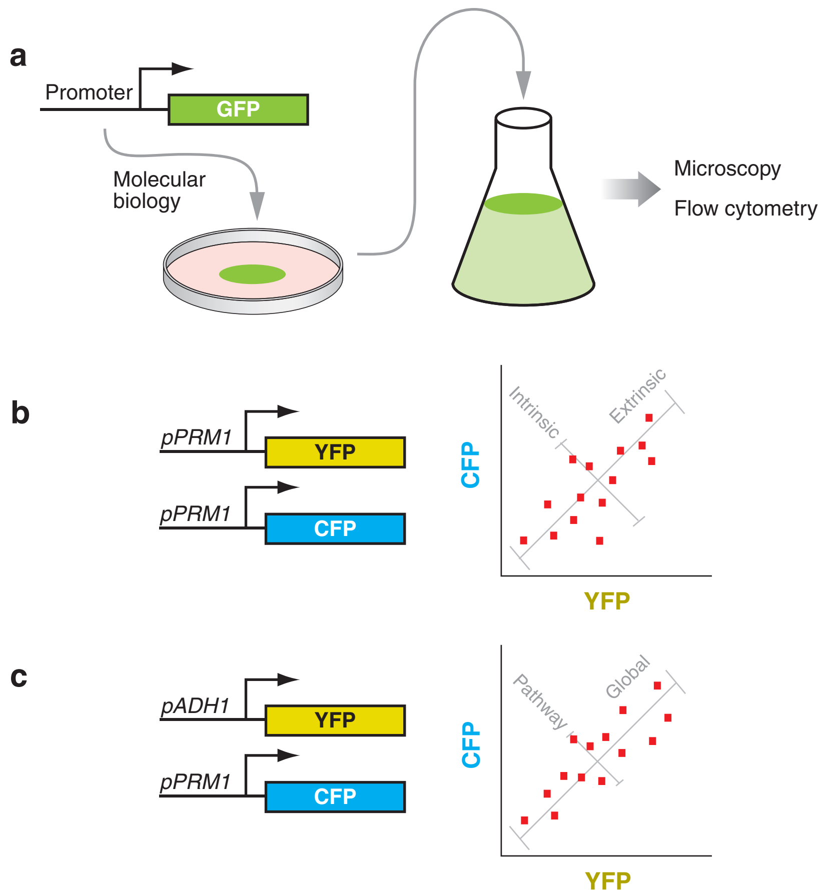


Figure 1

Measuring noise. (*a*) The gene encoding a fluorescent protein is placed under the control of a promoter, allowing gene expression to be monitored. GFP, green fluorescent protein. (*b*) Extrinsic and intrinsic noise contributions are measured by looking at the correlation in the expression of two identical copies of the pheromone-responsive *PRM1* gene promoter, within a cell and between different cells. CFP, cyan fluorescent protein; YFP, yellow fluorescent protein. (*c*) Both pathway-specific and global expression noise are determined by comparing expression of the pheromone pathway-specific promoter (*PRM1*) with that of an unrelated, constitutively expressed promoter (*ACT1*). Correlated changes in expression result from global fluctuations in expression capacity, whereas uncorrelated fluctuations result from specific fluctuations in the pheromone pathway.

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

Intrinsic noise: an experimentally defined component of the total noise that consists of differences in the expression of identical genes in the same intracellular environment

Extrinsic noise: an experimentally defined component of the total noise that consists of differences in identical genes due to differences between cellular environments

Pathway-specific noise: an experimentally defined component of the extrinsic noise that consists of extrinsic variation resulting from cell-to-cell differences in the signaling pathway(s) dictating a gene's expression

output of each module can be partitioned into two sources: (*a*) the intrinsic noise, originating from fluctuations internal to the module, and (*b*) the extrinsic noise, originating from external fluctuations that impinge on the module. Gene expression is one module of the cellular system and its boundary is operationally drawn around the processes of promoter remodeling, transcription, and translation of a given gene. Random fluctuations in these chemical transformations contribute to the intrinsic noise in the expression of that gene. Variation in external factors such as the number of ribosomes or transcription factors contributes to the extrinsic noise in gene expression (62). Each noise contribution can be measured using two copies of the same promoter expressing two distinct fluorescent proteins in a single cell (24, 49). If both reporters are independent and equivalent in their expression, then the difference between the number of each distinct fluorescent protein within a single cell, averaged over many cells, is proportional to the intrinsic noise. The extrinsic noise is proportional to cell-to-cell differences in total fluorescent protein expression (**Figure 1*b***). One can also distinguish between these two noise sources using multiple copies of a promoter fused to a single fluorescent protein (7, 69). If each copy is independent and equivalent, the degree of correlation between the copies corresponds to the relative contribution of extrinsic noise (69). We note that the terms intrinsic noise and extrinsic noise usually refer to steady-state fluctuations.

Gene-specific biochemical events lead to intrinsic noise, but many factors could contribute to extrinsic noise. For example, fluctuations in both the activity of an upstream transcription factor or in ribosome number are extrinsic to the promoter and lead to correlated fluctuations in identical downstream promoters. To distinguish the fluctuations of the former type from the latter, extrinsic noise can be further subdivided into pathway-specific noise and global expression noise (19, 34). The degree of correlation between promoters controlled by the same pathway gives pathway-specific noise, whereas correlation between promoters from independent pathways gives the relative contribution of global noise (**Figure 1*c***).

GENE EXPRESSION: MOLECULES, STATISTICS, AND MECHANISMS

In Vivo Single-Molecule Measurements of Gene Expression and the Statistics of Intrinsic Noise

Researchers have used stochastic models of gene expression to infer promoter, mRNA, and protein dynamics based on static snapshots of protein distribution (see Reference 45 for a review). Recent in vivo single-molecule experiments now provide information on mRNA and protein dynamics to support these models. In these experiments, investigators measured absolute mRNA numbers and relative protein levels simultaneously in single *Escherichia coli* cells (28). They measured RNA abundance by introducing an RNA sequence that allowed the mRNAs to bind a fluorescent RNA binding protein and measured relative protein levels with a fluorescent protein. These experiments revealed that transcription from a *lac/ara* hybrid promoter occurred in geometrically distributed bursts of mRNA. Such statistics are consistent with a process in which the promoter switches randomly between an ON and OFF state. The *lac/ara* promoter produced average bursts of ~4 mRNA, and it was inferred that additional induction increased burst frequency. Furthermore, the statistics of mRNA partitioning after cell division were close to a binomial model in which each mRNA had an equal chance of ending up in either daughter cell. Steady-state protein levels were proportional to mRNA levels, demonstrating the validity of the common assumption that protein production is proportional to mRNA levels (28).

A set of complementary studies focused on measuring protein numbers in single cells

(15, 71). Whereas actual protein numbers had been inferred previously (51; see below), this set of studies used two complementary methods to measure the statistics of protein production from a repressed *lac* promoter in *E. coli*. In the first study, single cells were trapped in microfluidic chambers containing a fluorogenic substrate that became fluorescent only on β-gal-mediated hydrolysis. This assay allowed single-molecule detection of β-gal tetramers by monitoring the hydrolysis rate. Protein expression occurred in exponentially distributed bursts, with an exponential distribution of protein molecules produced per burst. This is similar to the geometric distribution of mRNA bursts (the exponential distribution is the continuous time analog of the discrete geometric distribution). Moreover, steady-state protein numbers were gamma-distributed, as expected if production is exponentially distributed and protein numbers are halved every cell division. The results correspond to a burst frequency of ∼0.1–0.2 per cell cycle and a burst size of ∼7 enzymes (15). In the second study, researchers fused the fluorescent protein Venus to Tsr, a bacterial membrane protein. The same *lac* promoter controlled expression of the Tsr-Venus fusion protein. Because the protein was localized to the membrane, its diffusion was reduced, enabling the detection of single Tsr-Venus molecules. Fluorescent protein molecules were photobleached on detection so that further production could be monitored. Protein expression was also geometrically distributed with an average burst size of 4.2 molecules and a burst frequency of 1.2 per cell cycle (71).

The statistics of mRNA production have been monitored in two eukaryotic systems: the social amoeba *Dictyostelium* (16) and a mammalian cell line (48). Both studies observed bursts of transcription, again measured using an RNA sequence recognized by a fluorescently labeled RNA binding protein (16) or probe (48). In the study of mammalian cells, a tetracycline-inducible promoter (29) was integrated at a specific location in the genome, allowing the measurement of transcriptional burst statistics at various levels of gene expression. Surprisingly, an increased level of gene expression corresponded to an increase in the burst size but not the burst frequency of mRNA, in contrast with the *E. coli* study (28). One possible model put forward was that periodic waves of chromatin decondensation set the burst frequency, and the activity of a transcriptional activator set the burst size. Regardless of mechanism, the increased burst size implies that the intrinsic noise does not necessarily decrease monotonically with increased expression (48), as had been the case for all other genes studied. Finally, transcription from an endogenous gene encoding the large subunit of RNA polymerase II occurred in bursts of ∼120 mRNAs and with the same statistics as the integrated genes above.

These in vivo single-molecule studies directly confirm the assumptions made regarding the kinetic mechanisms of gene expression (**Figure 2*a***). Promoters shift randomly between ON and OFF states, with exponentially (geometrically) distributed times in each state. Production of mRNA during the ON state then occurs in a Poisson (binomial) manner, leading to an exponential (geometric) distribution in the burst size of mRNA and a gamma (negative binomial) steady-state distribution of mRNA, when the time spent in the ON state is small relative to the mRNA half-life (see Reference 48 for the mRNA distribution without this assumption). If a promoter is always in the ON state, steady-state mRNA numbers are Poisson distributed. Protein production occurs in exponentially (geometrically) distributed bursts, leading to a gamma (negative binomial) steady-state distribution in protein. In the combined case in which both mRNA and proteins are produced in bursts, protein distributions still remain gamma distributed (see References 11, 15, 45). Because these models only account for intrinsic noise, the match between these models and the data supports the idea that the variability in the single-molecule experiments arises from intrinsic sources; this is likely a

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

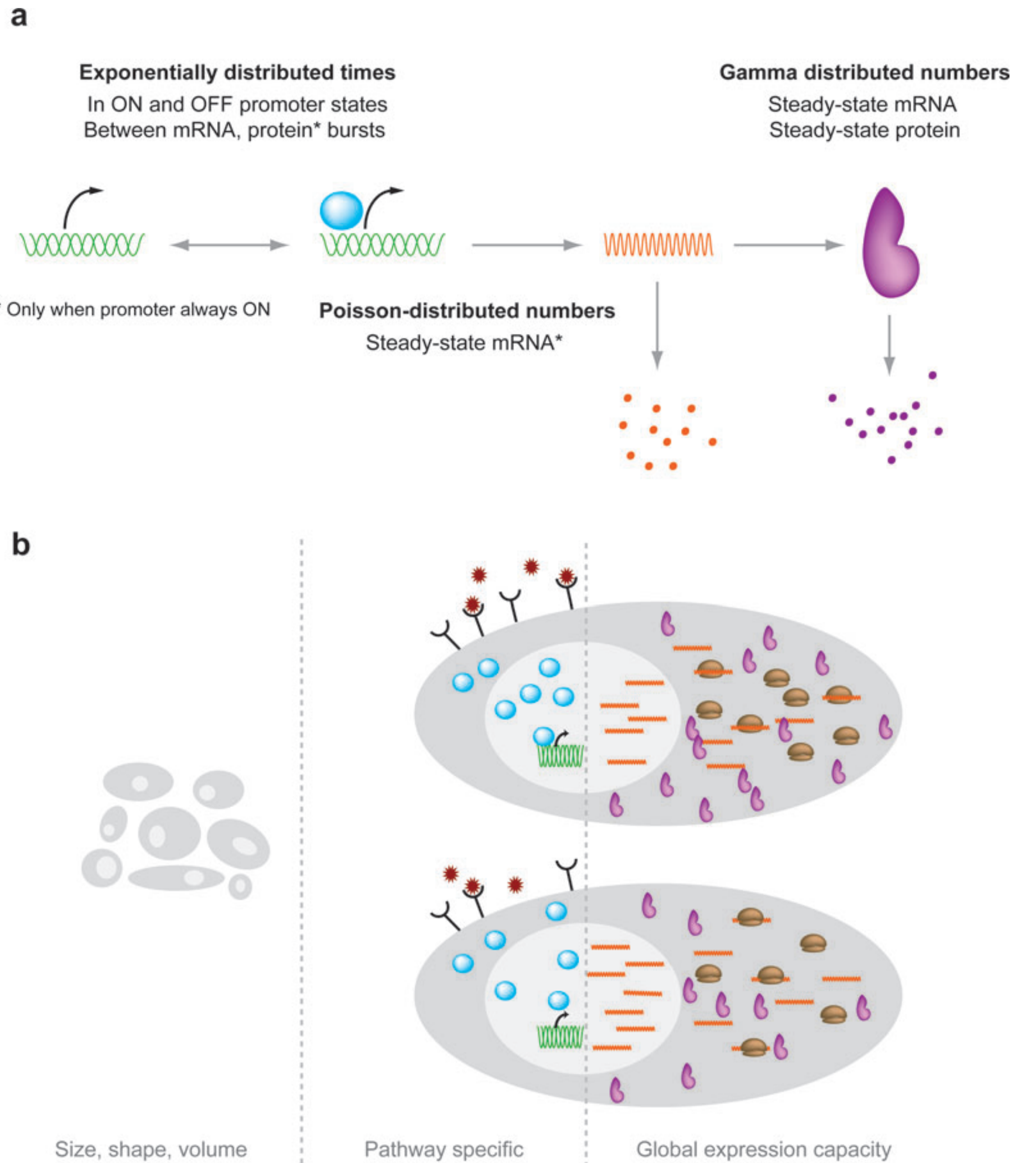


Figure 2

Intrinsic and extrinsic noise. (*a*) Stochastic model for intrinsic noise in the expression of a single gene. (*b*) One source of extrinsic noise in the expression of a single gene is variation in population structure, including differences in cell size, shape, and organelle composition. However, even in two identically structured cells, there can still be pathway-specific variation owing to intrinsic or extrinsic fluctuations in the activities of members of the upstream signaling pathway that affect that gene's expression. In addition, there can be global variation owing to cell-to-cell differences in overall expression capacity, which may be related to the amount of transcriptional and translational machinery and/or the cell's metabolic state.

result of the low level of expression for the protein counting experiments and a possible robustness of mRNA to extrinsic fluctuations in the RNA counting assays (28, 48). These models also suggest that the burst size, frequency, and associated rates are a more natural parameterization for the intrinsic noise than the CV.

Steady-State Protein Distributions of Single Genes and the Nature of Extrinsic Noise

The above stochastic models only describe the statistics of intrinsic noise. Early investigations in both *E. coli* and *Saccharomyces cerevisiae* demonstrated that the absolute value of extrinsic noise is maximal at intermediate levels of gene expression and is the dominant source of noise when gene expression is high (12, 24, 49). Recent studies focused on identifying the sources of extrinsic noise (**Figure 2*b***), often by posing models that incorporate both intrinsic and extrinsic sources, and attempted to predict the distribution of protein expression.

Mathematical models that couple gene expression to cell cycle–driven population dynamics can be useful in distinguishing between different sources of extrinsic noise (51, 69). A study in budding yeast used a structured population model to account for extrinsic variation owing to population dynamics. The researchers found the presence of an extrinsic noise floor in exponentially growing yeast cells that resulted from the steady-state structure (e.g., cell size, shape) of the population (69). However, extrinsic variability in the expression of the *GAL1* gene, which is dominated by extrinsic noise (49), could not be completely accounted for by differences in population structure. The additional variability was suggested to result from fluctuations in an upstream regulator (69). Similarly, a study in *E. coli* used a mathematical model to account for cell cycle–dependent changes in expression owing to gene copy number. Even after such a correction, there was still substantial extrinsic variation (CV of 40%) (51). Clearly, extrinsic noise results from more than just differences in population structure.

Experimentally, steady-state protein distributions have been described as both normal and log-normal-like or long-tailed (5, 42, 51, 54). If the noise is dominated by random births and deaths of mRNA and protein, and mRNAs are short-lived, then the protein number should be gamma-distributed (as seen in References 15 and 71). At low abundance, a gamma distribution would be long-tailed, whereas at higher abundance it would approach a normal distribution. However, when extrinsic noise is dominant, it is difficult to predict what the distribution will be without knowing the source of the noise. A recent study assumed that extrinsic noise was dominated by low-frequency fluctuations in the expression capacity of cells (see **Figure 3**). A stochastic model that incorporated a random, slowly fluctuating protein production rate was able to predict the long-tailed steady-state distribution of *lacY*-promoted green fluorescent protein (GFP) in *E. coli* (40). One biological interpretation is that some determinant of expression capacity is normally distributed among different cells, leading to the log-normal-like distribution observed. Protein synthesis components such as ribosomes are more likely the determinant because mRNA noise measured in the single-molecule studies was primarily intrinsic (28, 48). Although no factors responsible for these global cell-to-cell differences have been identified, we speculate that they may be linked to central metabolism.

Expression capacity: an experimentally defined quantity that is measured by comparing the expression of a strong, global, constitutively active promoter in different cells

GFP: green fluorescent protein

Single-Cell Tracking of Gene Expression and the Dynamics of Noise

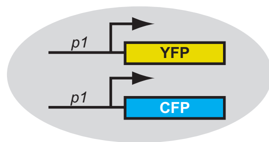
Real-time tracking of gene expression in single cells over many generations has been accomplished by combining fluorescent reporters, fluorescence microscopy, and either agarose pads or microfluidic chambers (3, 14, 19, 40, 51). Investigators recently used this technique to study the quantitative

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org by Johns Hopkins University on 10/31/11. For personal use only.

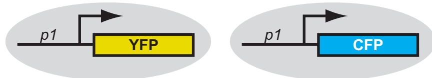
Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

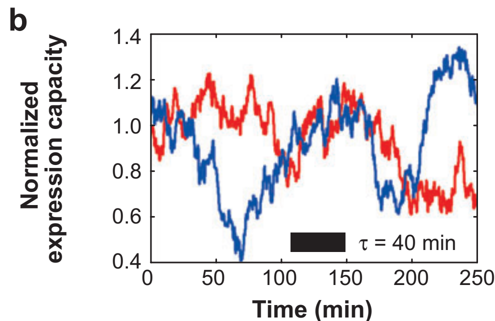
a

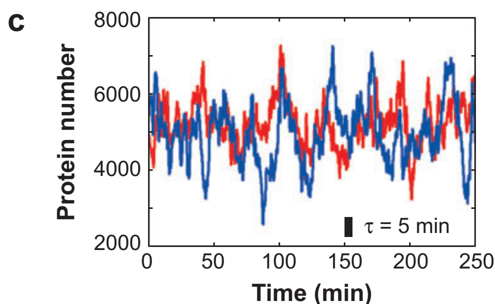
Intrinsic noise:
comparing expression in the same cell

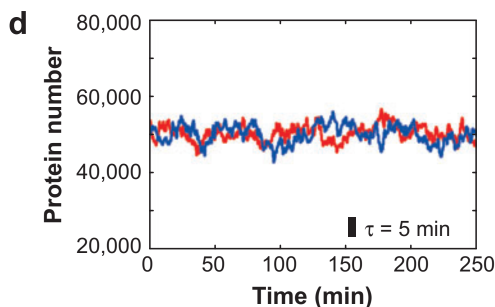


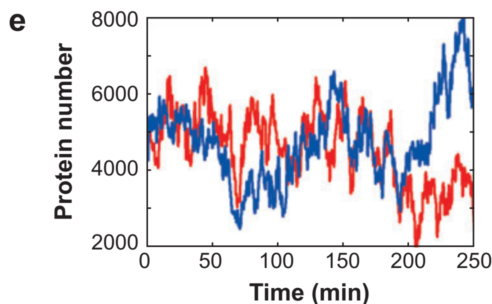
Total noise:
comparing expression in different cells











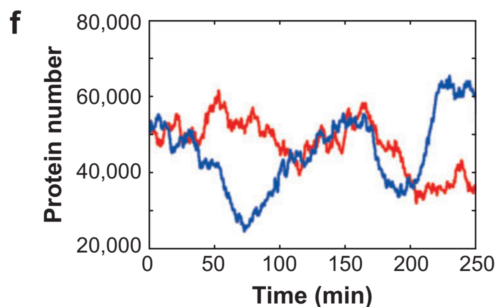


Figure 3

Simulating intrinsic and global expression noise. (*a*) Red and blue traces represent two different dynamic realizations of the same stochastic simulation. For intrinsic noise traces (*c* and *d*), this is analogous to the expression of two identical promoters in one cell. Total noise traces (*e* and *f*) are analogous to expression from promoters in different cells. Global expression capacity fluctuates slowly (*b*), with a long autocorrelation time (associated with division time in *E. coli*, but not in all cell types). Intrinsic noise (*c* and *d*) fluctuates quickly with a short autocorrelation time (associated with mRNA degradation if generated by translational bursting). Intrinsic noise is greater at lower abundances ($\langle p \rangle = 5000$; panel *c*) versus higher abundances ($\langle p \rangle = 50{,}000$; panel *d*). Total noise (*e* and *f*) can be modeled by introducing low-frequency fluctuations in expression capacity that enter multiplicatively in the mRNA or protein production rate. Trajectories in *e* and *f* were generated using the global fluctuations in *b*. Both *e* and *f* follow the trajectory in *b*, but the differences are washed out to a much greater extent in *e* where low-abundance mean intrinsic fluctuations are comparable in magnitude. CFP, cyan fluorescent protein; YFP, yellow fluorescent protein.

relationship between transcription factor concentration and the steady-state rate of gene expression from a target promoter, a relationship described as a gene regulatory or transfer function (51). Single-cell gene regulatory functions were calculated by following lineages of bacteria growing in real time. Even after compensating for cell cycle–related effects, there still remained significant variation in the expression rate at any particular level of transcription factor (CV of 40%). Measurement of the expression rate from two identical promoters indicated that the extrinsic component dominated the total noise.

These authors then calculated how long a fluctuation in gene expression persisted by measuring the autocorrelation function of the noise. Fluctuations that were extrinsic in origin had an autocorrelation time of approximately 40 min or one cell cycle, suggesting that cell division contributes significantly to extrinsic variation. In contrast, intrinsic fluctuations died off quickly with an autocorrelation time of less than 9 min, the time resolution of the experiment; this is not surprising given the fast death of mRNA that drives intrinsic fluctuations (51). Similar experiments done in *E. coli* grown at various temperatures demonstrated that the measured autocorrelation time of the noise coincided with the different growth rates (3).

A Genome-Wide View of Noise in Single Genes

The first genome-wide view of noise estimated intrinsic noise in the expression of each gene in budding yeast, assuming that noise arose from translational bursting (Equation 9) and using experimentally determined estimates of transcription and translation rates. The results indicated that essential proteins and those involved in large protein complexes tended to have less variation, suggesting a functional need for noise minimization (26). Two recent experimental studies in budding yeast now provide a global view of noise in gene expression. Both studies used flow cytometry to measure protein abundance and variability at the single-cell level, employing a collection of yeast strains expressing GFP fusion proteins (33). These studies used a single reporter, but by restricting analysis to a small, homogenous subset of cells, they eliminated most (but not all) of the extrinsic noise, which in this study results from population structure. Both studies verified minimization of the extrinsic noise contribution by using the two-color approach to analyze a smaller subset of genes (5, 42).

One study focused on 43 genes involved in four coexpressed transcription modules: stress, proteasome, ergosterol synthesis, and rRNA processing (5). These genes were subject to 11 different environmental conditions, and the transient response to each perturbation was measured. A second study analyzed noise in ~50% of the yeast proteome under two steady-state conditions—exponential growth in rich or minimal medium (42). Both studies reported that intrinsic noise scales inversely with mean protein abundance with a burst size of ~1200 proteins. This burst could be explained theoretically as corresponding to the average number of proteins made per transcript if noise were driven by mRNA fluctuations (Equation 9) or the average number of proteins made when the promoter is in the ON state, if the noise were driven by promoter fluctuations (5). [For comparison, translational burst sizes in *E. coli* for *lac*-promoted GFP are ~35 proteins per mRNA (40).] This observation leads to the remarkable conclusion that post-transcriptional events occur at similar rates for at least half the budding yeast genome in exponential growth conditions.

When cells were subject to a perturbation, the extent to which the noise in each gene changed depended greatly on the module to which it belonged (5). The authors defined a "noise residual," which compared the measured noise with the noise predicted by Equation 9 using the measured protein abundance. Stress genes had large noise residuals in contrast with the small noise residuals of genes in other modules. The second

Autocorrelation: the correlation of a quantity (the gene expression rate or protein number) with itself over successive time intervals

Autocorrelation time: the characteristic time it takes for a fluctuation (in gene expression rate) to decay to the mean (gene expression rate)

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

study used a slightly different noise residual in which they compared the noise of any one gene's expression with the median noise of all genes with similar expression levels (42). They looked for correlations between factors known to influence noise and the noise residual. Most notably, the type of transcriptional regulation correlated strongly with the noise residual. Many noisy genes tend to be regulated by chromatin remodeling factors, which are thought to play a dominant role in promoter fluctuations (26). Studies in both yeast (7) and mammalian cells (48) suggest that chromatin remodeling is locally correlated so that contiguous or homologous genes tend to have correlated fluctuations, and the absolute genomic position affects the extent of those correlations. A strong TATA box, which has already been shown experimentally to increase noise (49), also correlates strongly with noisy genes. In contrast, transcription factors known to disrupt chromatin structure correlate with low noise genes. Moreover, essential proteins and proteins related to translation, the ribosome, the proteasome, and the secretory pathway exhibit low noise as predicted (26).

A striking conclusion of these studies is the lower limit seen for the noise residual, what one might call an intrinsic noise floor (5, 42). Genetic networks can mitigate noise further (see below), but at least in budding yeast, intrinsic fluctuations are not reduced below the Poisson expectation based on protein abundance given by Equation 9. This finding should prove invaluable, allowing one to apply a lower bound to the intrinsic noise in larger models of genetic networks in budding yeast. However, we do not expect this bound to exist in general, especially in higher metazoans. In fact, lower abundance does not necessarily result in higher intrinsic noise in mammalian cells (48). At the oncogenic *c-myc* promoter, multiple molecular mechanisms could possibly function to reduce noise in its expression (17). Requiring a (nearly) irreversible sequence of reactions and multiple factors prior to transcription at the promoter can reduce noisy effects of any one component. The *c-myc* promoter may also have real-time feedback control of transcription through several AT-rich enhancer sequences that melt owing to torsional forces produced by RNA polymerase II during transcription, possibly preventing transcription factor binding and affecting transcription (17).

At the other end of the spectrum, many genes can be noisier than this floor, and they are often highly regulated. There is a natural trade-off between the need to regulate a gene's expression, providing a wide dynamic range, and the need to control its noise. First, regulation makes gene expression susceptible to fluctuations in upstream factors. Second, a highly regulated gene tends to have multiple promoter states in which slow fluctuations, often controlled by chromatin remodeling (12, 34, 49), cause expression noise. Third, a large change in expression rate requires amplification at the level of transcription, further amplifying noisy promoter fluctuations (49). Fourth, there is the potential for regulation at the mRNA level. Although this is less important in budding yeast, antisense RNAs, microRNAs, RNAi, and RNA splicing are all involved in regulating expression and could potentially be noisy (see Reference 59 for a recent review). Fifth, post-translational mechanisms could also play a role (61).

PUTTING IT ALL TOGETHER: SIMPLE NETWORKS

Linear Networks and the Importance of Correlated Noise Sources

A linear genetic network can be defined as a sequence of genes regulating each other in series. The extrinsic noise in each component arises from two sources: pathway-specific noise that is transmitted from the upstream regulators and global noise in the expression capacity of a cell. Investigators studied noise propagation through a linear network in a simple system in *E. coli* (**Figure 4*a***) using three components: (component 1) a titratable

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

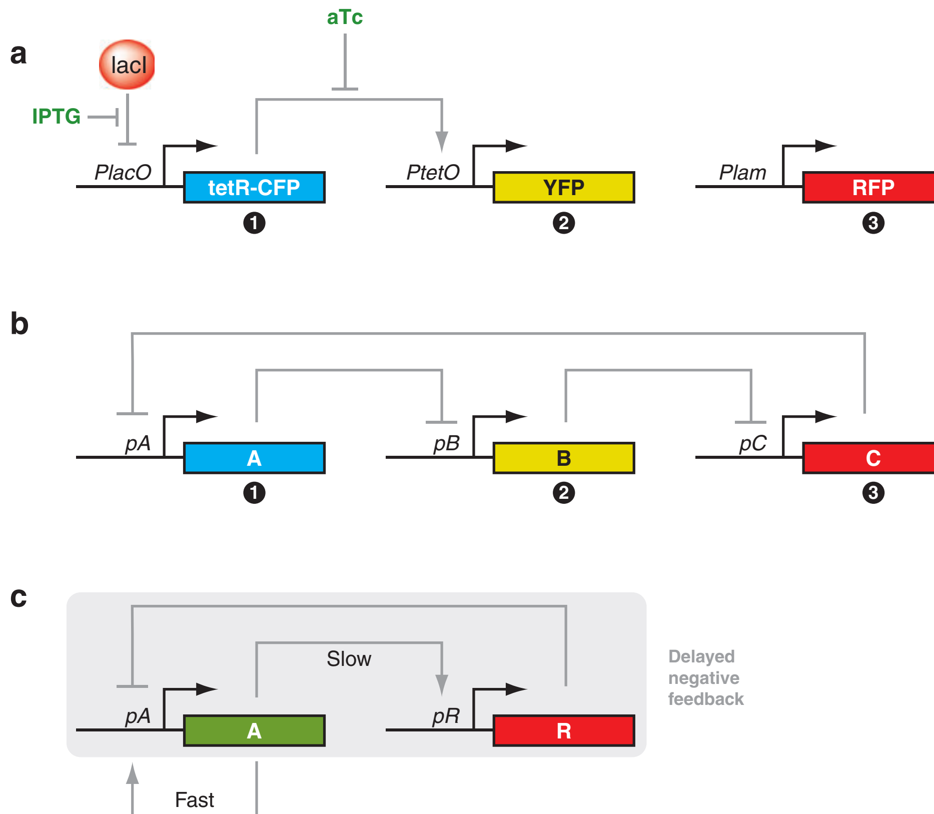


Figure 4

(*a*) The two-gene cascade consists of (1) a *lac*-promoted tetR–cyan fluorescent protein (CFP) fusion driving (2) a *tet*-promoted yellow fluorescent protein (YFP), and (3) a lambda-promoted red fluorescent protein (RFP) that serves as a global reporter. The coupling and expression level of (1) and (2) are easily modified by varying [IPTG] and [aTc]. Figure adapted from Reference 46. (*b*) Design of a ring oscillator consisting of a ring of three transcriptional repressors, termed the "repressilator" (23). (*c*) The boxed section of this design is a delayed negative feedback loop. An activator represses itself through slow activation of a repressor. Embedding a fast positive feedback loop converts this to a relaxation oscillator that functions reliably for a wider range of parameters when noise is significant.

promoter driving expression of a repressor-fluorescent protein fusion (R), (component 2) an R repressible promoter driving expression of a second fluorescent protein, and (component 3) a separate strong promoter driving expression of a third fluorescent protein (46). Components 1 and 2 represented the linear network, and component 3 was used to control for global fluctuations in gene expression. Using a Langevin approach, they derived equations (44) to predict the noise in the expression of component 2 as a function of both global noise in component 3 and pathway-specific noise in component 1. The model quantitatively predicted the noise in component 2 as a function of the other components and provided a straightforward theoretical and experimental recipe for analyzing how noise is transmitted in linear networks (46). Perhaps the most interesting qualitative insight from this study was that the correlation of connected noise sources can have either amplifying or suppressing effects on the noise, and this depends on promoter logic. In the case of

Expression-specific noise: an experimentally defined component of the extrinsic noise that consists of extrinsic variation due to cell-to-cell differences in the expression capacity

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

a repressor, a global fluctuation in cellular expression capacity increases expression of component 2 but is suppressed by the increased expression of the repressor, component 1. Although not demonstrated, global fluctuations should be amplifying for an activator (46).

Investigators have studied experimentally and computationally the dynamics of linear cascades of one, two, or three repressors in *E. coli* (30). An inducible promoter controlled expression of the first repressor in the cascade, and each repressor acted on the gene immediately downstream. Longer cascades had more ultrasensitive steady-state transfer functions but much slower response times. Moreover, at intermediate induction levels, the transmitted noise became significant in these systems. It is tempting to speculate that much of the transmitted noise resulted from intrinsic fluctuations in members of the cascade, as a repressor network would be robust to global fluctuations, but only the total noise was measured.

Researchers have also studied noise transmission in a natural biological network, the pheromone response pathway in budding yeast (19). In this study, the authors developed a strategy to measure the intrinsic noise, pathway-specific noise, and global expression noise components as discussed above (see Experimental Measurement and Classification of Noise) and depicted in **Figures 1*b*** and **2*c***. The system was subdivided into a linear network consisting of a pheromone input, a pathway subsystem, and an expression subsystem. A pheromone receptor coupled the pheromone input to the pathway subsystem, and pheromone-responsive genes coupled the pathway subsystem to the expression subsystem. The subsystems were assumed to be otherwise insulated. At high pheromone levels, almost all (>75%) the cell-to-cell variation could be explained by differences in expression capacity and cell cycle position, similar to other studies in budding yeast (see above). At lower pheromone levels, pathway-specific noise played a significant role (~60%) in the observed variation, but intriguingly, the total noise was similar to the high pheromone experiment. In fact, the expression capacity was not affected by the differences in pathway activation, but the expression-specific noise was negatively affected. In this system, expression-specific noise appears negatively correlated with pathway-specific noise. The mechanism proposed to account for these observations is conceptually similar to what was demonstrated in the *E. coli* study (46)—large fluctuations in expression capacity are compensated for by correlated fluctuations in pathway-specific regulators (19).

Finally, two redundant mitogen-activated protein kinases, Fus3 and Kss1, lie within the pheromone-response pathway. Signaling through Fus3 resulted in half as much pathway-specific noise as signaling through Kss1, leading the authors to suggest that the extent of cell-to-cell variation is regulated by the choice of mitogen-activated protein kinase (19). This is a compelling example of how the choice of signaling component in a branched pathway could affect pathway-specific noise but because the total noise remains constant regardless of choice, the functional relevance is unclear.

Negative Feedback

Both theoretical and experimental investigations demonstrate the ability of negative feedback to suppress noise (9, 64) (see Reference 34 for review). This is an intuitive idea, as potentially large fluctuations are quickly stifled by feedback. Recently, the master equation for negative autoregulation of a single gene was solved exactly (31). The solution clearly demonstrates that when the rate of repressor binding/unbinding is slow compared with its degradation rate, fewer repressors are required to achieve the same average extent of inhibition. At least three factors contribute to the ratio of these two rates: speed of response, energetic requirements, and fidelity. Fast binding/unbinding requires more proteins but results in less noise. The benefits

of slower binding/unbinding are a switch-like transfer function, a binary response, and fewer proteins needed, but these come at the cost of a noisier outcome, more energy expended owing to the high protein degradation rate, and a slower transient response.

As with any dynamic process, noise can be studied both in the time and frequency domain. A series of studies (57, 58) demonstrated that negative feedback not only reduces noise, but also shifts it to higher frequencies: The stronger the feedback and the slower the dynamics of the repressor (primarily through longer repressor half-life), the greater the shift. Because a downstream process only sees frequency components lower than its own characteristic frequency, shifting some of the noise to an invisible high-frequency regime aids in attenuation. This spectral shift was recently demonstrated experimentally by measuring the autocorrelation time of noise in gene expression from the *tetR* promoter in *E. coli* in the absence and presence of negative feedback (3). Because low-frequency extrinsic noise is often dominant, a realistic biological implementation of this strategy would have to involve long, slow, cascaded negative feedback systems that may not be able to respond quickly enough to changing stimuli.

In the sections above, we focus on examining the noise at the steady state. However, researchers have also analyzed the intrinsic noise in the transient response of a negative feedback system (63). In this study, the authors solved numerically linearized Fokker-Planck equations using the phenomenological Hill equation to describe the feedback. In contrast to the steady-state condition, increasing feedback strength by either changing the Hill coefficient or the (repressor) binding constant did not always lead to a reduction in noise. In addition, initial conditions dramatically changed the dynamics of the noise. This study illustrates the general challenge of modeling nonlinear processes far from steady state and the danger in extrapolating steady-state results to transient phenomena.

Positive Feedback

Nonlinear, cooperative positive feedback can lead to multiple stationary solutions of a gene's expression rate (1, 8, 25, 43, 59, 70). Stochastic fluctuations have the potential to switch the system between these multiple stable states, with a calculable average frequency based on rates (35). Detailed experimental analysis in two sugar uptake systems in two different organisms, lactose in *E. coli* (43) and galactose in *S. cerevisiae* (1), demonstrates a shared architecture and the existence of two stable sugar utilization states. The sugar (or a chemical analog) activates the expression of sugar transporters that positively feedback to increase sugar uptake. In the case of galactose, an embedded negative feedback loop is also activated that reduces the strength of the positive feedback. With strong positive feedback, the two stable states are less susceptible to noise-induced transitions. In the galactose system, the strength of positive feedback could be affected by modifying the strength of the negative feedback, enabling the discovery of a regime in which noise-induced transitions between the two stable states are fast enough to effectively erase the deterministic bistability predicted (1).

In the lactose system, investigators observed two populations corresponding to each stationary state over an approximately tenfold range of TMG (thio-methyl-β-D-galactosidase) concentration (a lactose analog used here) (43). For such bistable systems, two populations are often seen in the bistable region, even after careful preparation of the system in the induced (or uninduced) state. The deterministic description would not predict these two populations, and their presence is explained by extrinsic differences between cells. These extrinsic differences were explicitly modeled as global fluctuations in expression capacity that die out on the order of the cell cycle time (40) (**Figure 3**). This allowed the correct dynamical prediction of the protein distribution. When deterministic dynamics dominate, a ballistic transition

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

occurs as the entire protein distribution moves from one state to the other. When stochastic fluctuations are important, a transient bimodal distribution occurs as cells fluctuate between the two states as they approach the unimodal steady-state distribution. In the lac system, ON cells rapidly introduced into an OFF environment exhibited ballistic dynamics. In contrast, OFF cells rapidly introduced into an ON environment exhibited stochastic dynamics (40).

Noise not only can eliminate deterministic bistability, it can also create a bistable system when a deterministic description has only one stationary solution. Noncooperative positive transcriptional feedback leads to only one deterministic solution, but introducing noise into the transcription factor promoter interaction results in a bistable system (35). In fact, even in systems in which deterministic bistability is possible, large stochastic fluctuations lead to bistability for a larger range of biochemical parameter values. A recent theoretical study considers how noise can induce bistability when enzymatic futile cycles are involved (52). Let us consider the dynamic interconversion of protein X to a phosphoisoform X^* by a relevant kinase and phosphatase. If there is noise in either enzyme's activity at a frequency lower than the usually fast interconversion of X and X^*, then a stochastic description leads to a bimodal distribution of X^* (and X) that corresponds to two bistable states (52) (**Figure 5**). Noise often changes the behavior qualitatively when the diffusive term in the associated Fokker-Planck equation depends on the state variable in a nonlinear fashion.

Noise and Oscillations

Periodic oscillations in both time and space are present and utilized in biological systems. Whether their role is to keep time, encode information for cell signaling, coordinate diverse cellular processes, or simply be the consequence of a periodic external input, the underlying molecular mechanism causing the oscillation has been of great interest. How do biological oscillations remain reliable in the face of noise? Noise can result in unreliable oscillations, but, conversely, sometimes noise is required for reliable oscillations. It all depends on the design.

A ring, or phase oscillator, consists of a closed chain of negative feedback elements (**Figure 4*b***). A synthetic genetic circuit that realizes this design has been built in bacteria (23). The presence and strength of sinusoidal-like oscillations rely on the short half-life of each repressor in the circuit. Oscillations tend to be noisy, showing both cell-to-cell and single-cell variation in period. This variation likely arises both from intrinsic fluctuations in the expression of each repressor that dominate during periods of low expression and from extrinsic fluctuations that affect repressor half-life.

When negative feedback is coupled with a large enough time delay associated with transcription, translation, protein modification, and/or a genetic network, the stationary solution is destabilized, resulting in oscillatory behavior (13) that may be desirable. In fact, negative feedback loops in which the first element activates the second element, which then downregulates the first, are fairly common in nature (41) (**Figure 4*c***).

One well-studied delayed negative feedback loop is p53-Mdm2, in which (slow) p53 expression activates Mdm2 through a (fast) protein-protein interaction, ultimately leading to a reduction in p53 expression. An early model (37) described a complete set of parameters that would admit oscillations, and several subsequent mechanisms have been proposed (18, 38, 41). One study demonstrated oscillations in p53 levels in MCF-7 (human breast cancer epithelial) cells on irradiation, characterized by repeated pulses of p53 expression (36). Increased irradiation raised the probability of multiple pulses; the timing of the initial pulses was somewhat variable but became more precise with subsequent pulses, the pulse height varied threefold, and the pulse width was fairly constant (36). A second study

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

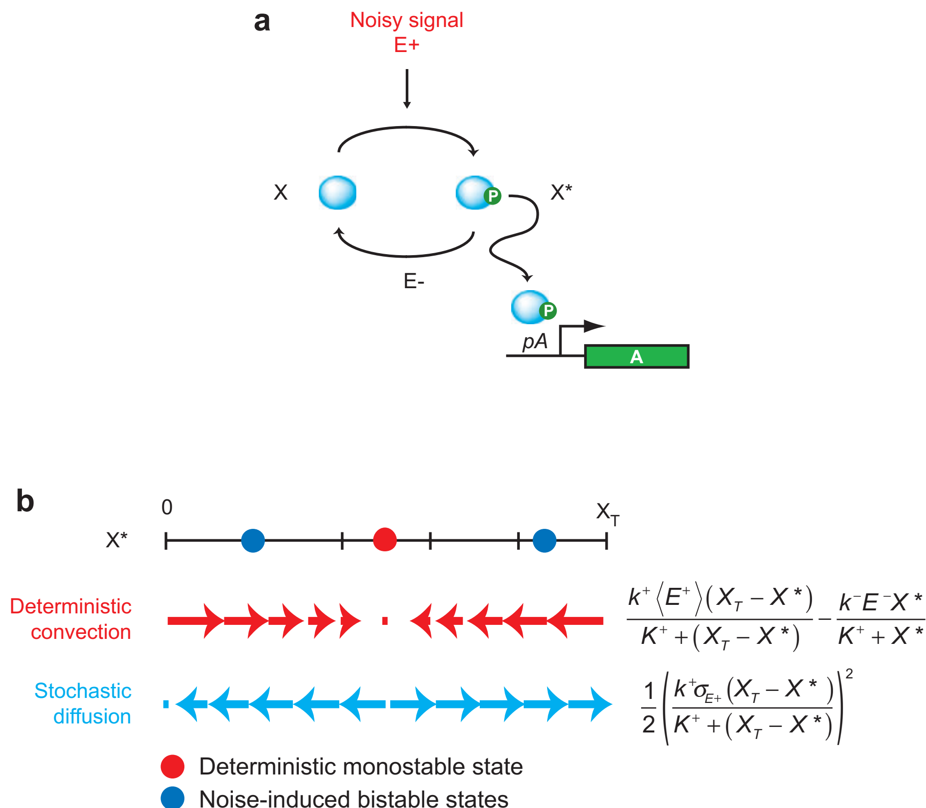


Figure 5

Noise-induced bistability. (*a*) Many transcription factors are regulated by an enzymatic futile cycle. Here, the transcription factor X is phosphorylated by kinase E^+ to yield the active phosphoform X* that is competent for activating a downstream gene. Phosphate is removed from X* by the phosphatase E^-, completing the cycle. We assume the source of noise is low-frequency extrinsic fluctuations in the kinase E^+, although one could also include noise in the phosphatase E^-. The relative amount of each phosphoform has an influence on gene expression and function. (*b*) The black line depicts the state space of numbers of X*, which can vary from 0 to X_T. The size and direction of the arrows below correspond to the magnitude and sign of the dynamic forces that govern X*. The red arrows illustrate how the Michaelis-Menten-like deterministic rate equations, listed to the right, drive the numbers of X* to a single stable steady state at $X_T/2$. The blue arrows represent the noise, a diffusive term driving X* away from the deterministic steady state, in the nonlinear manner listed to the right. This leads to two stable steady states. Figure based on work in Reference 52.

analyzed in depth the variability in the oscillation dynamics observed for many cells (27). These authors found that pulse height was highly variable (CV of 75%), but the peak width and period were less so (CV of 30% and 20%, respectively). Moreover, the autocorrelation function of the oscillatory dynamics suggested that the extrinsic noise causing these differences died off at ∼11 h, or half a cell cycle. Three deterministic models were able to generate the mean behavior of the observed data. To model the variability, the authors recast the deterministic equations into stochastic Langevin equations. Interestingly, when the authors used low-frequency noise in the level of production/transcription rate on the order of the autocorrelation time, they were able to recapitulate the observed

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

variability in pulse height and the smaller variability in pulse width and timing (27).

Embedding a positive feedback loop in the delayed negative feedback circuit results in a relaxation oscillator (**Figure 4*c***). This is the molecular oscillator found to drive circadian rhythms across many organisms (for a review, see Reference 10). The activator activates both itself and the repressor, but the dynamics of positive feedback is much faster than the dynamics of negative feedback. Therefore, the activator quickly flips from a low to a high level. The repressor slowly builds up until it crosses a critical level that causes the activator to quickly flip back to the low state. The existing repressor is eliminated slowly, but when it drops below a threshold, the activator flips to a high state again, driving the oscillation. Theoretical work shows that embedding the positive feedback loop renders the oscillator more robust to stochastic fluctuations (6). This design not only drives circadian rhythms, but also the cell cycle (20). Eliminating the embedded positive feedback in the *Xenopus laevis* cell cycle led to damped cell cycle oscillations (47). Moreover, even when the positive feedback is weak and in a deterministic setting the activator is never able to flip from the low state to high state, stochastic fluctuations can flip this switch and allow the system to oscillate (68). This is the hallmark of an excitable system, in which stochastic fluctuations quickly excite it away from the stationary state and then it slowly relaxes back. In fact, the competent state in *Bacillus subtilis* corresponds to the slow relaxation of an excitable system driven by noise (60).

OUTLOOK

In this review, we focus on noise in gene expression and its role in transcriptional regulatory networks. With the exception of our discussion of enzymatic futile cycles (**Figure 5**), we neglect the effects of fluctuations in protein-protein and protein-metabolite interactions. We hypothesize that some of these fluctuations will affect gene expression, especially the low-frequency fluctuations in expression capacity. One theoretical study (22) attempts to connect fluctuations in translation rate, a primary determinant of expression capacity, to fluctuations of metabolites, but a thorough investigation will require the measurement of global fluctuations of expression capacity in cells in different metabolic conditions and the connection of fluctuations in expression capacity to variation in the relevant metabolic pathway. In addition, protein-protein and protein-metabolite fluctuations occur fast enough that transport processes can play a dominant role. For example, the spatiotemporal oscillations of certain proteins involved in *E. coli* cytokinesis can be explained robustly using a stochastic description (32). Finally, both transport (4, 67) and reactions (55, 56) of molecules are influenced by the three-dimensional structure of the environment in which they take place, but knowledge of the important structural features and structural dynamics is currently lacking.

Most of the examples discussed in this review pertain to free-living, fast-growing unicellular microbes. Because the importance of the size and timescale of noisy fluctuations depends on the length scale of the reaction environment and timescale of affected dynamical processes, we expect investigations in different organisms to alter our understanding of noise in gene expression. For example, the autocorrelation time of extrinsic fluctuations is not on the order of cell cycle in mammalian systems (27), and possibly not in budding yeast (19), in contrast to observations in rapidly dividing bacteria (51). Protein noise can also depend on subcellular localization of the protein (42). This should be further investigated in organisms with heterogeneous intracellular environments at the organelle level. For example, the packaging of low-abundance proteins in the secretory pathway might result in significant variation in their vesicular concentration, as in the case of neurotransmitter release (53). Finally, protein concentration, not protein number, is relevant for intracellular functions. Methods for

controlling population structure will not control for concentration in cells in which organelle composition has a large impact on cytosolic volume.

This review begins with the molecular details of noisy gene expression and progresses to its impact on global cellular processes and ultimately a selectable phenotype. Will a stochastic description be necessary in any realistic model of biological systems and their evolution? The answer undoubtedly depends on the system being modeled. It may be possible to use a higher-level stochastic description that dispenses with the mechanistic details of how fluctuations arise using a stochastic description, but determining the correct extent of abstraction remains difficult. With current and future advances in experimental technology allowing us to spy on fluctuations, and the creative application of modeling and simulation, our intuition can only improve.

SUMMARY POINTS

1. Both theoretical and experimental investigations have led to stochastic chemical kinetic models that describe the statistics of intrinsic noise in gene expression. Intrinsic fluctuations in protein number are driven by random births and deaths of mRNA and/or random promoter transitions from an inactive to a transcriptionally competent state. Recent experiments in both prokaryotes and eukaryotes have confirmed these statistics by directly measuring the production of single mRNA and protein molecules.
2. Sources of extrinsic noise in gene expression include the following: cell-to-cell variation in cell morphology, volume, and size; expression capacity (the rate of gene expression owing to differences in transcriptional, translational, and/or metabolic components); and fluctuations in pathway-specific molecular components that lie upstream of that gene.
3. Gene expression noise dynamics can be characterized by its autocorrelation time—the characteristic timescale over which the effects of a single fluctuation persist. Intrinsic noise has a fast autocorrelation time consistent with the mRNA half-life. Extrinsic noise has a longer-lived autocorrelation time that is organism specific. Modeling expression capacity using the appropriate autocorrelation time has accurately predicted protein distributions.
4. Global analysis of intrinsic noise in budding yeast indicates that highly regulated genes exhibit expression that is noisier than average. Remarkably, noise is not reduced more than what would be expected for intrinsic noise owing to random births and deaths of mRNA.
5. Expression noise can be shaped by the architecture of transcriptional regulatory schemes, such as feedback or feedforward loops. Certain combinations of feedback loops drive oscillations in gene expression. Noise in gene expression can either increase or decrease the reliability of these oscillations, depending on the regulatory scheme.

LITERATURE CITED

1. Acar M, Becskei A, van Oudenaarden A. 2005. Enhancement of cellular memory by reducing stochastic transitions. *Nature* 435:228–32

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

2. Arkin A, Ross J, McAdams HH. 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics* 149:1633–48
3. Austin DW, Allen MS, McCollum JM, Dar RD, Wilgus JR, et al. 2006. Gene network shaping of inherent noise spectra. *Nature* 439:608–11
4. Banks DS, Fradin C. 2005. Anomalous diffusion of proteins due to molecular crowding. *Biophys. J.* 89:2960–71
5. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, et al. 2006. Noise in protein expression scales with natural protein abundance. *Nat. Genet.* 38:636–43
6. Barkai N, Leibler S. 2000. Circadian clocks limited by noise. *Nature* 403:267–68
7. Becskei A, Kaufmann BB, van Oudenaarden A. 2005. Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat. Genet.* 37:937–44
8. Becskei A, Seraphin B, Serrano L. 2001. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* 20:2528–35
9. Becskei A, Serrano L. 2000. Engineering stability in gene networks by autoregulation. *Nature* 405:590–93
10. Bell-Pedersen D, Cassone VM, Earnest DJ, Golden SS, Hardin PE, et al. 2005. Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nat. Rev. Genet.* 6:544–56
11. Berg OG. 1978. A model for the statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.* 71:587–603
12. Blake WJ, Kaern M, Cantor CR, Collins JJ. 2003. Noise in eukaryotic gene expression. *Nature* 422:633–37
13. Bratsun D, Volfson D, Tsimring LS, Hasty J. 2005. Delay-induced stochastic oscillations in gene regulation. *Proc. Natl. Acad. Sci. USA* 102:14593–98
14. Brehm-Stecher BF, Johnson EA. 2004. Single-cell microbiology: tools, technologies, and applications. *Microbiol. Mol. Biol. Rev.* 68:538–59
15. Cai L, Friedman N, Xie XS. 2006. Stochastic protein expression in individual cells at the single molecule level. *Nature* 440:358–62
16. Chubb JR, Trcek T, Shenoy SM, Singer RH. 2006. Transcriptional pulsing of a developmental gene. *Curr. Biol.* 16:1018–25
17. Chung HJ, Levens D. 2005. *c-myc* expression: keep the noise down! *Mol. Cells* 20:157–66
18. Ciliberto A, Novak B, Tyson JJ. 2005. Steady states and oscillations in the p53/Mdm2 network. *Cell Cycle* 4:488–93
19. Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, et al. 2005. Regulated cell-to-cell variation in a cell-fate decision system. *Nature* 437:699–706
20. Cross FR, Siggia ED. 2005. Shake it, don't break it: positive feedback and the evolution of oscillator design. *Dev. Cell* 9:309–10
21. Elf J, Ehrenberg M. 2003. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Res.* 13:2475–84
22. Elf J, Ehrenberg M. 2005. Near-critical behavior of aminoacyl-tRNA pools in *E. coli* at rate-limiting supply of amino acids. *Biophys. J.* 88:132–46
23. Elowitz MB, Leibler S. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–38
24. Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002. Stochastic gene expression in a single cell. *Science* 297:1183–86
25. Ferrell JE Jr, Machleder EM. 1998. The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science* 280:895–98
26. Fraser HB, Hirsh AE, Giaever G, Kumm J, Eisen MB. 2004. Noise minimization in eukaryotic gene expression. *PLoS Biol.* 2:e137

27. Geva-Zatorsky N, Rosenfeld N, Itzkovitz S, Milo R, Sigal A, et al. 2006. Oscillations and variability in the p53 system. *Mol. Syst. Biol.* 2:2006.0033
28. Golding I, Paulsson J, Zawilski SM, Cox EC. 2005. Real-time kinetics of gene activity in individual bacteria. *Cell* 123:1025–36
29. Gossen M, Bujard H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA* 89:5547–51
30. Hooshangi S, Thiberge S, Weiss R. 2005. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl. Acad. Sci. USA* 102:3581–86
31. Hornos JE, Schultz D, Innocentini GC, Wang J, Walczak AM, et al. 2005. Self-regulating gene: an exact solution. *Phys. Rev. E* 72:051907
32. Howard M, Rutenberg AD. 2003. Pattern formation inside bacteria: fluctuations due to the low copy number of proteins. *Phys. Rev. Lett.* 90:128102
33. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, et al. 2003. Global analysis of protein localization in budding yeast. *Nature* 425:686–91
34. Kaern M, Elston TC, Blake WJ, Collins JJ. 2005. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* 6:451–64
35. Kepler TB, Elston TC. 2001. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* 81:3116–36
36. Lahav G, Rosenfeld N, Sigal A, Geva-Zatorsky N, Levine AJ, et al. 2004. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat. Genet.* 36:147–50
37. Lev Bar-Or R, Maya R, Segel LA, Alon U, Levine AJ, Oren M. 2000. Generation of oscillations by the p53-Mdm2 feedback loop: a theoretical and experimental study. *Proc. Natl. Acad. Sci. USA* 97:11250–55
38. Ma L, Wagner J, Rice JJ, Hu W, Levine AJ, Stolovitzky GA. 2005. A plausible model for the digital response of p53 to DNA damage. *Proc. Natl. Acad. Sci. USA* 102:14266–71
39. McAdams HH, Arkin A. 1999. It's a noisy business: genetic regulation at the nanomolar scale. *Trends Genet.* 15:65–69
40. Mettetal JT, Muzzey D, Pedraza JM, Ozbudak EM, van Oudenaarden A. 2006. Predicting stochastic gene expression dynamics in single cells. *Proc. Natl. Acad. Sci. USA* 103:7304–9
41. Monk NAM. 2003. Oscillatory expression of Hes1, p53, and NF-κB driven by transcriptional time delays. *Curr. Biol.* 13:1409–13
42. Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, et al. 2006. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 441:840–46
43. Ozbudak EM, Thattai M, Lim HN, Shraiman BI, Van Oudenaarden A. 2004. Multistability in the lactose utilization network of *Escherichia coli*. *Nature* 427:737–40
44. Paulsson J. 2004. Summing up the noise in gene networks. *Nature* 427:415–18
45. Paulsson J. 2005. Models of stochastic gene expression. *Phys. Life Rev.* 2:157–75
46. Pedraza JM, van Oudenaarden A. 2005. Noise propagation in gene networks. *Science* 307:1965–69
47. Pomerening JR, Kim SY, Ferrell J, James E. 2005. Systems-level dissection of the cell-cycle oscillator: Bypassing positive feedback produces damped oscillations. *Cell* 122:565–78
48. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* 4:e309
49. Raser JM, O'Shea EK. 2005. Noise in gene expression: origins, consequences, and control. *Science* 309:2010–13
50. Risken H. 1989. *The Fokker-Planck Equation: Methods of Solution and Applications*. Berlin: Springer-Verlag. 2nd ed.

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

51. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. 2005. Gene regulation at the single-cell level. *Science* 307:1962–65
52. Samoilov M, Plyasunov S, Arkin AP. 2005. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl. Acad. Sci. USA* 102:2310–15
53. Sargent PB, Saviane C, Nielsen TA, DiGregorio DA, Silver RA. 2005. Rapid vesicular release, quantal variability, and spillover contribute to the precision and reliability of transmission at a glomerular synapse. *J. Neurosci*. 25:8173–87
54. Sato K, Ito Y, Yomo T, Kaneko K. 2003. On the relation between fluctuation and response in biological systems. *Proc. Natl. Acad. Sci. USA* 100:14086–90
55. Savageau MA. 1995. Michaelis-Menten mechanism reconsidered: implications of fractal kinetics. *J. Theor. Biol*. 176:115–24
56. Schnell S, Turner TE. 2004. Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Prog. Biophys. Mol. Biol*. 85:235–60
57. Simpson ML, Cox CD, Sayler GS. 2003. Frequency domain analysis of noise in autoregulated gene circuits. *Proc. Natl. Acad. Sci. USA* 100:4551–56
58. Simpson ML, Cox CD, Sayler GS. 2004. Frequency domain chemical Langevin analysis of stochasticity in gene transcriptional regulation. *J. Theor. Biol*. 229:383–94
59. Soller M. 2006. Pre-messenger RNA processing and its regulation: a genomic perspective. *Cell Mol. Life Sci*. 63:796–819
60. Süel GM, Garcia-Ojalvo J, Liberman LM, Elowitz MB. 2006. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* 440:545–50
61. Swain PS. 2004. Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *J. Mol. Biol*. 344:965–76
62. Swain PS, Elowitz MB, Siggia ED. 2002. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA* 99:12795–800
63. Tao Y, Jia Y, Dewey TG. 2005. Stochastic fluctuations in gene expression far from equilibrium: Ω expansion and linear noise approximation. *J. Chem. Phys*. 122:124108
64. Thattai M, van Oudenaarden A. 2001. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA* 98:8614–19
65. Turner TE, Schnell S, Burrage K. 2004. Stochastic approaches for modelling in vivo reactions. *Comput. Biol. Chem*. 28:165–78
66. van Kampen NG. 1992. *Stochastic Processes in Physics and Chemistry*. Amsterdam: North-Holland
67. Verkman AS. 2002. Solute and macromolecule diffusion in cellular aqueous compartments. *Trends Biochem. Sci*. 27:27–33
68. Vilar JM, Kueh HY, Barkai N, Leibler S. 2002. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl. Acad. Sci. USA* 99:5988–92
69. Volfson D, Marciniak J, Blake WJ, Ostroff N, Tsimring LS, Hasty J. 2006. Origins of extrinsic variability in eukaryotic gene expression. *Nature* 439:861–64
70. Weinberger LS, Burnett JC, Toettcher JE, Arkin AP, Schaffer DV. 2005. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell* 122:169–82
71. Yu J, Xiao J, Ren XJ, Lao KQ, Xie XS. 2006. Probing gene expression in live cells, one protein molecule at a time. *Science* 311:1600–3

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

Annual Review of Biophysics and Biomolecular Structure

Volume 35, 2006

Contents

Frontispiece
Martin Karplus xii

Spinach on the Ceiling: A Theoretical Chemist's Return to Biology
Martin Karplus 1

Computer-Based Design of Novel Protein Structures
Glenn L. Butterfoss and Brian Kuhlman 49

Lessons from Lactose Permease
Lan Guan and H. Ronald Kaback 67

Evolutionary Relationships and Structural Mechanisms of AAA+ Proteins
Jan P. Erzberger and James M. Berger 93

Symmetry, Form, and Shape: Guiding Principles for Robustness in Macromolecular Machines
Florence Tama and Charles L. Brooks, III 115

Fusion Pores and Fusion Machines in Ca^{2+}-Triggered Exocytosis
Meyer B. Jackson and Edwin R. Chapman 135

RNA Folding During Transcription
Tao Pan and Tobin Sosnick 161

Roles of Bilayer Material Properties in Function and Distribution of Membrane Proteins
Thomas J. McIntosh and Sidney A. Simon 177

Electron Tomography of Membrane-Bound Cellular Organelles
Terrence G. Frey, Guy A. Perkins, and Mark H. Ellisman 199

Expanding the Genetic Code
Lei Wang, Jianming Xie, and Peter G. Schultz 225

Radiolytic Protein Footprinting with Mass Spectrometry to Probe the Structure of Macromolecular Complexes
Keiji Takamoto and Mark R. Chance 251

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org
by Johns Hopkins University on 10/31/11. For personal use only.

The ESCRT Complexes: Structure and Mechanism of a Membrane-Trafficking Network
James H. Hurley and Scott D. Emr 277

Ribosome Dynamics: Insights from Atomic Structure Modeling into Cryo-Electron Microscopy Maps
Kakoli Mitra and Joachim Frank 299

NMR Techniques for Very Large Proteins and RNAs in Solution
Andreas G. Tzakos, Christy R.R. Grace, Peter J. Lukavsky, and Roland Riek 319

Single-Molecule Analysis of RNA Polymerase Transcription
Lu Bai, Thomas J. Santangelo, and Michelle D. Wang 343

Quantitative Fluorescent Speckle Microscopy of Cytoskeleton Dynamics
Gaudenz Danuser and Clare M. Waterman-Storer 361

Water Mediation in Protein Folding and Molecular Recognition
Yaakov Levy and José N. Onuchic 389

Continuous Membrane-Cytoskeleton Adhesion Requires Continuous Accommodation to Lipid and Cytoskeleton Dynamics
Michael P. Sheetz, Julia E. Sable, and Hans-Günther Döbereiner 417

Cryo-Electron Microscopy of Spliceosomal Components
Holger Stark and Reinhard Lührmann 435

Mechanotransduction Involving Multimodular Proteins: Converting Force into Biochemical Signals
Viola Vogel 459

INDEX

Subject Index 489

Cumulative Index of Contributing Authors, Volumes 31–35 509

Cumulative Index of Chapter Titles, Volumes 31–35 512

ERRATA

An online log of corrections to *Annual Review of Biophysics and Biomolecular Structure* chapters (if any, 1997 to the present) may be found at http://biophys.annualreviews.org/errata.shtml

Annu. Rev. Biophys. Biomol. Struct. 2007.36:413-434. Downloaded from www.annualreviews.org by Johns Hopkins University on 10/31/11. For personal use only.