

Translating the noise

Jeff Hasty & James J. Collins

Center for BioDynamics and Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215, USA.
e-mail: hasty@bu.edu and jcollins@bu.edu

A systematic exploration of noise in gene expression demonstrates the value of integrating novel experiments with computational modeling.

The fundamental importance of cellular protein levels would seem to dictate that transcription and translation operate with clockwork-like regularity. Natural selection should prune any 'sloppiness' occurring at the genetic level. As a consequence, the observed reliability of gene regulation should stem from an underlying determinism. On the other hand, an optimally selected regulatory system could still show significant random variation. Theoretically, with the number of promoter sites and mRNA molecules of a given type typically numbering fewer than ten per cell, the statistical physics implies that large fluctuations must exist. That is, as the biochemical rates of transcription and translation are proportional to the number of promoter sites and mRNA molecules, the small number of molecules implies relatively infrequent transcriptional and translational events, leading to large fluctuations in the number of mRNA and protein molecules.

Population measurements typically show that the level of expression from the same gene varies significantly from one cell to another within a genetically identical colony. Such variations are routinely observed in the cells of organisms ranging in complexity from bacteria to mammals, yet there has been little work aimed at deducing the sources of these variations. On page 69 of this issue, Ertugrul Ozbudak and colleagues¹ investigate how variations in gene expression arise from molecular fluctuations whose origin is the stochastic nature of the underlying biochemical reactions.

Lost in translation

In a series of experiments on a single-gene system, Ozbudak *et al.*¹ measured the amount of the random variation observed in the fluorescence distributions of a reporter protein as a function of the transcriptional and translational efficiencies. At the transcriptional level, they changed the efficiency both by introducing mutations in the promoter sequence and by using an inducible promoter. In the latter



Gene noise

case, they inserted a *Lac* operator site into the P_{spac} promoter so that *Lac* tetramers would repress transcription. Isopropyl- β -D-thiogalactopyranoside (IPTG) was then used to render *Lac* proteins unable to bind to the operator and repress transcription, thereby increasing the transcriptional efficiency with increasing amounts of IPTG. The translational efficiency was varied by constructing a battery of mutants that produced mRNAs with differing affinities for ribosome binding sites. By independently varying the transcriptional and translational efficiencies, Ozbudak *et al.*¹ found that the stochastic variation in protein levels is predominately generated at the translational level.

Theory and experiment

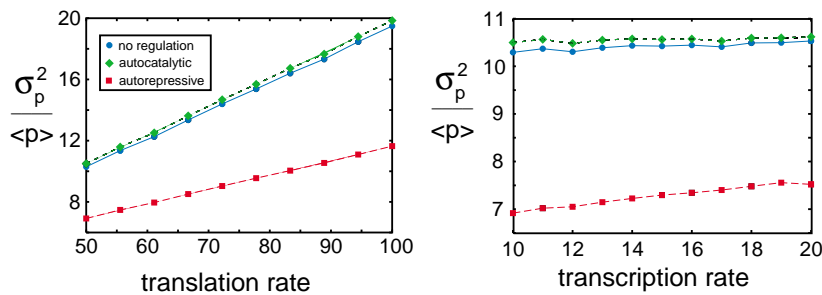
An important aspect of this study is its relationship to theoretical results generated from a biochemical model. Whereas most biological modeling studies attempt to describe a given set of experimental data post hoc, this study differs in that the experiments were designed in light of theoretical predictions reported earlier by the same group². In that modeling work, it was

predicted that the random variation in expression from a single gene should scale linearly with the translational rate and be independent of the transcriptional rate (see figure on next page). The experimental results in the present study are consistent with this prediction. The Fano factor, defined as the ratio of variance to mean, is observed to increase linearly with the translational efficiency and shows only a very mild increase with the transcriptional efficiency. As in the theoretical model, the source of the fluctuations is the small number of reactant molecules. The two studies taken together strongly support the conclusion that the origin of the noise is molecular fluctuations arising from the small number of reactant molecules.

Although the experimental findings are generally consistent with the theoretical predictions, one intriguing discrepancy is that the Fano factor is observed to increase modestly with the transcriptional efficiency. As this result is not expected to arise from fluctuations originating from the small number of molecules, a question arises as to the source of this mild, yet discernible, increase. A potential answer can be found in recent modeling work on the origins of biochemical noise in gene regulation³. This work showed that an additional source of noise can arise from the finite lifetime of an operator-inducer complex. In the context of the present study, this effect would originate from a relatively long half-life for the *Lac*-DNA complex. That is, if a *Lac* tetramer binds to its operator site in the P_{spac} promoter for a significant amount of time, then these 'operator fluctuations' become important, because the off-on rate is such that the effect cannot be averaged away. Notably, previous results³ predict an increase in random variation as the transcriptional rate is increased, consistent with the increase observed in the present study.

Gene regulation

The quantification of the consequences of the two-step process of transcription and translation in an unregulated single-gene network opens the door for the



Model results of unregulated and regulated single-gene networks. The Fano factor (the value on the y axis) is a measure of the degree of variability in the distribution of the number of intracellular proteins taken over a population of cells. Autocatalytic activation of the promoter by a factor of 100 has little effect on the Fano factor as compared with the unregulated system. In the autorepressive case, variations are suppressed^{2,4,5}. Activation results when a protein dimer docks at the promoter region and amplifies the transcriptional rate, whereas repression occurs when dimer binding inhibits transcription. The results are for a fluorescent reporter that is translated along with the regulatory protein. Models used in generating these figures are analogous to those used by Ozbudak *et al.*¹, and thus incorporate small molecular numbers as the source of variability.

exploration of the source of noise in a network where transcription factors can regulate expression. Early modeling work⁴ explored the effect of autoregulation on equilibrium stability, comparing networks with varying degrees of repression with those of varying activation. When a gene product is initially in an equilibrium state, small perturbations to the system will be followed by an exponential decay or rise back to equilibrium. The stability of the equilibrium state is related to the time constant in the exponential response; high stability corresponds to a fast return to equilibrium, whereas low stability correlates with a slow return. The stability was found to increase with repression and decrease with activation⁴. In the context of noise, there is an inverse relationship between stability and variability in which increased stability corresponds to

decreased variability and vice versa. This dictates that autorepressive circuits should be less noisy and autocatalytic circuits more so.

Both a negatively controlled and an unregulated promoter have been used to examine the effect of repression on variations in cellular protein concentrations. Within this framework, it was found that increased stability indeed corresponded to tighter distributions (that is, less variability). The present work by Ozbudak *et al.*¹ sets the stage for an investigation of fluctuations arising at both the transcriptional and translational levels in a regulated system. As in the unregulated case, analogous predictions have been generated (see figure) and reported by the same research group², and it will be interesting to see how these predictions fare in the laboratory.

Does noise matter?

Given the importance of genetic regulation, the very existence of such large fluctuations in protein levels seems counterintuitive. How can important, genetically driven processes such as circadian oscillators and environmentally induced switches reliably operate in such a noisy environment? The work by Ozbudak *et al.*¹ provides an experimental basis for several recent theoretical studies of this issue. For example, in the context of circadian rhythms, it has been suggested that the architecture of a given gene network has evolved to minimize the effects of noise⁶. In addition, it has been shown theoretically that switches that are stable for as long as years and switchable in as little as milliseconds can be built from fewer than a hundred molecules⁷. In other words, it may be that the large fluctuations are not that crucial to network behavior. Another proposition is that genetic noise can actually be used by an organism in deciding between alternative 'states,' such as a particular developmental pathway⁸. A definitive resolution to these issues awaits systematic experimental and modeling efforts such as those undertaken in the present study. □

- Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D. & van Oudenaarden, A. *Nature Genet.* **31**, 69–73 (2002).
- Thattai, M. & van Oudenaarden, A. *Proc. Natl Acad. Sci. USA* **98**, 8614–8619 (2001).
- Kepler, T. & Elston, T. *Biophys. J.* **81**, 3116–3136 (2001).
- Savageau, M.A. *Nature* **252**, 546–549 (1974).
- Becskei, A. & Serrano, L. *Nature* **405**, 590–593 (2000).
- Barkai, N. & Leibler, S. *Nature* **403**, 267–268 (2000).
- Bialek, W. in *Advances in Neural Information Processing Systems 13* (MIT Press, Cambridge, Massachusetts, 2001).
- Arkin, A., Ross, J. & McAdams, H.H. *Genetics* **149**, 1633–1648 (1998).